FOR THE
IB DIPLOMA
PROGRAMME

IN COOPERATION WITH ib

Economics

Paul Hoang
Sean Wray
Tanusankar Chakraborty

HODDER EDUCATION

DYNAMIC LEARNING

Dynamic Learning is an online subscription solution that supports teachers and students with quality content and unique tools. Dynamic Learning includes Teaching and Learning Resources, Whiteboard eTextbooks and Student eTextbooks.

Economics for the IB Diploma is also available as a **Whiteboard eTextbook** which is ideal for lesson planning.

Whiteboard eTextbooks are zoomable, flickable, searchable pages of traditional printed textbooks that enable teachers to:

- display pages to their class
- export sections of the book to their VLE
- add notes and highlight areas
- bookmark key pages.

Additionally the **Student eTextbook** of ***Economics for the IB Diploma*** is a downloadable version of the printed textbook that teachers can assign to students so they can:

- download and view on any device or browser
- add, edit and synchronise notes across two devices
- access their personal copy on the move.

Find out more and sign up for free trials – visit: **www.hoddereducation.com/dynamiclearning**

FOR THE
IB DIPLOMA
PROGRAMME

Economics

Paul Hoang
Sean Wray
Tanusankar Chakraborty

HODDER
EDUCATION
AN HACHETTE UK COMPANY

Dedication

Paul Hoang – *Dedicated to Christine Rowlands and Greg Thornton, my amigos*.

Sean Wray – *To the teachers, my colleagues and friends at the English Schools Foundation, Hong Kong.*

Tanu Chakraborty – *This book is dedicated to my parents, Bimal and Kaberi, my wife, Monami, and my daughter, Toya.*

The authors would also like to thank So-Shan Au, Kelvin Knight and Catherine Perks from Hodder Education for their endless hours of support, advice and encouragement. It really would not have been possible to have completed the project in the time frame we had without them.

Acknowledgements

Every effort has been made to trace all copyright holders, but if any have been inadvertently overlooked, the Publishers will be pleased to make the necessary arrangements at the first opportunity.

Although every effort has been made to ensure that website addresses are correct at time of going to press, Hodder Education cannot be held responsible for the content of any website mentioned in this book. It is sometimes possible to find a relocated web page by typing in the address of the home page for a website in the URL window of your browser.

Hachette UK's policy is to use papers that are natural, renewable and recyclable products and made from wood grown in well-managed forests and other controlled sources. The logging and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

Orders: please contact Bookpoint Ltd, 130 Park Drive, Milton Park, Abingdon, Oxon OX14 4SE. Telephone: +44 (0)1235 827827. Fax: +44 (0)1235 400401. Email education@bookpoint.co.uk Lines are open from 9 a.m. to 5 p.m., Monday to Saturday, with a 24-hour message answering service. You can also order through our website: www.hoddereducation.co.uk

ISBN: 978 1 5104 7914 2

© Paul Hoang, Sean Wray and Tanusankar Chakraborty 2020

First published in 2020 by
Hodder Education,
An Hachette UK Company
Carmelite House
50 Victoria Embankment
London EC4Y 0DZ

www.hoddereducation.co.uk

Impression number 10 9 8 7 6 5 4 3 2 1

Year 2024 2023 2022 2021 2020

All rights reserved. Apart from any use permitted under UK copyright law, no part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or held within any information storage and retrieval system, without permission in writing from the publisher or under licence from the Copyright Licensing Agency Limited. Further details of such licences (for reprographic reproduction) may be obtained from the Copyright Licensing Agency Limited, www.cla.co.uk

Cover photo © pierrick – stock.adobe.com

Illustrations by Integra Software Services Pvt. Ltd., Pondicherry, India

Typeset in 10/14pt Berkeley Oldstyle by Integra Software Services Ltd.

Printed in Slovenia

A catalogue record for this title is available from the British Library.

FSC www.fsc.org MIX Paper from responsible sources FSC™ C104740

Contents

Introduction v

Introduction to economics 1

CHAPTER 1 What is economics? 2

CHAPTER 2 How do economists approach the world? 28

Microeconomics 45

CHAPTER 3 Demand 46

CHAPTER 4 Supply 56

CHAPTER 5 Competitive market equilibrium 67

CHAPTER 6 Critique of the maximizing behaviour of consumers and producers (HL only) 77

CHAPTER 7 Elasticity of demand – price elasticity of demand (PED) . . . 92

CHAPTER 8 Elasticity of demand – income elasticity of demand (YED) 104

CHAPTER 9 Elasticity of supply 111

CHAPTER 10 Role of government in microeconomics 121

CHAPTER 11 Market failure – externalities and common pool (common access) resources 140

CHAPTER 12 Market failure – public goods 165

CHAPTER 13 Market failure – asymmetric information (HL only) 171

CHAPTER 14 Market failure – market power (HL only) 178

CHAPTER 15 The market's inability to achieve equity (HL only) 217

Macroeconomics 225

CHAPTER 16 Measuring economic activity and illustrating its variations 226

CHAPTER 17 Variations in economic activity: aggregate demand and aggregate supply 243

CHAPTER 18 Macroeconomic objectives – economic growth 264

CHAPTER 19 Macroeconomic objectives – low unemployment 275

CHAPTER 20 Macroeconomic objectives – low and stable rate of inflation 289

CHAPTER 21 Macroeconomic objectives – sustainable level of government (national debt) (HL only) 310
CHAPTER 22 Macroeconomic objectives – potential conflict between macroeconomic objectives 316
CHAPTER 23 Economics of inequality and poverty 325
CHAPTER 24 Demand management (demand-side policies) – monetary policy 364
CHAPTER 25 Demand management – fiscal policy 380
CHAPTER 26 Supply-side policies 401

The global economy 419

CHAPTER 27 Benefits of international trade 420
CHAPTER 28 Types of trade protection 436
CHAPTER 29 Arguments for and against trade controls/protection . . . 455
CHAPTER 30 Economic integration 469
CHAPTER 31 Exchange rates 485
CHAPTER 32 Balance of payments 506
CHAPTER 33 Sustainable development 527
CHAPTER 34 Measuring development 542
CHAPTER 35 Barriers to growth and/or economic development 564
CHAPTER 36 Economic growth and/or economic development strategies 590

Glossary 636

Index 656

Acknowledgements 662

Introduction

This textbook is written for students and teachers of IB DP Economics, or anyone with a general interest in economics. It provides readers with a clear and detailed introduction to economics, with real-world applications to engage learners. The 'In co-operation with IB' logo signifies the content in this textbook has been reviewed by the IB to ensure it fully aligns with the current IB curriculum and offers high-quality guidance and support for IB teaching and learning.

The textbook provides complete coverage of the new IB DP Economics syllabus (first teaching in August 2020). Differentiated content for SL and HL students is clearly identified. As the study of economics relies on the use of diagrams to facilitate analysis and discussion, this textbook also includes the use of clearly constructed and explained diagrams and visual graphs. Note: all currency values are expressed in US dollars ($) unless otherwise stated. Other features of this book to help students to consolidate and develop their understanding of economics include:

SYLLABUS CONTENT

The textbook follows the exact order of the contents of the IB Economics syllabus. At the beginning of each chapter is a list of the content to be covered, with all subsections clearly linked to the Assessment Objectives (AO), and showing the breadth and depth of teaching required. For example, AO4 skills highlight the specific sections of the syllabus that require students to make calculations or to construct relevant economic diagrams.

CASE STUDY

Real-world international examples and case studies are used to bring economics to life. Real-life examples are important to help students understand economics in the real world. These case studies provide context to improve understanding of the IB Economics syllabus content, such as providing context to explain why monopolists really do exist in the real world.

Real-world issue

Each chapter starts with one of the six real-world issues from the IB Economics syllabus (see page 7 of the syllabus guide). The course allows students to explore models and theories used in Economics, and to apply them using empirical data, through the examination of these real-world issues, which are split as follows:

Microeconomics	• Real-world issue 1 – How do consumers and producers make choices in trying to meet their economic objectives? • Real-world issue 2 – When are markets unable to satisfy important economic objectives and does government intervention help?
Macroeconomics	• Real-world issue 3 – Why does economic activity vary over time and why does this matter? • Real-world issue 4 – How do governments manage their economy and how effective are their policies?
The global economy	• Real-world issue 5 – Who are the winners and losers of the integration of the world's economies? • Real-world issue 6 – Why is economic development uneven?

ATL ACTIVITY

Approaches to learning, including learning through inquiry, are integral to IB pedagogy. These ATL skills activities get students to critically think about real-world economics such as how and why large airline companies avoid price rigidity.

Common mistake

These boxes contain advice regarding some of the common misunderstandings and typical errors made by students so that you can avoid making such mistakes.

◆ **Key terms –** There are over 500 key term definitions that appear throughout the textbook, in the relevant section, to help students grasp the language of economics. There is also a glossary of all key terms at the back of the book.

Chapter summary

At the end of each chapter there is a summary of the key points addressed to allow you to consolidate your learning.

Top tip!

These boxes include advice relating to the content being discussed and tips to help you retain the knowledge you need, including some useful acronyms. Read on to find out what these (and many others) mean: CELL (Chapter 1), PIGLETS (Chapter 11), BLADE (Chapter 14), BUDGIE (Chapter 23), EPIC (Chapter 25), CUPEL (Chapter 30), and GIST (Chapter 31).

WORKED EXAMPLE

These provide a step-by-step guide showing you how to answer the kind of questions that you might encounter in your studies.

PRACTICE QUESTION

Practice questions are included to allow you to check your understanding and prepare for the assessments. Some questions are in the style of exam questions – the particular paper they relate to is indicated in each case.

Theory of Knowledge

Links to Theory of Knowledge allow you to develop critical thinking skills and deepen economic understanding by bringing discussions about the subject beyond the scope of the content of the curriculum.

WISE ChoICES

Prompts and questions are provided in every chapter to challenge teachers and students to take on an authentic approach to concept-based learning and teaching through inquiry. While these questions are not explicitly asked in the final assessments, they promote discussion and develop critical thinking skills. After all, the teaching of economics is conceptually focused and grounded in real-world issues. The WISE ChoICES questions that appear throughout this textbook extend learning beyond the syllabus while remaining in line with IB pedagogy and enabling students to look at IB Economics from multiple perspectives and to deepen their level of understanding.

The nine key concepts that underpin the new economics course are: Scarcity, Choice, Efficiency, Equity, Economic well-being, Sustainability, Change, Interdependence and Intervention. We have developed an easy to remember acronym for the key concepts so that it is for ever at the forefront of learning: WISE ChoICES:

- **W**ell-being
- **I**nterdependence
- **S**carcity
- **E**fficiency
- **C**hoice
- **I**ntervention
- **C**hange
- **E**quity
- **S**ustainability

The WISE ChoICES **concepts** are anchored in the **content** of economics – its theories, models, ideas and tools – and are given **context** through the study of real-world issues and examples. Connections are made and context is given to the concepts in our WISE ChoICES boxes, which appear throughout the topics and chapters.

Assessment guidance

Guidance relating to each of the three exam papers, the internal assessment and the extended essay can be accessed at hoddereducation.com/ib-extras. From there, you will need to navigate to the heading for Economics and find the document titled Assessment guidance.

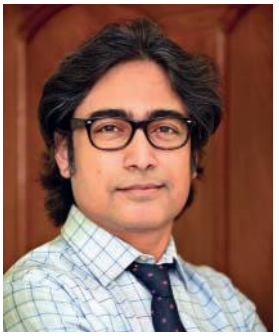
■ Visit the Hodder Education website to access advice for IB Economics assessment

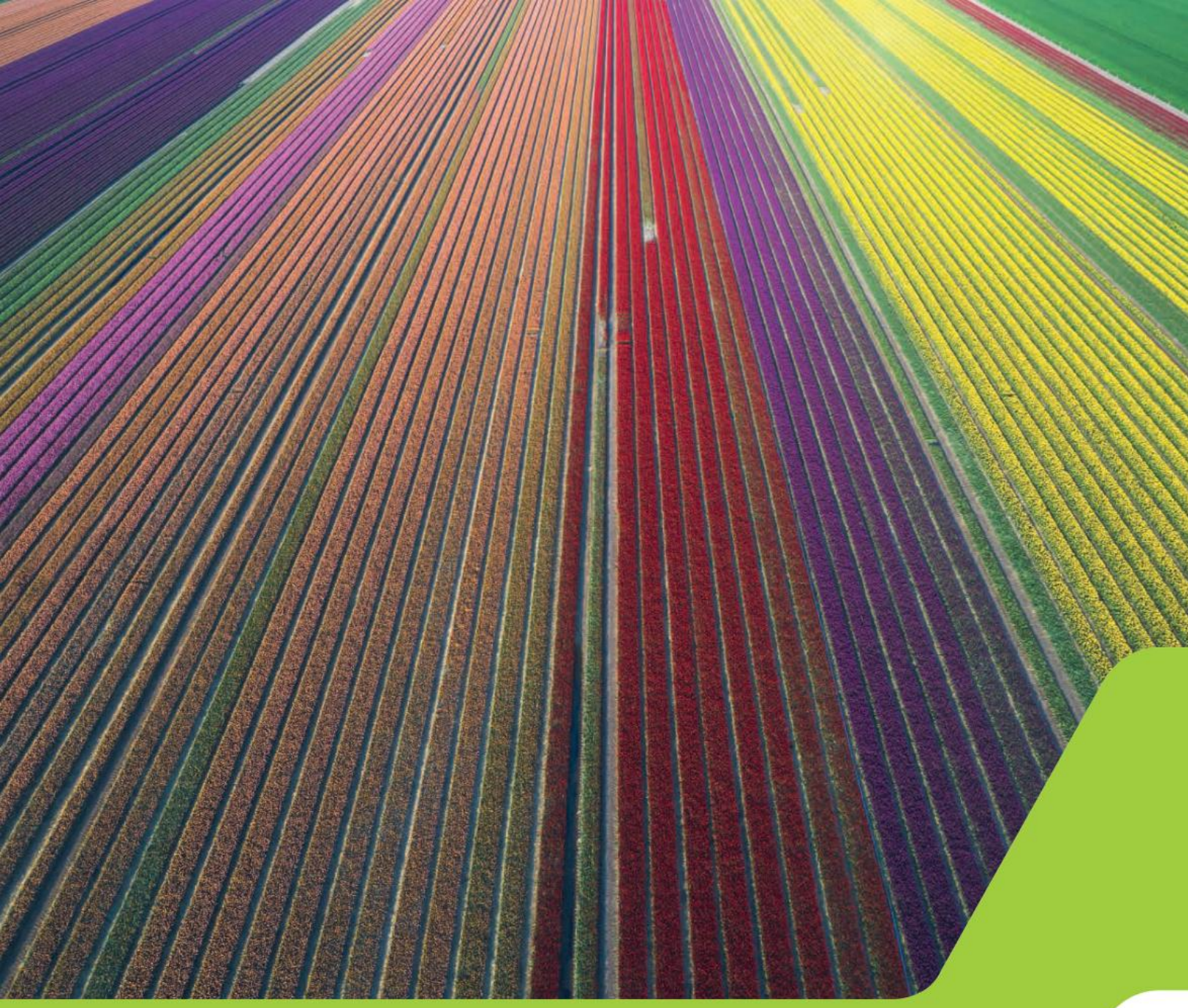
About the authors

Paul Hoang is a highly experienced educator having taught Economics for over 25 years. He has held the positions of Vice Principal (Curriculum), IB Diploma Programme Coordinator, and Head of Faculty at Sha Tin College, an internationally renowned IB World School in Hong Kong. Paul is the author of several best-selling publications for IB and IGCSE Economics and the Extended Essay (Hodder Education). He has experience as an examiner and workshop leader for various examination boards. Paul is also Educational Partner Consultant at *IB School Improvement*, working with school leaders to improve teaching and learning with specific focus on the use of assessment data. Outside of work, Paul enjoys cinema, football, Tae Kwon Do, and the outdoors. He is a life-long Arsenal fan but acknowledges that no one is perfect.

Sean Wray teaches Economics and is a Head of House at the English Schools Foundation's King George V School, Hong Kong. He grew up in Western Australia and Hong Kong, and attended secondary school in both places. Sean graduated from The University of Manchester, United Kingdom, with a BA (Hons) in Modern History with Economics. After working in finance in Hong Kong, he studied for his PGDE at The University of Hong Kong. He is the current ESF Curriculum Group Leader for Business Management and Economics and is also the current Chair of the ESF Committee of Teachers. He enjoys leading collaboration on teaching and learning through both of these wider school roles. Sean has also written regularly for IB Review, the flagship magazine publication from Hodder Education. Outside of the classroom, Sean enjoys coaching rugby and he is an enthusiastic supporter of the West Coast Eagles in the Australian Rules Football competition.

Tanusankar Chakraborty hails from a small university town, about 50 km away from Kolkata, India. He has taught IB Economics and Business Management since 2005 at leading international schools in India, UAE and China. He has held the position of Department Head, Model United Nations Director, Extended Essay Coordinator and Diploma Programme Coordinator in different IB World Schools. Currently, he teaches IB Economics and Global Politics at the Western Academy of Beijing. He also teaches DP Business Management for Pamoja Education, an IB approved online provider, and has a wealth of experience as an examiner and workshop leader. He is interested in developing a genuine concern for others among his students, through shared life experiences.





Introduction to economics

1 What is economics?

Conceptual understandings (WISE ChoICES)

- Economics is a social science characterized by **interdependence**, which focuses on how people interact with each other to improve their **economic well-being**.
- The economic world is dynamic in nature and constantly subject to **change**.
- The central problems of economics are **scarcity** and **choice**, which forces societies to face opportunity costs and the challenge of **sustainability**.
- Debates exist in economics regarding the potential conflicts between economic growth and **equity** and between free markets and government **intervention**.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- economics as a social science: the social nature of economics, the basis of the study of economics (microeconomics and macroeconomics) and introduction to the nine central concepts (WISE ChoICES) (AO2)
- the problem of choice: factors of production, scarcity, unlimited human needs and wants to be met by limited resources, and scarcity and sustainability (AO2)
- opportunity cost and free goods (AO2)
- the basic economic questions and the means of answering the economic questions (AO2)
- market versus government intervention and economic systems (AO2)
- the production possibilities curve model (PPC) (AO2)
- a PPC diagram illustrating choice and opportunity cost, unemployment of resources, actual growth, growth in production possibilities and increasing versus constant opportunity cost (AO4)
- modelling of the economy through the circular flow of income model (AO2)
- a circular flow of income diagram with leakages and injections (AO4).

Economics as a social science (AO2)

Economics is the study of how resources are allocated to satisfy the unlimited needs and wants of individuals, governments and firms in an economy. Economics is considered to be a **social science** because it examines the diverse social behaviours of individuals and societies in the allocation of scarce resources to meet the endless needs and wants of **economic agents** (decision makers in the economy, comprising consumers, employees, firms and governments). The choices made by economic agents can generate positive or negative outcomes, thereby affecting the relative well-being of individuals and societies. As a social science, economics demonstrates an appreciation of the impacts of economic activity on the behaviour and well-being of individuals, firms and governments. Examples include:

- the problem of choice (Chapter 1)
- behavioural economics (Chapter 6)
- the role of government in economics (Chapter 10)

◆ **Economics** is an academic discipline under the category of social sciences, examining how resource allocation can best meet the needs and wants of individuals and societies.

◆ **Social science** is the study of human interactions and relationships between individuals and societies.

◆ **Economic agents** are economic decision makers who interact for the purpose of production and consumption. They comprise households, firms and the government.

Figure 1.1 Economics is about resource allocation and making wise choices

- market failure (Chapters 11–14)
- statistics to measure economic well-being (Chapter 16)
- economics of inequality and poverty (Chapter 23)
- sustainable development (Chapter 33).

The social nature of economics

As a social science, economics examines the choices of consumers, producers and governments by using appropriate models and theories. While economics is often associated with facts and empirical evidence, it also examines the different social behaviours of individuals, firms and nations in different situations. One of the goals of the IB Economics course is for students to develop an understanding of the complexities and interdependence of economic activities in a continually and rapidly changing world.

The study of economics examines the prescribed **WISE ChoICES** concepts at three levels:

- at the level of households and firms (or producers and consumers) in individual markets – studied under the topic of **microeconomics**
- at the level of the government and the economy as a whole – studied under the topic of **macroeconomics**
- at the level of international governments and the growing interdependence between countries through freer and fairer international trade and the international movement of labour and capital – studied under the topic of the **global economy**.

As a social science, economics enables IB students to examine models and theories using the following six **real-world issues**:

1. How do consumers and producers make choices in trying to meet their economic objectives?
2. When are markets unable to satisfy important economic objectives and does government intervention help?
3. Why does economic activity vary over time and why does this matter?
4. How do governments manage their economy and how effective are their policies?
5. Who are the winners and losers of the integration of the world's economies?
6. Why is economic development uneven?

Source: IB Diploma Programme Economics guide, page 7

Top tip!

The IB Economics course requires you to focus on the prescribed six real-world issues by using the nine key concepts (see **WISE ChoICES** on the left). This enables students to develop the knowledge, skills, values and attitudes needed to become responsible and global citizens.

Figure 1.2 Economists study the reasons for unequal wealth distribution

Theory of Knowledge

Are there fundamental differences between economics and other disciplines or areas of knowledge? If so, are these differences more than just methodological?

The basis of the study of economics: microeconomics and macroeconomics

As core aspects of economics, both microeconomics and macroeconomics focus on the allocation of scarce resources in an efficient way. Resources are scarce because they are limited in supply. Both disciplines of economics are concerned with the notion of maximization: microeconomics is about profit maximization for individual firms and welfare maximization for consumers, while macroeconomics is about maximizing the economic goals of society as a whole, such as national income and economic growth.

Microeconomics is a category of economics concerned with the behaviour of individuals and firms in distinct markets and segments of the economy, rather than the operations of the economy as a whole. For example, microeconomics examines how prices are determined in particular markets and the determination of how best to allocate scarce resources given their different alternative uses. Much of microeconomics assumes rational economic behaviour and the decisions of individual consumers and producers, which would ultimately lead to scarce resources being allocated efficiently.

By contrast, **macroeconomics** examines the operations of the economy as a whole. It looks at aggregate variables, such as national output. For example, this will include looking at the causes and consequences of short run fluctuations in economic activity (known as the business cycle) and what determines long run economic growth and development. Macroeconomics considers how the economy is affected by unemployment (people who are willing to work but are unable to find work), inflation (an increase in the general price level) and the distribution of income and wealth. Hence, macroeconomics deals with broad economic issues and topics, based on the decisions made by governments and countries.

◆ **Microeconomics** is the area of economics focused on the actions of individuals and firms in particular markets rather than the operations of the economy as a whole.

◆ **Macroeconomics** is the area of economics focused on the operations of the economy as a whole, rather than the behaviour of economic agents in individual markets.

Figure 1.3 Micro and macroeconomics are core aspects in the study of economics

There is far less debate about the basic principles of microeconomics, such as the law of demand (see Chapter 3) or how changes in the level of real income affect (people's income after adjusting for inflation or changes in the general price level) consumer demand for goods and services (see Chapter 8). By contrast, there is far less consensus when studying macroeconomics, with its different schools of thought, such as Keynesians and monetarists (see Chapter 2). In addition, macroeconomics has greater emphasis on empirical data or key economic indicators, such as national income statistics (see Chapter 16), inflation and unemployment (both covered in Chapter 19). In addition, microeconomics examines the actions and decisions of economic agents in individual markets as part of the domestic economy, whereas macroeconomics considers outcomes on national, regional and global scales.

Introduction to the nine central concepts (WISE ChoICES)

Figure 1.4 WISE ChoICES

As part of the IB's overarching pedagogy of concept-based learning (CBL), the economics syllabus incorporates the use of nine concepts: scarcity, choice, efficiency, equity, economic well-being, sustainability, change, interdependence and intervention. CBL, as defined by education consultant Dr H Lynn Erickson, is a three-dimensional model that frames **facts** (content) and **skills** with **concepts** of the IB Economics course. The rigours of the economics course create two approaches to teaching and learning: either students can be told what they need to know (based on factual knowledge and skills in traditional economics) or teachers can get their students excited enough to discover and deepen their understanding of economics for themselves. The IB Economics syllabus (first examinations 2022) emphasizes the teaching of economics being conceptually focused, and grounded in real-world issues. Essentially, a conceptually focused approach to teaching and learning is core to the discipline of IB Economics. It is also explicitly examined in the Internal Assessments (for both SL and HL students).

The authors and publisher have constructed a mnemonic that will help you to think about concepts in the study of IB Economics: **WISE ChoICES** (Figure 1.5).

ATL ACTIVITY 1.1

Investigate how and why concept based learning (CBL) may be deemed to be a better way for you to understand every day economic problems and real-world issues.

Theory of Knowledge

Economic theory suggests that the material well-being of individuals and societies is directly and best related to the quantity of goods and services available to people. To what extent is economic well-being a true indication of people's psychological well-being? How do consumers make choices in trying to meet their economic objectives?

Well-being – economic well-being is about the level of economic prosperity and quality of life (standards of living) in the economy. Economic agents interact with each other in order to improve their economic well-being. The choices made by economic agents have both positive and economic consequences on individuals and societies, thus affecting their relative economic well-being.

↓

Interdependence – this refers to the growing interaction and reliance on others in order to achieve economic goals, because individuals and societies are not self-sufficient in a rapidly changing world. As economics is the study of how individuals and societies work together to satisfy their unlimited wants with the world's limited resources, interdependency is an essential concept in the subject. Globalization has led to a highly interdependent economic world, so decisions made in one part of the world have economic consequences for individuals and firms in other locations.

↓

Scarcity – this concept is central to the study of economics, namely that resources are scarce (finite or limited in supply) relative to the infinite (unlimited) needs and wants of individuals and societies. It means there is an inadequate amount or shortage of resources in the economy. Hence, rational economic choices are made when addressing the basic economic questions of what, how, and for whom production should take place.

↓

Efficiency – this refers to how well things are done, determined by the input–output ratio, i.e. the input of scarce factor resources to generate the output of goods and services in order to meet the infinite needs and wants of individuals and societies. Hence, efficiency is a quantifiable concept in economics. Efficiency generates a socially optimum level of output of goods and services in the economy, thus improving economic well-being.

↓

Choice – economics is essentially the study of choice due to finite resources and infinite wants of individuals and societies. Hence, economic agents have to make choices, thus giving rise to opportunity costs. Economists study the opportunity costs and consequences of choice at the microeconomic, macroeconomic and global level. Essentially, economics is about making wise choices, due to the central concept of scarcity. In general, these choices are made in an efficient and sustainable way.

↓

Intervention – this refers to the roles and responsibilities of governments in terms of monitoring and regulating the behaviour of the workings of different markets in the private sector. It involves public sector involvement when markets fail to achieve an efficient allocation of scare resources (from society's point of view), thus jeopardizing economic well-being, efficiency, equity and sustainability.

↓

Change – economic change is inevitable, and the economic world is continually evolving, so economic agents need to be aware of change and thus adapt their thinking. In this discipline, change can be brought about by institutional, structural, technological, economic and social factors. Hence, an understanding of change as a concept is essential to the understanding of the subject.

↓

Equity – this concept is about the idea of perceived fairness. The very nature of fairness is subjective, so is an aspect of normative economics, because what is deemed fair by an individual, firm or government might not necessarily be the case for other stakeholders. It is about the challenges faced by individuals and societies in terms of fair access to the world's scarce resources. As a normative concept, equity means different things to different individuals and societies but enables economists to explore significant global issues.

↓

Sustainability – from the perspective of economists, sustainability is about meeting the needs and wants of the current generation without jeopardizing those of future generations. It is about the ways in which economic activity impacts the natural environment and the world's scarce resources. Hence, sustainability is about intergenerational equity. As we strive for economic well-being and prosperity, much damage has been done to the Earth's scarce resources, and hence, sustainability is of growing importance to the global economy.

■ **Figure 1.5** WISE ChoiCES

Concept-based learning in IB Economics uses the **triangle model**, which consists of the **content** of the syllabus, integrated with real-word issues to provide a **context** to the learning, and is embedded in **concepts** to transcend understanding across all aspects of the economics course. An example of a classroom activity that embraces these concepts (as building blocks for learning) might require you to to investigate how macroeconomic policy is influenced by WISE ChoICES: economic well-being, interdependence, scarcity, efficiency, choice, intervention, change, equity and sustainability for a country of your choice.

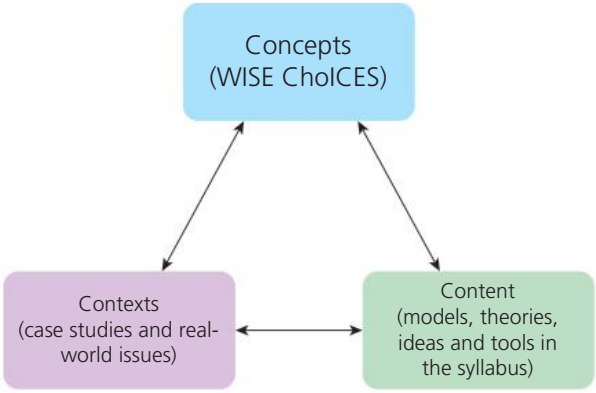


Figure 1.6 The triangle model in IB Economics

Concept-based learning is *not* about learning facts and figures in isolation. Instead, CBL is about encouraging students to:

- think beyond the content (tools, theories and models) specified in the IB Economics syllabus
- be active, curious and open-minded rather than passive learners
- be inquirers who want to further their understanding and knowledge using real-world issues that they are interested in
- make connections of knowledge in economics to the WISE ChoICES concepts (**W**ell-being, **I**nterdependence, **S**carcity, **E**fficiency, **Cho**ice, **I**ntervention, **C**hange, **E**quity and **S**ustainability)
- recognize the transferability of knowledge and skills in economics, and apply this to both familiar and unfamiliar situations
- organize and make sense of their own learning (possibly using the WISE ChoICES framework)
- think critically and creatively to solve problems
- extend their knowledge to understanding, in a synergistic way – as Albert Einstein once said, *'Any fool can know. The point is to understand.'*

Refer to the Introduction for more details about the WISE ChoICES concepts and how these can be incorporated into the teaching and learning of economics.

Theory of Knowledge

If people 'know' that smoking is bad for their health or that they should exercise more, what role do *reason* and *intuition* play in the understanding of real-world economics?

ATL ACTIVITY 1.2

Many hotels and restaurants offer all-you-can-eat buffets. With a partner, decide on the three most applicable WISE ChoICES concepts that apply to such economic activity. Be prepared to share your ideas and discussion points with the rest of the class.

Figure 1.7 Which concepts are best aligned with buffets at hotels and restaurants?

The problem of choice (AO2)

Central to the study of economics is the concept and problem of scarcity. In every country, resources are limited in supply so decisions have to be made by individuals, firms (businesses) and governments about how to allocate scarce resources to satisfy their unlimited needs and

wants. As a result of scarcity, **choices** have to be made. This is known as the **basic economic problem**, which exists in every economy: how best to allocate scarce resources in order to satisfy people's infinite needs and wants.

◆ The **basic economic problem** refers to the issue of how best to allocate an economy's scarce resources in order to satisfy the unlimited needs and wants of individuals, firms and governments.

Needs are the necessities people must have in order to survive. This includes: food, water, shelter, clothing, warmth and sleep. Wants are desires, that is, things people would like to have or have more of. Wants are infinite (unlimited). There are several reasons for this, including:

- It is human nature to want more things; once people have something, they want something else too.
- Marketers use advertising and promotional techniques to inform and persuade customers to buy more.
- Peer pressure and societal pressures cause people to buy what their friends, neighbours or other acquaintances have, so they do not feel left out or left behind.
- Physical goods eventually need to be replaced or upgraded; for example, cars, mobile phones, furniture and clothes. Perishable products need to be replaced regularly; for example, food, flowers, tobacco products, medication and petrol. Services also need to be replenished; for example, haircuts, public transport and utility bills (for gas, water, telecommunications and electricity services).

However, the economic resources to fulfil unlimited wants and needs are finite (limited). This limits people's options, so forces them to make economic choices.

Factors of production

Factors of production are the four categories of resources that are required to produce any good or service: land, labour, capital and entrepreneurship.

◆ **Factors of production** are the four categories of resources that are required to produce any good or service, namely land, labour, capital and enterprise.

- **Land** (or **natural capital**) refers to the natural resources used in the production process. Examples include crude oil, coal, water, wood, metal ores and agricultural products. Land can be further categorized as renewable and non-renewable resources. **Renewable resources** are those that can be used and replaced, such as fish stocks, water, forests and plants. **Non-renewable resources** are those that once used cannot be replaced, such as oil supplies, gold reserves, fossil fuels and metal ores.
- **Labour** (or **human capital**) refers to the human resources required in the production process. This includes physical human effort and intellectual input from the workforce. Examples include the services of an accountant, barista, chef, doctor or estate agent.
- **Capital** (or **physical capital**) refers to non-natural (manufactured) products used in the production process, such as machinery, tools, equipment and vehicles. Capital is used to facilitate production. For instance, infrastructure such as transportation and communications networks are vital for the growth and development of society.
- **Enterprise** (or **entrepreneurship**) refers to the skills, creativity and risk-taking ability that a businessperson

Figure 1.8 Glass is a renewable resource

requires to successfully combine and manage the other three factors of production. The entrepreneur has two main functions: to manage the overall production process (of producing physical goods and/or providing intangible services), and to take overall responsibility for the profits or losses of the firm.

Consider the example below of the factors of production required to produce aluminium cans of Coca-Cola beverages:

- Land: the natural resources required to make Coca-Cola, for example, sugar, water, carbon dioxide and caffeine.
- Labour: people hired to work on the production lines, perform administrative tasks, promote Coca-Cola's drinks to customers and to manage the company.
- Capital: factory buildings, machinery, computers, tools, and trucks to transport the drinks to warehouses and retailers.
- Enterprise: the risk-taking and business skills necessary to successfully organize the production and distribution processes, and to motivate employees so they work to the best of their ability, in the pursuit of profit for Coca-Cola's shareholders (owners).

ATL ACTIVITY 1.3

Discuss the resources (factors of production) required to deliver an IB Economics lesson.

Figure 1.9 Factors of production are required to deliver an economics lesson

Theory of Knowledge

Karl Marx (1818–1883) claimed that labour is the most important factor of production. What do you think about this claim? What is the most 'useful' factor of production in today's world. Does the answer depend on the period of time being referred to?

Top tip!

The four factors of production can be remembered by using the acronym **CELL**: capital, enterprise, land and labour.

The rewards for the factors of production

As the factors of production are productive resources, each has a reward for its use in the production process:

- The reward for land is called **rent**. Rental income or revenue comes from the ownership of land, paid by the users of the land resources.
- The rewards for labour are **wages** and **salaries**. Wages are paid to workers on an hourly basis, such as those who earn a **national minimum wage**. Salaried workers are paid a fixed amount, usually paid per month.
- The reward for capital is called **interest**. If the interest rate is low, it becomes more worthwhile for firms and households to borrow money for production and consumption purposes as the cost of borrowing money falls, all other factors remaining unchanged.
- The reward for enterprise is called **profit**. This is the return for the entrepreneur's business ideas and risk taking in starting up and running the organization. Profit is what remains after all costs of production have been accounted for, including payments to the other three factors of production.

Collectively, the four rewards for the factors of production are known as **income** (see Figure 1.10).

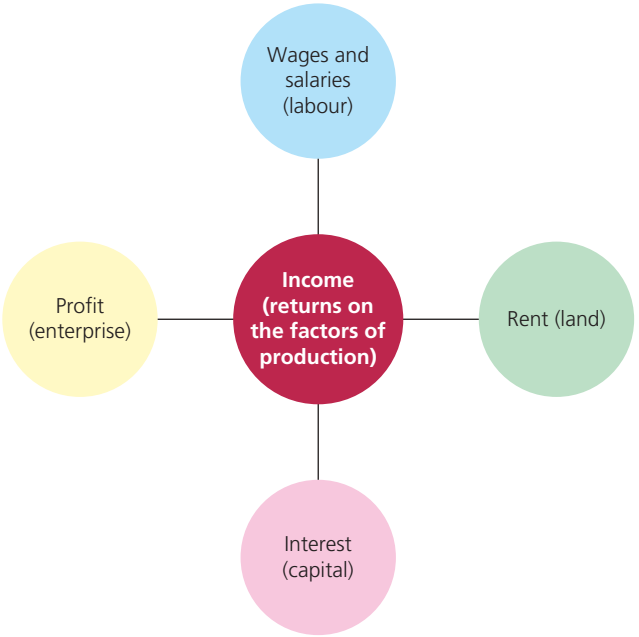


Figure 1.10 Income: the returns on the factors of production

◆ The **National Minimum Wage** (NMW) is a price floor in the labour market, requiring employers to pay at least this amount to their employees. They are not permitted to pay less than this.

◆ **Profit** is the positive difference between a firm's total revenues (TR) and its total costs (TC). It is calculated using the formula: TR – TC.

◆ **Income** is a flow concept referring to the money a person receives from the production process; for example, wages and salaries.

Common mistake

'Income refers to the money that labour receives for work.'

Students often refer to income as the money people receive from work. This is inaccurate because the reward for labour services is called 'wages' (although it can include 'salaries' too). The term 'income' refers to the collective rewards for all the factors of production, that is, the sum of rent, interest, profit and wages/salaries.

Common mistake

'Land refers to the physical land used in the production process, such as farm or office land space.'

Students often refer to land as physical land used in the production process. This is incorrect because in economics, land refers to natural resources in the production process; for example, water, raw sugar and caffeine are some of the ingredients used to make Coca-Cola.

Common mistake

'Labour refers to the number of workers used in the production process.'

Labour does not refer to the number of workers used in the production process. Labour refers to both the physical and mental human resources used in the production process, so it is more than just the physical number of workers.

Common mistake

'Capital is the amount of money used in the production process.'

Capital is not the amount of money used in the production process. In economics, capital refers to all non-natural (human-made) resources used in the production process, such as machinery, tools, buildings and vehicles.

Scarcity

British economist Lionel Robbins (1932) defined economics in terms of scarcity as: *'Economics is the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses.'*

The central purpose of economic activity is the production of goods and services to satisfy the needs and wants of individuals and societies. However, the fundamental economic problem faced by all countries is scarcity. **Scarcity** means there is a shortage of resources in the economy at any moment in time. Resources are scarce (finite or limited in supply) relative to the infinite (unlimited) needs and wants of individuals and societies for goods and services. The planet's natural resources are finite in supply, so if economies use them continuously in an unsustainable way, they will eventually run out. Even free goods and natural resources can become scarce if they become over-consumed (see common pool resources in Chapter 11). Hence, rational economic choices need to be made when addressing the basic economic questions of *what*, *how*, and *for whom* production should take place.

◆ **Scarcity** refers the finite resources (limited in supply) of an economy relative to the unlimited needs and wants of individuals and societies.

Needs are the essential goods and services required for human survival. These include nutritional food, clean water, shelter (housing), protection (safety), clothing and access to healthcare and education. All individuals have a right to have these needs met as is stated in Articles 25 and 26 of the United Nations Universal Declaration of Human Rights, drafted in December 1948.

◆ **Needs** are goods and services that are essential for survival; for example, nutritional food, clean water, shelter (housing), protection (safety), clothing, and access to healthcare and education.

> *Article 25*
>
> *Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family, including food, clothing, housing and medical care and necessary social services, and the right to security in the event of unemployment, sickness, disability, widowhood, old age or other lack of livelihood in circumstances beyond his control.*
>
> *Article 26*
>
> *Everyone has the right to education. Education shall be free, at least in the elementary and fundamental stages. Elementary education shall be compulsory. Technical and professional education shall be made generally available and higher education shall be equally accessible to all on the basis of merit.*
>
> *Source: Articles 25 and 26 of the United Nations Universal Declaration of Human Rights*

Wants are goods and services that are not deemed to be necessary for survival but are human desires, that is, things we would like to have. Wants are unlimited as most people are rarely satisfied with what they have and are always striving for more. Wants are a matter of personal choice and part of human nature.

◆ **Wants** are desires for goods and services that are not necessary for human survival but are demanded to fulfil the wishes of the consumer.

Goods are tangible (physical) items that can be produced, bought and sold. Examples include furniture, clothing, toothpaste and pencils. **Services** are intangible (non-physical) items provided by individuals and firms, and paid for by customers. Examples include haircuts, bus journeys, education, concerts, telephone calls and internet access. Collectively, goods and services are called **products** or **economic goods**. These resources are limited in supply (scarce) in relation to their demand and have economic value to society, so human effort is required to obtain such products and economic agents are willing to pay for them. As there is a degree of scarcity to all economic goods, the production and consumption of such products carries an opportunity cost. For example, the production of wheat may mean less land and other resources being available for the output of cotton.

◆ **Goods** are physical (tangible) products such as furniture, cars, toothpaste and stationery.

◆ **Services** are intangible (non-physical) products such as haircuts, bus journeys, telephone calls and internet access.

◆ **Economic goods** are the resources and products that are limited in supply and have economic value (utility) to society. The vast majority of products are economic goods.

Scarcity and sustainability are essential interlinking concepts in economics. There are several costs, including opportunity costs, associated with any economic activity. For example, economic growth

has led to fresh air becoming an increasingly scarce resource owing to traffic congestion and pollution, thus making breathing increasingly difficult for many people. In many large cities, poor air quality has been associated with high rates of chronic health problems, diseases and deaths. In order to avoid or minimize these huge costs to society and to ensure economic growth is sustainable, government intervention is needed through regulation of economic activity and investments in green technologies.

Renewable resources are not exempt from scarcity either. These are resources that can be used repeatedly and replenished naturally. Examples of renewable resources include fish stocks, cattle, water, natural oils, fruits and vegetables, timber, and renewable energy sources such as wind and hydroelectric power. Such resources can become irreplaceable if their usage rate exceeds that of their natural replenishment rate.

Figure 1.11 Bangkok, Thailand is one of the world's most congested cities

According to the most recent figures from the World Bank, in 2015, 3 billion people on the planet live on less than $2.50 per day, and more than 1.3 billion live in extreme poverty (less than $1.25 a day). These figures suggest that their basic human needs are not being met. By contrast, the 10 richest people on the planet have wealth equivalent to the poorest half of the world's population! The study of economics can help to explain why this happens and offer possible solutions to the basic economic problem of scarcity and unlimited human needs and wants.

◆ **Renewable resources** are factors of production that can regenerate and replenish, such as water, air, solar energy and cultivated plants.

Hence, it is because of scarcity that economic agents need to make sensible choices. This is because every economy faces the basic problem of scarcity and choice. The study of economics helps us to understand the decisions that households, firms and governments make, given that there are never enough resources to address all the needs and desires of individuals and societies.

Top tip!

Scarcity is a relative concept, rather than an absolute one. For example, fresh water is extremely scarce in the middle of a desert but far less scarce in a rainforest or near the coast. Similarly, fresh air is relatively scarce in Beijing, Bangkok, New Delhi and Jakarta (some of the world's most polluted cities) but less so in rural areas of Wales and New Zealand.

ATL ACTIVITY 1.4

Make a list of 10 scarce resources in the world. Compare your list with that of a partner to see how many similarities and differences you have.

ATL ACTIVITY 1.5

1. Make a list of your top 10 wants and needs.
2. Are there any 'needs' that you could actually survive without?
3. Are there any wants that might be considered as needs?
4. Identify a shortage of any good or service in your economy. Explain why the shortage has occurred.

Theory of Knowledge

Is the study of economics irrelevant in the absence of scarcity? Why does economic activity vary over time and why does this matter?

WISE ChoICES

The depletion of the world's scarce resources is a growing and intensifying economic problem. Many fish stocks have collapsed, and deforestation threatens ecosystems and our natural environment. Discuss whether the drive for economic prosperity can ever preserve the Earth's limited supply of natural resources so that economies can growth in a **sustainable** way to improve **economic well-being**.

WISE ChoICES

To what extent is **intervention** needed to ensure the economy grows in a **sustainable** way?

Opportunity cost

Opportunity cost refers to costs of an economic decision measured in terms of the best alternative choice forgone. In other words, it is the economic cost of choice. Owing to scarcity, there is always an opportunity cost when making an economic decision. For example, the opportunity cost of constructing a new airport is the other projects that could have been produced if the airport had not been built, such as the funds being spent on building roads, motorways, schools or hospitals.

◆ **Opportunity cost** refers to costs of an economic decision measured in terms of the best alternative choice forgone, that is, it is the economic cost of choice.

Figure 1.12 All aspects of government expenditure incur opportunity costs

ATL ACTIVITY 1.6

Investigate how different businesses make use of opportunity cost for the customers waiting in long queues, such as at a supermarket, beauty salon or theme park.

Some examples of opportunity cost are illustrated below:

- The opportunity cost of choosing to study IB Economics is another Individuals and societies subject you could be studying instead.
- The opportunity cost of visiting the cinema on Saturday night could be the money you would have earned from a part-time job, instead of going to the cinema.
- The opportunity cost of spending more money on national defence is using the same funds for other government priorities such as public housing for low-income families.
- The opportunity cost of a school purchasing 100 laptops for use in classrooms might be the science equipment that could not be bought as a result.
- The opportunity cost of going to university to earn a degree is the loss of income during that time if the student chose to work instead.

Economist Milton Friedman stated that *'there's no such thing as a free lunch'* (1975) to suggest that it is not possible to get something for nothing owing to the concept of opportunity cost. This is because whatever goods and services are provided, they must be paid for by someone. Therefore, opportunity cost directly influences the decisions made by consumers, workers, producers

and governments. Referring to the basic economic problem, there are competing uses for the economy's scarce resources. Thus, there is always an opportunity cost involved when allocating scarce resources, which helps economists to view the true costs of decision-making.

ATL ACTIVITY 1.7

1. Make a list of the goods and services provided by the public sector in your economy or a country of your choice.
2. Identify the goods and services that are free to individuals and those for which you have to pay.
3. List which products could be provided by a private sector firm as well as by the public (government) sector.
4. Compare and contrast the aims and objectives of a government-funded swimming pool and a private health and leisure club.

Theory of Knowledge

To what extent do you agree with the view that everything has a **cost**, but not necessarily a **price**? What does this say about the relationship between knowledge and language?

CASE STUDY

The hidden costs of 'Gangnam Style'

The music video of 'Gangnam Style' was hugely popular when it was launched on YouTube. It became the first ever video to surpass 2,000,000,000 (2 billion) views on YouTube in June 2014. At 4:12 minutes, that equates to more than 140 million hours – or more than 16,000 years! By comparison, it took 50 million person-hours to build the supercarrier USS *Gerald R. Ford* in 2013. The Empire State Building in Midtown Manhattan, New York, took 'just' 7 million person-hours to build. By the start of 2020, 'Despacito' by Luis Fonsi featuring Daddy Yankee was the most-viewed video on YouTube, with more than 6.5 billion views.

Adapted from The Economist, *June 2014, and* YouTube, *January 2020*

WORKED EXAMPLE

Suppose a firm can produce either 5,000 units of consumer goods plus 4,000 units of producer goods *or* 6,000 units of consumer goods plus 3,000 units of producer goods. Explain the opportunity cost of producing the extra 1,000 units of consumer goods. *[2 marks]*

Solution

The firm gains 1,000 units of consumer goods (from 5,000 to 6,000) at the opportunity cost of 1,000 units of producer goods (from 4,000 to 3,000).

PAPER 2 AND 3 EXAM PRACTICE QUESTION 1.1

Consider the table below, which shows the maximum combination of consumer and producer goods in an economy:

Consumer goods		Producer goods
100 units	plus	50 units
80 units	plus	60 units

Calculate the opportunity cost of one unit of producer goods in terms of units of consumer goods. *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 1.2

A farmer has land on which carrots and corn are grown. The production possibilities are shown in the table below:

Carrots ('000 kg)	Corn ('000 kg)
2	4
6	10
10	16
14	20

Calculate the farmer's opportunity cost of producing 1 kilogram of carrots in terms of kilograms of corn. *[2 marks]*

While the production and consumption of economic goods incur an opportunity cost, **free goods** are products with a natural abundance of supply and do not require any deliberate effort to obtain. Hence, people do not place a value on their production or consumption. Thus, free goods are not tradeable owing to their unlimited supply and the absence of demand. Examples include air, desert sand, seawater, airwaves, rainwater, sunlight and (to some extent) public domain web pages. There is no opportunity cost in the production or consumption of free goods.

◆ **Free goods** are products with a natural abundance of supply and do not require any deliberate effort to obtain them. Unlike economic goods, there is no opportunity cost of free goods in terms of their output.

■ **Figure 1.13** Desert sand is an example of a free good

Common mistake

'Free goods are those provided by the government, without consumers having to pay for them, for example, public healthcare, street lighting, public roads and state education.'

This statement is incorrect as free goods are those that do not have an opportunity cost due to their naturally abundant supply. The examples provided by the student are all economic goods – even though the government provides these to the general public, there is an associated opportunity cost involved in the provision of these services. Furthermore, consumers may indirectly fund these products through the tax system.

ATL ACTIVITY 1.8

Make a list of 10 goods that are limited in supply (economic goods) and a second list of goods that are unlimited in supply (free goods). How many goods can your classmates think of that are unlimited in supply?

ATL ACTIVITY 1.9

Is there anything that money cannot buy? According to some economists, money can buy anything – you just need to know where to shop! For example, some people argue that money cannot buy time, although those who are willing and able to pay can certainly save time. For example, theme parks such as Disneyland and Universal Studios offer 'fast passes' or 'express passes', which enable these customers to save time by not having to queue in line with visitors who hold ordinary entrance tickets.

Other examples of things that money cannot buy, apparently, include: health, happiness, intelligence, love and friendship. However, what economic arguments can you think of that suggest money *can* indeed buy these things?

Figure 1.14 Can money buy knowledge and intelligence?

Theory of Knowledge

An underlying assumption in economics is that individuals, firms and governments make rational decisions and choose the option that gives them the greatest benefit.

Discuss how economists and policymakers might use intuition and reasoning to make the knowledge claim that they know what is best for individuals and societies.

The basic economic questions

The three **basic economic questions** addressed by an economy are:

1. **What** production should take place? This question is about deciding which goods and services should be provided in the economy at any moment in time. For example, is it better for the economy to have more roads and airports or to have more schools and hospitals? As resources are limited in supply, decision makers realize there is an opportunity cost in answering this question.
2. **How** should production take place? This question is about the methods and processes used to produce the goods and services desired by individuals and societies. For example, decision makers will have to decide which combination of factors of production (land, labour, capital and enterprise) should be used in the production process. The economy will also need to decide whether it is feasible to use capital-intensive technologies or whether there is there a preference for specialization of labour, and which raw materials should be used or should be imported for use in the production process.
3. **For whom** should production take place? This question is about which economic agents receive goods and services. For example, should any goods and services be provided free to everyone in society, irrespective of their willingness and ability to pay for these? Should the government provide healthcare services to everyone within the country such as non-taxpayers, asylum seekers, immigrants and overseas tourists? Or should goods and services be produced only for those who can afford to pay?

ATL ACTIVITY 1.10

Using the basic economic questions, investigate how these are relevant to three different contemporary economic problems or issues in your country or a country of your choice.

Firms and individuals contribute to economic activity in the **private sector** of the economy by producing goods and providing services to meet the needs and requirements of their customers. The aim of most private sector firms is to earn a profit from economic activity.

◆ The **private sector** is the sector of the economy where private firms and individuals produce goods and services.

By contrast, the government produces or supplies certain goods and services in the **public sector**. For example, the government might provide education and public healthcare services for

◆ The **public sector** is the sector of the economy where the government produces or supplies certain goods and services.

the general public. Its aim is not to earn a profit, but to provide essential goods and services to improve the economic well-being of society as a whole. All economic agents (governments, firms and individuals) produce and consume goods and services.

ATL ACTIVITY 1.11

Discuss how a private firm that produces running shoes (sneakers) would answer the three basic economic questions: What to produce? How to produce? For whom to produce?

Figure 1.15 How do private firms address the three basic economic questions?

Means of answering the economic questions

Intervention is a key theme and concept in economics, which examines the extent to which a government should get involved in the economy in order to address the fundamental economic questions of *what*, *how* and *for whom* production should take place. Economies can address the basic economic problem of scarcity and opportunity costs by using either market forces of demand and supply or government intervention, or a combination of the two methods.

The degree of government intervention in economic activity is determined by the economic system that is implemented in the country. An **economic system** describes the way in which an economy is organized and run, including alternative views on how resources are best allocated. The challenging decision is the extent to which a government should intervene in the economy or leave economic agents (households and firms) to operate independently. There are three types of economic system (see Figure 1.16).

◆ An **economic system** describes the way in which an economy is organized and run, including alternative views on how resources are best allocated.

Figure 1.16 Economic systems

Free market economy

This economic system relies on the market forces of demand and supply to allocate scarce resources via the private sector of the economy. For example, an increase in the demand for a particular good or service will tend to cause an increase in price. Hence, in a free market economy there is minimal government intervention. According to The Heritage Foundation, based in Washington, Hong Kong has been ranked as the world's freest economy since its annual rankings (see Table 1.1) were first published in 1995, although Singapore become the world's freest economy in 2020.

◆ The **free market economy** is an economic system that relies on the market forces of demand and supply to allocate scarce resources in the economy.

Table 1.1 The world's 'freest' economies, 2020

Ranking	Economy
1	Singapore
2	Hong Kong
3	New Zealand
4	Australia
5	Switzerland
6	Ireland
7	United Kingdom
8	Denmark
9	Canada
10	Estonia

Source: The Heritage Foundation, **https://www.heritage.org/index/ranking**

According to The Heritage Foundation, there is a high correlation between a country's level of economic freedom and its standard of living (see Chapter 19). Its research data suggests that free market economies substantially outperform others in terms of economic growth, healthcare, education, protection of the environment and the reduction of poverty.

Planned economy

In this economic system, the government (or public sector) allocates scarce resources. The **planned economy** (also known as the **command economy**) is often associated with communist or socialist economies that strive for social equality. Examples include North Korea, Venezuela, Cuba and the Republic of the Congo.

Mixed economy

As its name suggests, this economic system is a combination of the planned and market economic system, with some resources being owned and controlled by private individuals and firms while others are owned and controlled by the government in the public sector. The government intervenes in economic activity to correct perceived market imperfections and market failures (see Chapters 11–14). Examples of mixed economies include Belgium, France, Italy, Spain, Mexico and the Philippines.

ATL ACTIVITY 1.12

Choose a country that has gone through a transition from a planned economy to a market based economy (or other way around). What were the economic and social reasons for change?

Top tip!

Students often quote the USA as an example of a free market economic system. In reality, State and Federal governments get involved in providing some basic services. In effect, the USA is a mixed economic system but with a much larger private sector than many other nations. Hence, it is acceptable to consider the USA as a free market economy.

Theory of Knowledge

Adam Smith's invisible hand

Adam Smith, often regarded as the Father of Economics, wrote about the 'invisible hand' in his book *The Wealth of Nations* (1776). He used this phrase to mean that incentives create efficiency gains in the economy, owing to consumers and producers acting in their own best interest. This assumes that individuals, firms and governments are rational decision makers.

How do consumers and producers make choices in trying to meet their economic objectives in a rational way?

WISE ChoICES

Discuss the extent to which governments should **intervene** in economic activity. How much **choice**, if any, should individuals have in such decision-making?

Top tip!

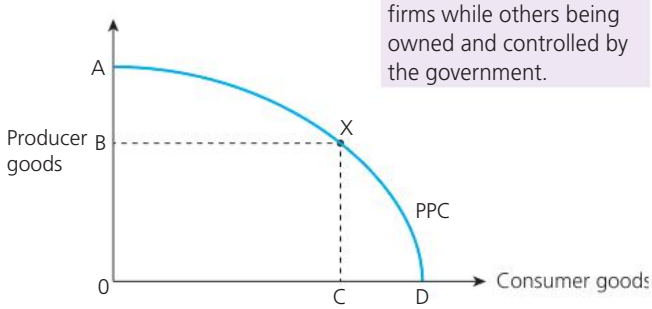
The WISE ChoICES concepts are integral to the IB Economics course, and appear throughout the syllabus, both explicitly and implicitly. For example, the concept of **intervention** appears throughout the course, providing opportunities for you to critically evaluate the balance between the case for government interference in markets and the case for market forces of the private sector to allocate resources.

◆ A **planned economy** is an economic system where the government (or public sector) allocates scarce resources.

◆ A **mixed economy** is an economic system that features aspects of both a planned and market economic system, with some resources being owned and controlled by private individuals and firms while others being owned and controlled by the government.

The production possibility curve model (AO2, AO4)

The **production possibility curve** (PPC), also referred to as the **production possibility frontier** (PPF), is a diagrammatic representation of the maximum combination of two products (goods and services) that an economy can produce, when all its resources are used efficiently, per time period. Essentially, it shows the productive capacity of the economy at any moment in time.



■ **Figure 1.17** The production possibility curve

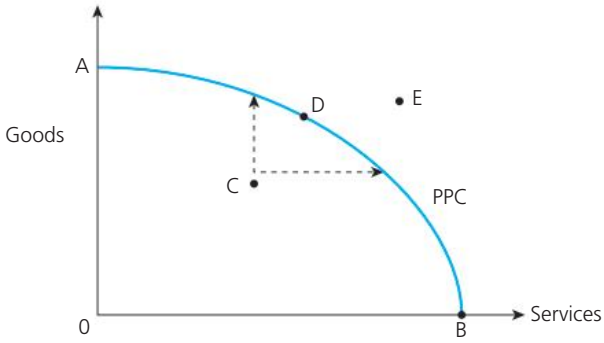
Figure 1.17 shows the production possibility curve for an economy. By operating at point X, the economy chooses to produce 0B units of producer goods and 0C units of consumer goods. Any point inside the curve suggests that not all resources are being utilized efficiently to achieve the potential output of the economy. Any point beyond the PPC is not attainable with the current level of resources and state of technology.

◆ The **production possibility curve** (PPC) or **production possibility frontier** (PPF) is a diagrammatic representation of the maximum combination of two products that a country can produce, given the efficient use of all its resources, at any moment in time.

However, to be at a point on the production possibility curve, the following conditions must be met:

1. **Full employment** – all factors of production are fully utilized. In other words, there are no unemployed resources.
2. **Efficiency** – all resources (such as land, labour and capital) are put to their best use and there is no wastage in the production process. Hence, the most output is generated from the available resources of the economy.

In Figure 1.18, points A, D and B are on the PPC, so these represent the maximum capacity of the economy if all resources are used efficiently. Point E is beyond the PPC so cannot be reached at this moment in time. Point C indicates that the economy can increase output without incurring an opportunity cost if it moves towards point D, for example. This is because the economy would be able to produce more units of both goods *and* services, as it moves towards full capacity on the production possibility curve.



■ **Figure 1.18** Features of the PPC

There are two main features when constructing a production possibility curve:

1. **The PPC slopes downwards from left to right.**

   This shows that the economy must forgo a certain quantity of one product in order to gain more quantity of the other product. For example, in Figure 1.18, if the economy moves from point D on the PPC towards point B, the economy has to give up more units of goods (shown on the *y*-axis) to get additional units of services (shown along the *x*-axis).

2 **The PPC is concave to the origin.**

A production possibility curve is typically drawn concave to the origin. This is owing to the notion of increasing opportunity cost (see below). Essentially, as an economy has more of a particular product (such as consumer goods), it sacrifices more of the other product (such as capital goods), so the opportunity cost increases.

Assumptions of the model

Like any economic model, the PPC model is based on certain preconditions or assumptions. There are four main assumptions behind the PPC model.

- **Fixed production possibilities** – The model assumes the economy only produces various combinations of two products, such as producer goods or consumer goods. Hence, the PPC is drawn as a two-dimensional diagram for ease of interpretation. In reality, this cannot represent the whole economy, but the assumption simplifies the workings of the market and it is impractical to take into account all production possibilities of all goods and services in the economy.
- **Scarcity** – There is a limited and fixed amount of resources in the economy (land, labour, capital and entrepreneurship) at any moment in time. As there is a trade-off between the two products on a PPC diagram, this illustrates the concept of scarcity in economics.
- **Constant state of technology** – The production techniques and technologies are assumed to be held constant because the economy has only a certain level of technology at any point in time. In reality, technology will improve over time, but the assumption makes it possible to analyse the consequences of any changes in output without the state of technology changing.
- **Efficiency** – The model assumes that all resources are fully utilized in an efficient way. In other words, there is no waste in the production process. The efficient use of an economy's scarce resources means that the maximum physical output of products is obtained from its fixed and limited resources. Note that this does not mean all the output is desired by consumers, only that the maximum amount of the goods and/or services is produced.

Increasing versus constant opportunity cost

At times, the production possibility curve is shown as a linear line (see Figure 1.19) rather than a curve concave to the origin. If the PPC is a straight downward sloping line, it means that the **marginal rate of transformation** (or the **rate of substitution**) between the two products in question is the same, so there is a constant opportunity cost. The marginal rate of transformation (MRT) is the gradient of the PPC, showing the opportunity cost of one product for another. A linear PPC suggests constant opportunity cost regardless of how much of each product is produced. While this is less realistic, it simplifies the idea behind the model and the opportunity cost of production.

◆ The **marginal rate of transformation** (or the **rate of substitution**) is the gradient of the PCC, showing the opportunity cost between the two products in question.

Figure 1.19 shows the trade-off between work hours and leisure hours. Hence, the opportunity cost of working more hours is fewer leisure hours. As there are only 168 hours in a week, individuals choosing to increase their working hours from 35 to 40 hours sacrifice 5 hours of leisure (133 – 128 hours per week). Thus, the opportunity cost is constant at 1:1, so, the marginal rate of transformation in this case is 1.

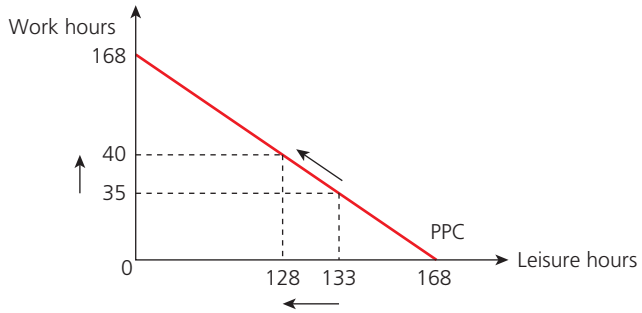


Figure 1.19 The linear PPC

Most PPC diagrams show a curve that is concave to the origin (see Figure 1.20), thus demonstrating increasing opportunity cost as more of a product is produced. For example, if the economy moves from point A to B, it gains 15 units of food (55 – 40) but sacrifices 10 units of clothes (an opportunity cost ratio of 10:15 or 1:1.5). As it moves from point C to point D, again it sacrifices 10 units of clothes but gains only 5 units of food (95 – 90). Hence, the opportunity cost ratio changes to 10:5 or 1:0.5. Previously, the economy needed to give up only 1 unit of clothes to gain 1.5 units of food, but now giving up 1 unit of clothes gives the economy only an extra 0.5 units of food. Hence, there is increasing opportunity cost as more of a particular product is produced and less of the item is made available.

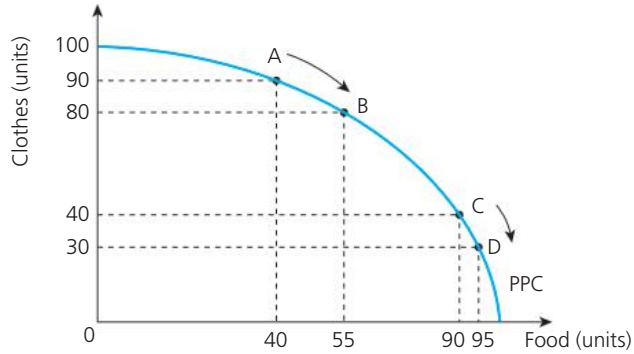


Figure 1.20 Increasing opportunity cost

Top tip!

Increasing opportunity cost can be observed by the tangent along the PPC, as shown by the gradient or the marginal rate of transformation of the PPC (Figure 1.21).

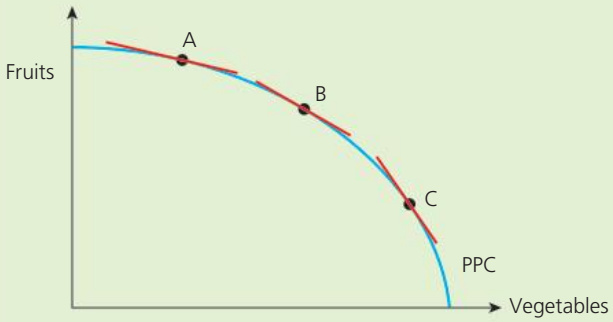


Figure 1.21 Increasing opportunity cost

Notice that the tangent at point A is flatter than at point B, and steeper still at point C. This illustrates the notion of increasing opportunity cost as more fruits need to be given up in order to produce the same additional units of vegetables as the economy moves along the *x*-axis.

Features of the model

The PPF model shows several distinct but interrelated features: opportunity cost, scarcity, choice, unemployment of resources, efficiency, actual growth and growth in production possibilities.

Opportunity cost

The PPC model assumes only two products are manufactured in different proportions, such as wheat and olive oil. The gradient of the PPC (the *marginal rate of transformation*) shows the opportunity cost of one product for another good or service. In Figure 1.22, if producers wish to increase production of olive oil from O_1 to O_2 then the amount of wheat produced will have to fall from W_1 to W_2. The opportunity cost of producing the extra O_1 to O_2 litres of olive oil is therefore W_1 to W_2 kilograms of wheat. In other words, the increased amount of olive oil is produced at the expense (opportunity cost) of wheat as the economy moves from point C to point D.

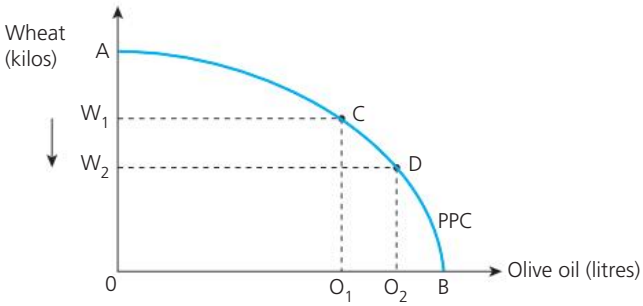


Figure 1.22 The opportunity cost of production possibilities

Scarcity

The supply of a country's factors of production are fixed because it has a limited amount of land, labour and capital, thereby limiting how much the economy can produce at any point in time. Scarcity is shown by the maximum values on the *y*-axis and *x*-axis in a PPC diagram (points A and B respectively in Figure 1.22) and the opportunity cost ratio of producing more of one product while sacrificing the output of the other good or service.

Choice

The PPC or production possibility frontier illustrates the issue and problem of choice. Any point along the PPC is a production possibility but choosing to produce more of one product (such as olive oil in Figure 1.22) means the indirect choice of less of the other product (such as wheat). In Figure 1.22, a movement along the PPC from point C to point D means the economy chooses to have more olive oil produced (from O_1 to O_2), but at the expense of fewer units of wheat (from W_1 to W_2). Governments also face a trade-off in their choice of how to use scarce resources and tax revenues. For example, if the government chooses to increase its spending on the military, then the opportunity cost will be less spending on another public service, such as healthcare or education.

Unemployment of resources

Recall that for a country to be operating on its PPC, two conditions need to be met: production is efficient and all resources are used. This means there is no unemployment of factors of production. Unemployment of resources is represented by all production possibilities within the PPC, such as point C in Figure 1.19. This is the result of resources being underutilized or not being used efficiently.

Efficiency

The PPC model demonstrates the concept of efficiency, in terms of how best to use scarce resources. **Pareto efficiency** (named after the Italian economist Vilfredo Pareto, 1848–1923) occurs when it is impossible to reallocate resources in such a way to make one party better off without making someone else worse off. For example, devoting more resources to produce capital goods for firms means fewer resources available to produce consumer goods for individuals and households. Pareto efficiency occurs at all points on a PPC.

◆ **Pareto efficiency** occurs when it is not possible to make one person better off without making someone else worse off.

Actual growth and growth in production possibilities

The PPC can shift outwards as a result of an increase in the productive capacity of the economy. This is the result of an increase in the quantity and/or quality of the economy's productive resources, such as:

- improved education and training of the workforce
- innovation of new products and services
- investments in research and development
- the discovery of new resources or the reclamation of land
- technological progress and improved production techniques
- greater population growth and the influx of skilled migrant workers
- freer and fairer international trade deals, which can lead to access to larger markets and more efficient production (see Chapter 30).

■ **Figure 1.23** Technical progress enables an economy to increase its production possibilities

For a country to shift its PPC outwards, such as shown in Figure 1.24, there must be economic growth (see Chapter 16), caused by an increase in the quantity and/or quality of productive resources of the country.

Figure 1.24 shows growth in production possibilities and actual growth. **Growth in production possibilities** is illustrated by moving towards the PPC, such as moving from point A to point B showing the economy is able to produce more output of consumer and capital products. However, **actual growth** is shown by an outward shift of the PPC, from PPC_1 to PPC_2, or shifting the economy's production from point B to point C. Economies strive to increase their productivity capacity. For example, advancements in technology will result in higher productivity and potential output for an economy. This means that even with the same amount of factors of production, more can be produced, resulting in an outwards shift of the PPC. Essentially, growth in production possibilities is caused by increased productivity of existing resources, whereas actual growth is caused by an increase in the economy's productive capacity.

By contrast, the PPC can shift inwards due to detrimental changes in the economy, such as a major natural disaster or war, which destroys a large proportion of the economy's farmland, factories and infrastructure. In Figure 1.25, the PPC shifts inwards from PPC_1 to PPC_2, representing a decrease in the productive capacity of the economy.

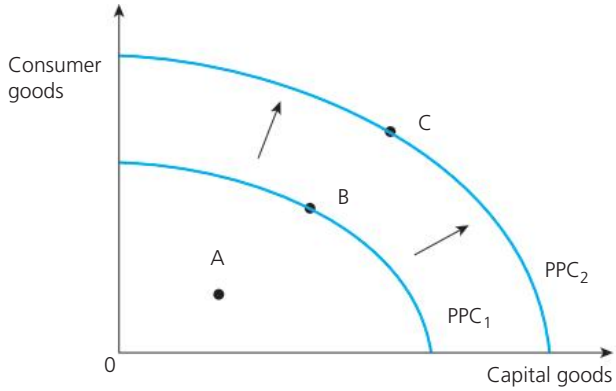


Figure 1.24 An outward shift of the PPC

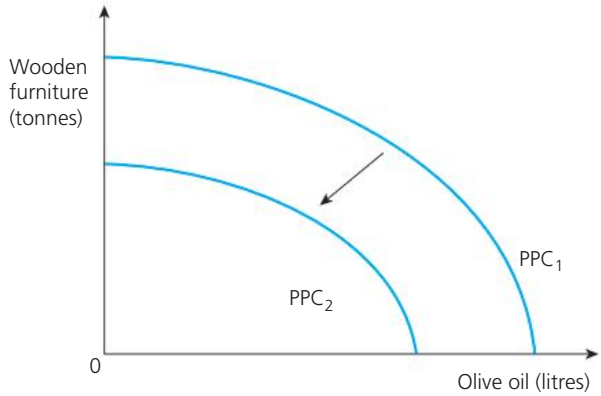


Figure 1.25 An inward shift of the PPC

ATL ACTIVITY 1.13

Investigate the economic costs of a major natural disaster of the twenty-first century. Some suggestions include:

- Hurricane Katrina in the Gulf Coast of the USA in 2005
- the 2010 Haiti earthquake
- the extensive bushfires across Australia in 2019–20.

Figure 1.26 Bushfire in the outback of Kakadu National Park, Northern Territory, Australia

ATL ACTIVITY 1.14

For each of the scenarios below, use a different PPC diagram and appropriate labels for the axes to explain:

1. the idea of opportunity cost
2. the concept of allocative efficiency
3. the concept of unemployment
4. the concept of economic development
5. what happens to the economy if there is a major natural disaster.

Top tip!

You need to be aware of the significance of the location of production possibilities, and be able to explain points under, on and beyond a PPC. The key to the production points in a PPC diagram is outlined below:

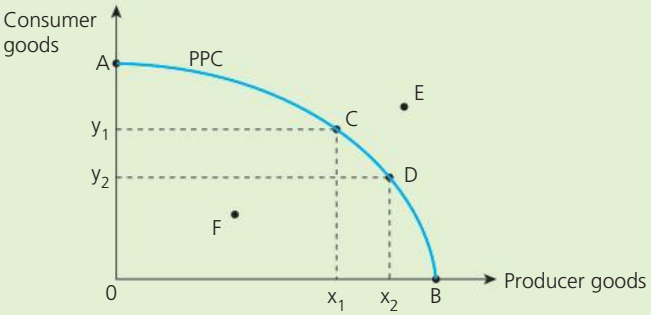


Figure 1.27 The Production Possibility Curve

- Point A – All resources are dedicated to the production of consumer products.
- Point B – All resources are dedicated to the production of producer products.
- Point C – y_1 units of consumer products are produced along with x_1 units of producer products.
- Point D – y_2 units of consumer products are produced along with x_2 units of producer products.
- Point E – This point is beyond the production possibility curve and lies outside the productive capacity of the economy, so it is currently unattainable.
- Point F – This point is within the productive capacity of the economy, so the production of both consumer products and producer products can increase (towards point C or D, for example) without any opportunity cost as some factors of production are currently not being used.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 1.3

Low-lying areas of Bangladesh are prone to flooding each year. Crops are lost and thousands of people lose their homes. During times of severe flooding, roads and railways are damaged and farmers find it impossible to transport their dwindled crops to market.

a Draw a relevant PPC diagram and explain the impact of flooding on the productive capacity of Bangladesh. *[4 marks]*

b On the diagram, draw and label a point where some of the factors of production will be idle. *[1 mark]*

c On the diagram, draw and label a point that is unattainable. *[1 mark]*

Top tip!

The PPC model is a very important one in the study of economics and can be applied in many areas of the discipline. Notice also that the features of the PPC model include several key **WISE ChoICES** concepts: scarcity, efficiency and choice. Although other concepts such as interdependence, change and economic well-being are not explicitly featured in this model, these concepts can also be seen in a PPC diagram.

Modelling the economy (AO2, AO4)

The circular flow of income model

As a social science, economics uses theories and models to explain the workings of an economy. One such model is the **circular flow of income**, which is used to explain how economic activity and national income are determined based on the interactions of economic decision makers.

Recall that income is the sum of the returns for use of the four factors of production:

- the return (or income) for land = rent
- the return for labour = wages and salaries
- the return for capital = interest
- the return for enterprise = profit
- therefore, income = rent + wages and salaries + interest + profit.

The interdependencies of economic agents in the basic circular flow of income model can be illustrated in a diagram (see Figure 1.28).

◆ The **circular flow of income model** is a macroeconomic tool used to explain how economic activity and national income are determined.

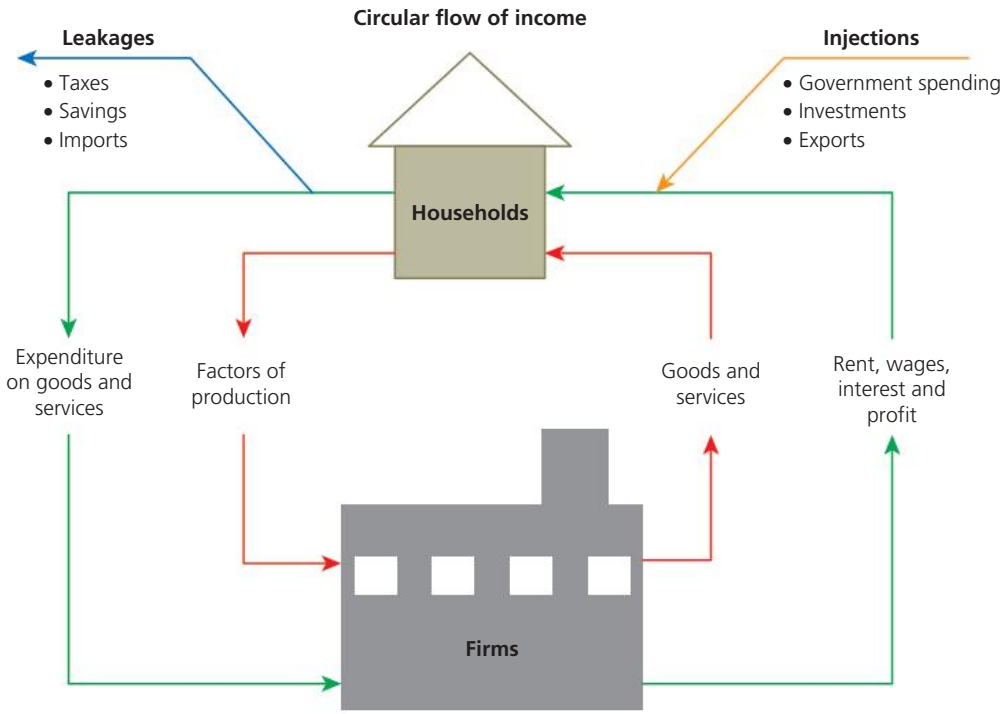


Figure 1.28 The circular flow of income

Interdependence between economic decision makers

The main **economic agents** (or economic decision makers) in an economy, as depicted in a circular flow of income diagram, are:

- **Households** – Individual consumers provide their labour services to firms, in return for income. With this, households aim to maximize their utility (the degree of satisfaction from their purchases). Rational economic behaviour results in individuals pursuing the choice that gives them the most utility. For example, with $20, individuals might prefer to visit the cinema rather than to go bowling if watching the movie brings them a greater degree of satisfaction.
- **Firms** – Businesses use factor of production to generate or supply the output of goods and services. Traditional economic theory assumes that firms aim to maximize profit. Profit is the reward to entrepreneurs for risk-taking in starting and running a business. In reality, not all firms operating in the private sector of the economy aim primarily for profit (such as charities and non-governmental organizations), but even these firms need to ensure they earn a surplus in order to keep their operations economically sustainable in the long run.
- **Government** – It is assumed in economics that the government exists to maximize social welfare for the general public, irrespective of political beliefs. The government taxes households and firms (by use of both indirect and direct taxes) in order to raise tax revenues to fund government spending on goods and services for the public, irrespective of whether they are taxpayers or not.

A **closed economy** is part of the circular flow of income model comprising domestic economic decision makers. In the closed economy model, households supply factors of production (land, labour, capital and enterprise) to domestic firms in order to produce or supply output of goods and services. In return, firms provide factor incomes to households (rent, wages, interest and profit). With the income, households spend their money on goods and services produced by firms, thus creating expenditure revenue for firms. Hence, the income flow is numerically the same value as the expenditure flow and the output flow.

◆ **Economic agents** are economic decision makers comprising households (private individuals in society), firms that operate in the private sector of an economy, and the government (the public sector of an economy).

◆ A **closed economy** is part of the circular flow of income model comprising domestic economic decision makers, that is, households, firms and the government.

ATL ACTIVITY 1.15

Historically, there have been many nations that chose to be closed, like Japan between the 1600s and the 1850s, Spain in the 1940s and 1950s, and Albania in the late 1970s to mid-1980s. Are there any contemporary countries you regard as closed? What are the potential benefits and costs of having a closed economy?

An **open economy** is part of the circular flow of income model comprising domestic and foreign economic decision makers (households, firms, the government and the foreign sector). The foreign sector facilitates international trade (exports and imports). This model shows that not all income is spent on domestically produced goods and services. This is because some of the income may be taxed, saved, or spent on purchasing imported products (see Figure 1.29). Similarly, additional funds can be injected into the economy in the form of government spending, export earnings and investment expenditure.

◆ The **open economy** is part of the circular flow of income model comprising domestic and foreign economic decision makers, that is, households, firms, the government and the foreign sector (which accounts for exports and imports).

Leakages and injections

The open economy of the circular flow of income model includes both leakages and injections. Leakages withdraw money from the economy whereas injections add money to the circular flow.

There are three categories of **leakage** (also known as **withdrawals**): savings (S), taxation (T) and import expenditure (M) as money leaves the circular flow to banks, the government and the foreign sector respectively. There are three types of **injection**: government spending (G), investment expenditure (I) and export earnings (X) as money enters the circular flow thereby stimulating economic activity. When there is national income equilibrium, the condition must hold: $S + T + M = G + I + X$.

◆ **Leakages** or **withdrawals** (W) take money out of the circular flow of the income. They comprise savings (S), taxation (T) and imports (M), that is, $W = S + T + M$.

◆ **Injections** (J) put money into the circular flow of the income. They comprise government spending (G), export earnings (X) and investment expenditure (I).

With reference to Figure 1.29 below:

- In the financial sector, savings (S) represent a withdrawal while investments (I) are injections.
- In the government sector, taxes (T) represent a withdrawal from the circular flow, whereas government spending (G) is an injection.
- Export earnings (X) represent an injection into the circular flow, whereas import expenditure (M) is a leakage.
- The circular flow of income and expenditure will change based on the relative size of all withdrawals ($W = S + T + M$) and all injections ($J = G + X + I$) at any moment in time.

The model shows that the level of economic activity depends on the relative size of injections (J) and withdrawals (W). If $W > J$ then economic activity declines, whereas economic activity increases if $J > W$.

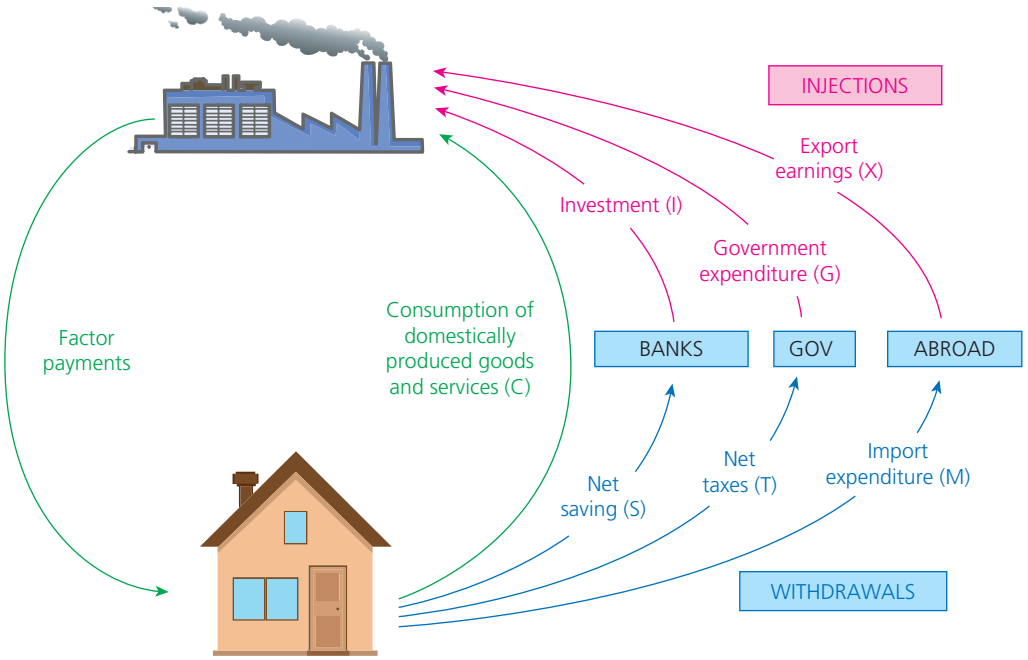


Figure 1.29 The open economy model

Top tip!

Through the skills of inquiry, you are expected to be able to evaluate and therefore appreciate both the values and limitations of economic models in explaining real-world economics. You can use the **WISE ChoICES** concepts in your discussions and explanations of real-world economic behaviour and outcomes.

WISE ChoICES

The circular flow model is a good example of the concept of **interdependence** between economic decision makers interacting and making choices in an economy. The actions of households, firms, the government, the financial sector and the foreign sector (foreign firms and households) can have profound implications on **economic well-being**. Can you provide examples to explain why this is the case?

Chapter summary

- Economics is the study of how limited resources are allocated to satisfy the unlimited needs and wants of individuals, firms and the government.
- Every economy faces the basic problem of scarcity and choice, which are two of the concepts at the heart of the study of economics. Choice is the result of scarcity in the world's economies.
- The basic economic problem is concerned with how to allocate scarce resources efficiently in order to satisfy people's unlimited needs and desires.
- Three fundamental questions arise from the basic economic problem: what, how and for whom production should take place?
- Economic agents comprise individuals (households), businesses (firms) and governments.
- Firms produce goods and services in the private sector, whereas the government operates in the public sector.
- Goods are physical (tangible) items that can be produced, bought or sold. Services are non-physical (intangible) items that are provided by firms and paid for by customers.
- Needs are the essential goods and services required for human survival. Wants are goods and services that are not necessary for survival but are human desires. The demand for wants is infinite.
- Economic goods are limited in supply, whereas free goods are unlimited in supply so have no opportunity cost in terms of production and consumption.
- The production possibility curve (PPC) is useful to demonstrate economic growth and opportunity cost.
- For a country to be on its PPC requires two conditions to be met: all resources are used (that is, there is no unemployment of factors of production), and there is efficiency in the use of these resources (that is, factors of production are allocated to their best use/purpose).
- Constant opportunity cost occurs when the PPC is linear.
- Growth can be caused by increased productivity of existing resources, whereas actual growth is caused by an increase in productive capacity.
- Factors of production (land, labour, capital and enterprise) are the resources needed to produce any good or service.
- Land (natural capital) refers to the natural resources required in the production process; for example, oil, coal, water, wood, metal ores and agricultural products.
- Labour (human capital) refers to the human resources required in the production process, in terms of both the mental (intellectual) and physical human effort.
- Capital (physical capital) refers to the manufactured or non-natural resources required in the production process; for example, machinery, tools, equipment and vehicles.
- Enterprise (entrepreneurship) refers to the skills a business owner requires to successfully combine and manage the other three factors of production.
- The returns on the factors of production are collectively called 'income', and comprise rent (land), wages and salaries (labour), interest (capital) and profit (enterprise).
- Geographical mobility of labour measures the willingness and ability of a person to relocate from one area to another for employment purposes.
- Occupational mobility of labour measures the extent to which a person is able to switch to different jobs (occupations or careers).
- The quantity and quality of factors of production will change if there is a change in the demand for and/or supply of land, labour, capital or enterprise.

2 How do economists approach the world?

Conceptual understandings (WISE ChoICES)

- Economics is a social science characterized by **interdependence**, which focuses on how people interact with each other to improve their **economic well-being**.
- The economic world is dynamic in nature and constantly subject to **change**.
- The central problems of economics are **scarcity** and **choice**, which forces societies to face opportunity costs and the challenge of **sustainability**.
- Debates exist in economics regarding the potential conflicts between economic growth and **equity** and between free markets and government **intervention**.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- economic methodology in terms of the role of **positive economics**: the use of logic, the use of hypotheses, models, theories, the ceteris paribus assumption, empirical evidence and refutation (AO2)
- economic methodology in terms of the role of **normative economics**: value judgements in policymaking, and the meaning of *equity* and *equality* (AO2)
- economic thought, including the origin and evolution of economic ideas in a historical context:
  - eighteenth century: the ideas of Adam Smith and laissez-faire economics (AO2)
  - nineteenth century: the basic ideas of classical microeconomics, the concept of the margin, Say's law, and Marxist critique of classical economic thought (AO2)
  - twentieth century: the Keynesian revolution, the rise of macroeconomic policy, and the main ideas of the monetarist/new classical counter revolution (AO2)
  - twenty-first century: the growing role of behavioural economics, the growing awareness of interdependencies between the economy, society and environment, and the compelling reasons for moving towards a circular economy (AO2).

Economic methodology (AO2)

Economic methodology refers to the study of the processes, practices and principles in relation to the discipline of economics as a social science. It includes the models, theories and assumptions of underlying economic reasoning. It deals with what is and what is not economics. Although economics as an academic discipline goes back at least to the eighteenth century, the methodology of economics has evolved over time.

◆ **Economic methodology** is the study of the processes, practices and principles of economics. It includes the models, theories and assumptions underlying the discipline.

Economic methodology is concerned with three key aspects:

- how economics functions
- how it could function, and
- how it should function.

This section of the IB Economics syllabus considers two main dimensions of how economists approach the world:

- Economic methodology in terms of:
  - the role of positive (objective) economics, and
  - the role of normative (subjective) economics.
- Economic schools of thought from the past several centuries, including:
  - eighteenth century: Adam Smith and laissez-faire economics
  - nineteenth century: Jean-Baptiste Say, Karl Marx and Classical economics
  - twentieth century: John Maynard Keynes, Monetarism and New Classical economics
  - twenty-first century: the growing role of behavioural economics; increasing awareness of the interdependencies that exist between the economy.

British philosopher and economist John Neville Keynes (1852–1949) categorized economic methodology as positive and normative economics. Positive economics is the study of *what is* and the way the economy actually works. Normative economics is the study of *what should be* and the way the economy ought to work. John Neville Keynes was the father of John Maynard Keynes (1883–1946) – see Economic Schools of Thought later in this chapter.

Theory of Knowledge

To what extent does the distinction between positive and normative statements exist in other academic disciplines and areas of knowledge?

The role of positive economics (AO2)

The methodology of positive economics was featured as the first essay in Milton Friedman's book *Essays in Positive Economics* (1953), which explores John Neville Keynes's distinction between positive and normative economics. **Positive economics** is the study of economics that is provable. This means factual statements about the economy or statements of 'what is'. For example, unemployment in the economy will tend to increase if corporate taxes are continually raised as firms become unprofitable so will not need to hire so many employees. Positive economics relies on reasoning, logic and empirical evidence. Hence, such statements can be verified or refuted by referring to facts, evidence or further investigation.

◆ **Positive economics** is the study of economics that is provable. This means factual statements about the economy or statements of 'what is' rather than 'what ought to be'.

Assuming all other things remain unchanged, the following are positive economic statements:

- Scarcity is the foundation of the basic economic problem.
- Factors of production are required to generate economic output.
- Customers will buy more of a product if the price is reduced.
- A fall in average incomes will lead to higher unemployment in the economy.
- Increasing the national minimum wage will raise costs for businesses.
- A reduction in income tax rates will create incentives for people to find employment.
- A tax on plastic carrier bags will reduce the volume of plastic waste.
- Greater spending on healthcare services will help to increase average life expectancy.
- Pollution control is effective through an efficient system of penalties and fines.
- Globalization creates more winners than losers.
- Cutting interest rates is likely to increase consumption and investment expenditure, but can cause inflation in the economy.

■ **Figure 2.1** Taxing single-use plastics is likely to reduce the amount of plastic waste

Top tip!

It is possible for a positive economic statement to be incorrect or inaccurate. The important thing to remember is that positive economics focuses on objective statements about 'what is' at any moment in time rather than 'what should be'. So, such statements can be tested for truth.

The use of logic

Logic is integral to the study of positive economics. **Logic** refers to rationality and reasoning, rather than emotions or beliefs, in explaining economic phenomena and policymaking. For example, it is logical to assume that people will make rational choices based on the information available to them. It is also logical to assume that private sector firms act in their self-interest so most will aim to maximize profits.

◆ **Logic** refers to rationality and reasoning, rather than emotions or beliefs, in explaining economic phenomena and policymaking.

Logic involves presenting and analysing facts and data as they are, in an objective, and often measurable, way. Logic enables economists to explain a chain of statements to deduce truth (based on facts and empirical evidence) regarding a hypothesis, model or theory. For example, if bad weather causes the cost of harvesting the tomatoes used to make Heinz tomato ketchup to increase, the following logical sequence of events is likely to take place:

1. The cost of producing Heinz tomato ketchup increases.
2. This reduces the ability and willingness of the firm to supply its product.
3. The subsequent fall in supply makes Heinz tomato ketchup more scarce.
4. This causes the price of Heinz tomato ketchup to increase.
5. Therefore, the quantity of Heinz tomato ketchup demanded by consumers will fall.

The above chain of logical events can be tested using empirical evidence to accept or refute the hypothesis, model or theory. Logic means none of the sequences above consider subjectivity or value judgements.

Milton Friedman, the American economist and winner of the Nobel Memorial Prize in Economic Sciences (1976), argued that economics should be free of subjectivism and value judgements for it to be respected as a science. He believed that objectivity, reasoning and logic should inform normative economics, such as whether to reduce progressive tax rates or to increase the national minimum wage. Friedman argued value judgements distort policymaking as predictions and decisions are based on opinions rather than facts about the economic consequences of different policies.

Theory of Knowledge

Do **religion** and **indigenous societies**, as optional themes in Theory of Knowledge, have any role in the study of economics?

■ **Figure 2.2** Do differences in opinions and value judgements distort objectivity?

Theory of Knowledge

How do consumers and producers make economic choices based on faith as a form of knowledge?

Theory of Knowledge

'As social beings, humans are never truly logical in their thinking or behaviour.' To what extent do you agree with this statement?

The use of hypotheses, models and theories

In economics, a **hypothesis** is an assumption, notion or educated guess made before research has been conducted. Economists formulate a hypothesis, often based on an observation or idea that explains something. This enables economists to test the hypothesis in order to accept or refute the hypothesis, and to gain a clearer understanding and evaluation of economic behaviour. Two examples of hypotheses are:

◆ A **hypothesis** is an assumption, notion or educated guess made before research has been conducted.

- 'The introduction of a national minimum wage will improve the economy' or phrased as a question: 'To what extent will the introduction of a national minimum wage be good for the economy?'
- 'Men, on average, are paid more than women' or phrased as a question: 'Are men, on average, paid more than women?'

The hypothesis statement helps researchers to keep their investigation focused. Relevant data can then be collected to test the hypothesis to inform policy decision makers depending on whether the data prove the hypothesis or refutes it.

Theory of Knowledge

What assumptions do economists make when they apply economic hypotheses to the real world?

Hypotheses are classified as either alternative or null. The majority of hypotheses in economics are **null hypotheses**, which are assumed to be true (such as the example above about minimum wages). Any evidence that contradicts the null hypothesis is captured in the **alternative hypothesis** (such as 'The introduction of a national minimum wage will *not* improve the economy' or 'On average, men are *not* paid more than women'). Hence, the alternative hypothesis is a direct contradiction of the null hypothesis. Either of these hypotheses can be used in economics to test a research question, an idea or a claim, using empirical evidence and data to provide a possible explanation to the phenomenon.

The process of hypothesis testing is shown below:

1. The null hypothesis is formulated and assumed to be true.
2. Research data are gathered to test the null hypothesis.
3. If the empirical evidence is consistent with the null hypothesis, it is considered to be true (although data that are subsequently available could still change this).
4. If the empirical evidence is *not* consistent with the null hypothesis, it is refuted (rejected) in favour of the alternative hypothesis.

A **model** occurs when a hypothesis has been repeatedly tested and proven or rejected. Models are only accepted by economists once they have been thoroughly tested and can be used to explain the real world. For example, the price mechanism as an economic model explains the allocation of the world's scarce resources (see Chapter 5), using empirical evidence and data to explain resource allocation by creating signals and incentives to buyers and sellers in the market.

◆ A **model** is a hypothesis that has been repeatedly tested and proven or rejected and can be used to explain the real world.

Theory of Knowledge

Given that economic theories do not show exactly what happens in the real world, how can we know if they are of any use or value? Refer to one of the six real-word issues (page v) in your response.

The word 'theory' comes from the Greek word 'theorein', which means 'to look at' or 'to contemplate'. A **theory** is a broad generalization used to explain situations or scenarios already supported by economic evidence and data from economic models. For example, the theory of demand (see Chapter 3) generalizes that as the price of a product falls, the quantity demanded by consumers will increase. This theory has been repeatedly tested and is widely accepted so can be used to make predictions about economic behaviour. Economic theories are based on logic and

◆ A **theory** is a broad generalization used to explain situations or scenarios already supported by economic evidence and data from economic models.

backed by empirical evidence, using models to represent and analyse the complexities of a real phenomenon.

An economic hypothesis is accepted as a model once it has been tested enough times to confirm its validity (with sufficient empirical data and evidence) in successfully explaining a phenomenon. Similarly, an economic model becomes an accepted theory once it is sufficiently tested and confirmed to successfully explain the phenomenon.

Note: like any economic hypothesis, an economic model or theory can also be refuted after rigorous testing or when new evidence contradicts or invalidates the existing model or theory.

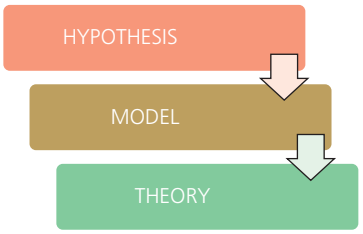


Figure 2.3 The relationship between hypotheses, models and theories

Hypotheses, models and theories are used in economics to explain:

- relationships – such as the relationship between the changes in household income and the quantity of a product demanded or supplied (see Chapter 7).
- cause and effect interactions – such as the impact of higher taxes or interest rates on the economy (see Chapter 17).

Friedman argued that for a hypothesis, model or theory to be useful, it must meet two criteria:

- **Simplicity** in predicting at least as much as an alternative hypothesis, model or theory, although requiring less information. Friedman rejected the idea of testing a hypothesis, model or theory by the realism of its assumptions, as it is the simplicity of economic hypotheses, models and theories that help to predict economic behaviour.
- **Fruitfulness** in the precision and scope of the predictions of the economic hypotheses, models, and theories and their ability to generate additional research and understanding of economics.

Theory of Knowledge

How realistic are economic theories and models in explaining real-world issues?

◆ **Ceteris paribus** is a Latin phrase meaning 'all other factors remaining constant' or 'all else unchanged'. It is used by economists to explain cause and effects of economic variables.

The ceteris paribus assumption

Economics is regarded as a social science, examining the behaviour of individuals and societies from an economic perspective. As it is not an exact science, it can prove difficult to isolate a particular variable or to hold it constant. For instance, the demand for ice cream is affected by many factors, such as the price, the income of customers, or even the weather. In reality, things are not constant and countries deal with the same economic problem or issue in different ways. Therefore, the concept of ceteris paribus enables economists to simplify and explain possible causes and effects, even if assumptions have to be made.

Milton Friedman (1953) argued that *'a theory or its assumptions cannot possibly be thoroughly realistic'* and nor do they need to be as it is impossible and unnecessary to examine every single *factor 'directly underlying the supply and demand for wheat'* (or any other product). Friedman argues that trying to make an economic theory overly realistic simply renders it useless owing to the ceteris paribus notion.

Ceteris paribus is a Latin phrase meaning *'all other factors remaining constant'* or *'all else unchanged'*. For example, the ceteris paribus condition allows economists to examine what is most likely to occur to the level of demand for a good or service if one variable affecting demand changes (such as income) while holding all other variables (that affect demand for the product) constant. For example, ceteris paribus, firms will increase the supply of flowers if the price of flowers increases.

Figure 2.4 Can you think of three factors that affect people's demand for flowers?

Empirical evidence

Empirical evidence refers to first-hand data and information acquired by observation or experimentation of certain behaviours and patterns. It is an important aspect of positive economics because it involves looking at what has happened in the economy and what is currently happening in order to make logical and rational predictions for the future. Any conclusions drawn from positive economic analyses can be tested and supported or refuted by facts, figures and other research data. Furthermore, what is observed or experienced can be different from what an economic hypothesis, model or theory might predict.

A positive economic statement (such as 'an increase in the price of a product causes an increase in the quantity demanded') can be accepted or rejected based on empirical evidence. In the study of positive economics, empirical evidence is vital for constructing hypotheses, models and theories and for testing their validity. The use of empirical evidence is also used to check and validate deduced factual statements and the use of logic in economics.

However, economics is not an exact science so controlled experiments are not viable to collect data to prove or refute a hypothesis. Instead, uncontrolled events and economic behaviour, influenced by value judgements, makes it more difficult to produce empirical evidence to justify the acceptance or rejection of a hypothesis. For example, reducing interest rates in the economy should lead to more consumption and investment (see Chapter 17) because it is relatively cheaper for households and firms to borrow money to consume and invest. However, households and firms may lack confidence in the economy so refrain from borrowing money to fund consumption and investment, despite the lower costs of borrowing.

◆ **Empirical evidence** refers to first-hand data and information acquired by observation or experimentation of certain behaviours and patterns.

ATL ACTIVITY 2.1

The field of education continually goes through change, in attempts to make it better. Discuss how the standards of teaching and learning at your school could be improved, and how empirical evidence could be collected and used to measure the effectiveness of your suggestions.

Theory of Knowledge

Discuss the extent to which empirical evidence in a continual world of change can ever allow economists to arrive at the truth about real-world issues.

Figure 2.5 Positive economics is based on evidence, data, and facts

Refutation

Positive economic statements can be verified or refuted by referring to facts, data and empirical evidence or further investigation. An economic statement is considered to be false if it is refuted (rejected) by the empirical evidence.

For example, economic theory suggests that more people will reduce savings and increase consumption if interest rates fall (because there is less of an incentive to save money and cheaper to borrow money to spend). This assumption or forecast of consumer behaviour is based on

◆ **Refutation** is the act of a statement or theory being proved to be wrong or false by the empirical evidence.

positive economics because past data can be used to support this. Whether this materializes in reality depends on many factors, such as the level of consumer confidence in the economy.

For example, Japan used a zero interest rate policy (ZIRP) for more than 20 years, starting in 1999, owing to the country's slow economic growth, yet the Bank of Japan's monetary policy measures (see Chapter 20) had a minimal impact on the economy. For Japan, the experience and evidence demonstrate the limitations of what central banks can do in isolation, so the theory that lower interest rates help to stimulate the economy can be refuted.

As an evolving social science, positive statements in economics are accepted or refuted by using ongoing research, data and empirical evidence. Examples of such statements that have been accepted and refuted at different times in different parts of the world include:

- Each country has an optimal tax rate; an increase in tax rates can bring in more tax revenues for the government up to a certain amount before disincentives set in.
- Globalization is beneficial for everyone in the long run.
- Individuals make rational expectations about the future that affect their current spending pattern.
- Fiscal stimulus through tax cuts and increased government spending can get countries out of a recession when there is a downturn in overall level of economic activity (see Chapter 16).

ATL ACTIVITY 2.2

The United States, Japan and several European Union member nations have all used a zero interest rate policy (ZIRP) to stimulate economic activity. Investigate the goals, potential benefits and limitations of using a ZIRP as a form of government intervention in the economy.

The role of normative economics

Recall that *positive economics* focuses on objective economic analysis, based on logic and hypotheses, models and theories that are backed by empirical evidence. This contrasts with **normative economics**, which considers opinions, beliefs and statements of 'what should be' (or what ought to be). Normative economic statements or claims are subjective, rather than objective (unbiased), expressing a value judgement about what is perceived to be desirable or undesirable about the economy. Normative economic statements cannot be verified or refuted by referring to facts, evidence or further investigation.

◆ **Normative economics** considers people's varying opinions and beliefs about what should be (or what ought to be). Such claims are subjective and expressed as value judgements.

Examples of normative statements include:

- The government should impose higher income taxes on wealthier people.
- The country must impose higher tariffs (import taxes) on foreign goods entering the country.
- The government should spend more money on education and less on national defence.
- Prosperous countries must donate more money to low-income countries.
- Society ought to ensure housing is provided for everyone.
- The state is supposed to ensure that everyone has a job.
- The retirement age in the country should be raised to 70 for women and 72 for men.

Although it would seem rational for economists to rely on evidence-based research to determine economic policies, the reality is that most

Figure 2.6 Should the retirement age be raised?

economic decisions are still influenced by value judgements (the opinions and beliefs of the decision makers and policymakers), at least to some extent. The reality is policymakers often base future predictions and decisions on value judgements regarding why things happen as they do.

Value judgements in policymaking

Value judgements are the beliefs of individuals and societies about what is right or wrong, or good or bad. These judgements are often influenced by morals, ethics, and cultural values and beliefs. For example, consider the following examples of normative economic statements and the role of value judgements in policymaking:

- On average, men should be paid more than women.
- It is wrong to have any form of discrimination when hiring and promoting people in the workplace.
- Children should study English, mathematics and science up to the age of 18.

There is much debate regarding whether value judgements should have any role in policymaking. For example, free market economists argue that positive discrimination occurs as a result of market forces (statistically, men might be paid more than women, on average, for several rational and logical reasons) rather than interventionist policies that may create equality but at the opportunity cost of economic efficiency.

In reality, value judgements often play a crucial role in policymaking especially because economic well-being is concerned with the desirability or undesirability of different economic policies. Judgeing the social welfare of economic policies involves value judgements so is not an easy task especially owing to individual perspectives and interpersonal comparisons of welfare and well-being. For example, the UK officially left the European Union on 31 January 2020 (having joined in 1973). The UK voted to leave the EU by a majority of 52 per cent to 48 per cent. The majority of voters in England and Wales opted to leave, while Scotland had voted to remain in the EU. The decision came four years after the UK's vote to 'Brexit', but took three different Prime Ministers to see it through. It remains to be seen whether the decision to Brexit really will benefit the British people and its economy.

◆ **Value judgements** are the beliefs of individuals and societies about what is right or wrong, or good or bad.

ATL ACTIVITY 2.3

The UK joined the European Union in 1973. Forty-seven years later, the UK permanently left the EU in January 2020. Investigate the purpose of EU membership, and evaluate the advantages and disadvantages of EU membership for the UK (or any other EU member state of your choice).

Figure 2.7 Brexit was the 'choice' of the British people – but was it the correct one?

Theory of Knowledge

What role, if any, do opinions and value judgements have in the creation of knowledge in economics?

Therefore, value judgements mean that individuals choose differently so there is not necessarily a unanimous social choice. Individuals make different choices based on their different tastes, preferences and beliefs. The existence of value judgements means that policymaking and welfare economics cannot be purely objective or free from subjectivity.

The meaning of equity and equality

In economics, the concept of **equity** is about fairness. It is concerned with economic fairness in the distribution of resources. In other words, people who are better qualified, have more work experience and work harder should be rewarded accordingly. Equity, although subjective to some extent, justifies the use of the price mechanism to rationalize the allocation of scarce resources. For example, doctors are paid more than junior nurses owing to the relatively higher demand and lower supply of doctors. The same applies to pilots and cabin crew, school principals and classroom teachers, or a head chef and servers. However, critics of free market economies raise concerns about such vast differences in income and wealth caused by the price mechanism in the rewarding of effort, ability and wealth creation.

◆ **Equity** is the concept concerned with economic fairness in the distribution of resources.

Another example of equity is the allocation of car park spaces in retail locations. Many retail operators allocate parking spots closest to supermarkets and shopping malls for disabled parking. This may be deemed to be fair (equitable) by society because a disadvantaged segment of the market is being given an advantage by reallocating resources in their favour. This strategy is different from equality as the parking bays are not equally distributed given the disproportionate number of people who qualify to use these spaces and there is no equal access (as qualifying members must have a legitimate disabled badge on display in their vehicle).

Figure 2.8 Car park allocations are often assigned in favour of disabled badge holders

CASE STUDY

Who gets to attend an IB Workshop?

The IB offers official upskilling workshops to teachers as part of their professional training and development. These workshops typically last three days and are run all over the world. They are facilitated by a Workshop Leader. In order to promote equity, the allocation of a Workshop Leader to particular IB workshop is based on the following criteria:

- Location (country of residence) – owing to visa and/or cost restrictions, such as airfares and hotel accommodation.
- Workshop Leaders' profile – based on the Workshop Leaders' level of experience, their own upskilling training, and feedback from past workshop participants.
- Rotation – to give IB Workshop Leaders the fairer chance of facilitating an IB Workshop, rather than the same few people getting the opportunity to do so.

However, equity is not the same as **equality**, which regards everyone being equal. Hence, equality is about social fairness and collectivism from the perspective of a planned economic system (see Chapter 1). A fundamental premise for discussions of equality is that individuals in society should have equal recognition. This means that every individual person matters, irrespective of factors such as gender, gender reassignment, age, ethnicity, religion, sexual orientation, disability, marital status, and so forth. **Inequality** is a major concern for many economies. This occurs

◆ **Equality** is the concept concerned with everyone being equal and should have equal recognition. It is about social fairness and collectivism.

when economic opportunities are unevenly distributed between different groups in society. For example, children from low-income families may be denied access to high-quality education and healthcare, which limits their ability to earn high levels of income in the future.

Common mistake

Students often confuse the concept of *equity* with the term *equality*. **Equity** is about perceived and justified economic fairness, whereas equality is about economic outcomes being the same or similar for different people. Equity is about empowering people to be successful, whereas **equality** is about treating everyone in the same way. In short, equity is about fairness and equality is about sameness. See Chapter 19 for more details about the relationship between equity and equality.

Overall, when considered together, positive economics and normative economics provide decision makers with a clearer understanding of the impact of different policies. The two economic methodologies cover actual facts combined with value judgements about equity and equality in formulating and executing policy decisions.

Economic thought (AO2)

This section of the IB Economics syllabus covers the different economic schools of thought from the eighteenth century to the twenty-first century. The assumptions that guide the study of economics as a social science have changed considerably throughout history, with direct and significant influence on the global economy and our everyday lives. Essentially, it is a historical account of economic ideas, beliefs and principles. Examining the origins of economic ideas in a historical context helps to explain and gain an understanding of how economists approach the world. This is important because economic ideas have been and continue to be instrumental in shaping the diverse policies of different countries throughout the world.

Eighteenth century

Adam Smith (1723–90) is widely credited as being the Father of Economics. In *The Wealth of Nations*, Smith laid the foundations of free market (or laissez-faire) economics. He believed in a **laissez-faire** approach to economics, which refers to the policy of leaving things to their own devices, without government interference. He argued that rational economic behaviour (people acting in their best self-interest) and competition are fundamental to economic well-being and prosperity. Smith called this idea the **invisible hand** (of the price mechanism). In other words, consumers and producers, each acting in their own interest, will ensure that the optimal output is produced. He believed that competition was self-regulating so government intervention should have no role to play in business activity. Hence, he argued against the use of tariffs and other barriers to international trade, and advocated the liberation of markets.

Tim Harford, author of *The Undercover Economist* (2005), used a modern-day example to illustrate the idea of Smith's invisible hand in action, that a laissez-faire approach enables market forces to allocate resources efficiently:

> *A cappuccino is intriguing to an economist: he doesn't know how to make a cappuccino, and he knows that nobody else does either. Who, after all, could boast of being able to grow, pick, roast, and blend coffee, raise and milk cows, roll steel and mold plastics and assemble them into an espresso machine, and, finally, shape ceramics into a cute mug? Your cappuccino reflects the outcome of a system of staggering complexity. There isn't a single person in the world who could produce what it takes to make a cappuccino.*
>
> *(Tim Harford page 4)*

Theory of Knowledge

What does inequality allow us to know about a society and its values?

WISE ChoICES

While hiring more women or people of different ethnic backgrounds might be deemed to be **equitable**, to what extent is does this practice jeopardize **efficiency**?

◆ **Economic thought** refers to a historical account of the different economic ideas, beliefs and principles.

◆ **Laissez-faire** is the approach to economics that allows transactions to take place without government interference. It is a French term that loosely translates as 'let it be'.

◆ The **invisible hand** is Adam Smith's metaphor describing how individuals' behaviour and decision-making ends up benefiting society as a whole because they are incentivized to act in their own self-interest.

Figure 2.9 The invisible hand ensures we can get a cappuccino

As a free market economist, Adams refuted the idea of mercantilism that had dominated economics and international trade throughout the sixteenth to eighteenth centuries. Mercantilism refers to the idea that a country's wealth and economic power were limited to its **finite resources**, so there was reliance on increasing the nation's exports and limiting its imports in an attempt to acquire precious metals such as gold and silver. Prior to the rise of Smith's classical economic theory, most economies were run using a 'command and control' approach, with the government tightly controlling all aspects of economic activity.

◆ **Finite resources** are part of the central problem of economics referring to factors of production being limited in supply.

Nineteenth century

Economics in the nineteenth century was dominated by classical microeconomic utility theory, the concept of the margin and classical macroeconomics (Say's law). Adam Smith is regarded by most people as the originator of classical economic theory. This section of the IB Economics syllabus also considers Karl Marx's critique of classical economic thought.

Classical economics is a broad term that refers to the main economic school of thought during the eighteenth and nineteenth centuries. Classical economic theory focused on the self-regulation of market forces to allocate resources in an efficient way. Classical economics is the opposite of a 'command and control' economic system and became closely associated with economic and political freedom (driven by incentives and minimal government intervention in economic activity). Classical economists were in favour of competition and free international trade among workers and firms in order for markets to work efficiently.

◆ **Classical economics** is the main economic school of thought during the eighteenth and nineteenth centuries, which focused on self-regulating markets to allocate resources efficiently.

Classical microeconomics – utility theory

An important development of classical economics was the concept of **utility theory**. **Utility** refers to the level of consumer satisfaction from the consumption or use of a product. A product has utility if it provides the consumer with satisfying use and/or enjoyment (happiness or usefulness). The theory, developed by British economist Alfred Marshall (1842–1924), assumes that utility maximization is the most important factor in determining the demand for a good or service. Classical economists believe that a consumer aims to maximize personal satisfaction so makes decisions based on their judgement of how much utility they get from the purchase and use of the product.

◆ **Utility** refers to the level of consumer satisfaction from the consumption or use of a good or service.

Utility theory also assumes rational behaviour. This means that people act rationally when making economic decisions to maximize their satisfaction. Many of the fundamental theories studied in economics, such as demand and supply (see Chapters 3 and 4) were devised by

classical economists based on utility theory. This is because economic utility (the usefulness of a good or service or the satisfaction that consumption of a product brings) directly influences the demand, and therefore price, of any given good or service. Essentially, classical economists of the nineteenth century had attempted to explain the economic reality of price, which they believed was determined by the utility of the good or service.

Figure 2.10 Diamond rings can bring a lot of utility so have a very high price

Top tip!

Note that utility tends to decline as the supply of a good or service increases because the resource becomes less scarce, so is valued less by consumers. This is one reason why highly scarce resources are very expensive, such as diamonds.

The concept of the margin

Classical microeconomics theory assumes that consumers purchase a good or service depending on the marginal utility of consumption. **Marginal utility** refers to the benefit or satisfaction gained from consuming an additional unit of a good or service. Classical economists assumed that the level of utility can be quantified in order to illustrate the varying degree of satisfaction from consumption. Marginal utility is measured by calculating the change in utility as an additional unit is consumed.

◆ **Marginal utility** refers to the benefit or satisfaction gained from consuming an additional unit of a good or service.

The **law of diminishing marginal utility** (see Chapter 3) states that as the level of consumption of a good or service increases, the marginal utility derived from the consumption of each additional unit of that product eventually declines. While marginal utility can increase incrementally from the consumption of an additional unit, the theory suggests that marginal utility will eventually diminish. For example, the utility for a consumer eating a first slice of pizza or drinking a glass of fizzy drink is likely to be quite high, whereas the utility of a second or third slice of the same pizza or drink is likely to be lower. Overeating or overdrinking can eventually cause marginal utility to be negative, that is, the consumer gets sick! Therefore, the law of diminishing marginal utility of consumption is a key determinant of a product's price as consumers are not willing to pay the same price for additional units if they experience diminishing marginal utility.

◆ The **law of diminishing marginal utility** states that as individuals consume more of a product, the satisfaction gained from each additional unit of consumption declines. Hence, customers would only purchase more units at lower prices.

Figure 2.11 The idea of free refill drinks is based on the law of diminishing marginal utility

ATL ACTIVITY 2.4

Economics is one way of looking at how the world achieves well-being, utility, benefit or satisfaction, as suggested by British classical economist and philosopher John Stuart Mill (1806–73). What is the difference between well-being and marginal utility? Discuss this question with a partner and be prepared to share your answers with the rest of the class.

Classical macroeconomics (Say's Law)

Jean-Baptiste Say (1767–1832) was a French classical economist and entrepreneur who advocated for the liberalization of markets in favour of competition and free international trade. He supported the laissez-faire view that governments should not interfere with the operations of the free market. In his book, *Treatise on Political Economy* (1803), Say developed the economic theory that stated the income from selling goods and services in the past becomes the source of expenditure to stimulate demand for current production. This became known as **Say's law**, or more formally Say's law of markets.

Say suggested that to have the means to buy something, economic agents must first have produced something to sell. In other words, an individual's ability to demand goods or services is based on the income they have produced (earned) from their own past acts in the production process. John Maynard Keynes subsequently interpreted and simplified this as 'supply creates its own demand'. Say's law implies that national output (or the country's gross domestic product) is the key to economic growth and prosperity, and that government policy should therefore focus on production rather than consumption. Say's law went on to influence supply-side economists of the twentieth century.

Theory of Knowledge

Given that a consumer's utility cannot be quantified accurately or truly, how is it possible to know anything about consumer happiness if it cannot be scientifically or mathematically measured?

◆ **Say's law** states that the ability to purchase a product depends on the ability to produce or supply, thereby generating income. Essentially, this means that supply can create its own demand.

Marxist critique of classical economic thought

Karl Marx (1818–83) was a German philosopher and sociologist who emerged to challenge the classical economic school of thought. He developed an alternative perspective to classical economics, by focusing on the surplus value created by a capitalist (free market) economic system. Marx saw the invisible hand of the market economy as more of an invisible fist or an iron fist of capitalism in the exploitation of workers. He accused classical economists of promoting capitalism in favour of those who enjoyed privileges and the power to ruthlessly exploit workers as industrialists maintain power in the workplace. This, according to Marx, would enable capitalists to earn monetary profits, with minimal if any concern for human worth and need.

Marx's **labour theory of value** (1867) made the basic claim that the value of a good or service can be objectively measured by the average number of hours required to produce that product. For example, if it takes twice as long to produce a pair of shoes compared to a shirt, then shoes are twice the value of shirts. Hence, the price of shoes is twice the price of shirts. A major criticism of this theory is that it ignores the different costs of production, such as the differences in costs of making shoes and shirts.

◆ **Marxism** is an approach to macroeconomic policy that focuses on meeting the needs and values of the masses, rather than for the privilege of a minority of capitalists.

Marxism thrived following the spread of communism throughout Eastern Europe during the first half of the twentieth century. He condemned capitalism as an economic system that ignores the needs and values of the masses. However, by the late nineteenth century, economists had rejected Marx's labour theory of value, arguing that capitalists do not earn profits by exploiting workers but that entrepreneurial capitalists earn profits by organizing production efficiently, by investments (sacrificing current consumption in favour of capital expenditure), and by taking calculated risks. He had also underestimated the flexibility of a free market economy as the world advanced through history. Furthermore, comprehensive planning, as advocated by Marx, became increasingly unrealistic as the world's economies grew and became more interdependent.

Theory of Knowledge

Marxist economics focuses on the role of labour in the development of an economy, but it ignores the role of entrepreneurs. To what extent does history reinforce and create new knowledge in economics?

Twentieth century

Following the Great Depression of 1929–39 (regarded as the worst economic downturn in the history of the industrialized world) and the Second World War, classical economics declined in popularity in favour of the Keynesian revolution. This school of economic thought (**Keynesian economics**) was named after British economist John Maynard Keynes (1883–1946), who advocated the use of interventionist macroeconomic policies to pull the world's economies out

of a global depression. This would be achieved by a combination of increasing government expenditure and lowering taxes in order to stimulate aggregate demand in the economy (see Chapter 17). Keynes was highly critical of the classical economic argument that natural forces (incentives of the market economy and the invisible hand) would be enough to help the economy recover from an economic depression and the problems of long-term unemployment.

Figure 2.12 Keynes reasoned that government intervention is needed to resolve the problems of unemployment

Keynes was the most well-known student of British economist Alfred Marshall (1842–1924), who with his mathematic stance on economics had developed the theories of demand and supply (see Chapters 3 and 4), marginal utility and costs of production.

Keynes also refuted Say's law. He argued that the Great Depression created over-production in the economy and the subsequent deficiency of demand. Keynes suggested that the depth and severity of the Great Depression meant that market forces could not correct themselves. He provided a more thorough examination and challenge to classical economics than Karl Marx had done. Keynes believed that government intervention is needed because free market economies lead to under-production and under-consumption of certain goods and services, and leave the economy vulnerable to the economic cycle of boom and slumps (economic depressions). He argued that government intervention plays a crucial role in stabilizing the economy during an economic recession (see Chapter 16) and economic interventionist policies can help to maximize social benefits.

◆ **Keynesian economics**, developed by British economist John Maynard Keynes during the 1930s, refers to an interventionist approach to macroeconomic policy by advocating for increased government expenditure and lower taxes in order to stimulate demand in the economy.

While Marx criticized capitalism, Keynes sought the opportunity for the government to justify its existence and the role it has in managing the economy to deal with the fluctuations in business cycles (see Chapter 16). Subsequently, Keynesian economics became considered a 'demand-side' theory that focused on interventionist macroeconomic policies in the immediate and short run to manage the problems of the economy as a whole, such as avoiding the troubles of a recession. In particular, Keynesian economists believe that fiscal policy (the use of taxation and government spending policies) is a more powerful and effective macroeconomic stabilization policy than monetary policy (the use of interest rate policies to control the supply of money in the economy). Nevertheless, Keynes' macroeconomic theories are still used today by governments around the world to manage the economy by using demand-side policies (see Chapters 20 and 21).

CASE STUDY

Coronavirus and fiscal stimulus

The outbreak of the coronavirus in 2019/20 caused governments across the world to adopt interventionist policies to support their economies. The USA, for example, saw unemployment increase to over 30 million people in just 5 weeks, and over 40 million by June 2020. The federal government injected over $2 trillion to support the unemployed and businesses that were forced to close during The Great Lockdown.

In the latter part of the twentieth century, monetarists and the new classical counter revolution (NCCR) refuted Keynes' economic theories. **Monetarists** are economists who believe monetary policy (see Chapter 21) is a more powerful and effective macroeconomic stabilization policy than fiscal policy. Monetary policy is a macroeconomic tool used to influence the overall performance of the economy. This is usually done by changing interest rates to adjust the amount of money in the economy.

◆ **Monetarists** believe that monetary policy is a the most powerful and effective macroeconomic stabilization policy to influence the overall level of economic activity.

Monetarism became a macroeconomic school of thought that emphasizes monetary policy analysis and the role of the money supply in the economy. It is mainly associated with 1976 Nobel Prize-winning American economist Milton Friedman (1912–2006). Friedman argued that poor management of the economy's money supply by the Federal Reserve (America's central bank) was the primary cause of the Great Depression in the 1930s. To monetarists, the money supply is the main determinant of national output (gross domestic product) in the short run and the general price level in the long run. Monetarism gained prominence during the 1970s and helped to combat inflationary pressures due to higher oil prices in the USA and UK. It has also been used by central banks and monetary authorities across the world to stimulate their economies following the global financial crisis between mid-2007 and early 2009.

During the 1980s, in the governments of the USA, UK, Canada and (West) Germany the **new classical counter revolution** (NCCR) came to the forefront. The NCCR economic school of thought favoured supply-side macroeconomic policies (see Chapter 25). This was in response to the interventionist arguments of theorists such as John Maynard Keynes. This led to the popularity of privatization of state-owned enterprises (the sale of public sector assets) and the withdrawal of government regulation of markets.

◆ The **new classical counter revolution** (NCCR) favoured supply-side macroeconomic policies over interventionist arguments, such as privatization of state-owned enterprises and deregulation of markets.

There was also a resurgence of the doctrines of classical economics with economies focusing on free markets and laissez-faire economics. The new classical counter revolutionists use the example of the Asian Tigers (Singapore, Hong Kong, South Korea and Taiwan) who used free market economics and an export-led strategy to propel their economies to wealth and economic prosperity. By contrast, many African and Latin American economies did not prosper despite their reliance on interventionist macroeconomic policies. Laissez-faire economics, privatization and deregulation (the removal of rules, restrictions and regulations in an industry, thereby promoting competition and efficiency) continued to gain importance into the twenty-first century, but such ideas were questioned for their effectiveness following the economic and social stresses of the global financial crisis (2007–09).

Theory of Knowledge

What is truth in the study of economics? How do we determine truth in the history of economic thought?

Twenty-first century

The twenty-first century has seen increasing dialogue with other disciplines such as psychology and the growing role of behavioural economics. In addition, the growing importance of globalization and the push for environmental protection of the planet has led to increasing awareness of the interdependencies that exist between the economy, society and environment.

This has also steered economics in the direction of having an understanding and appreciation of a circular economy, an economic system in which raw materials, components and other resources are used in sustainable ways to generate output of goods and services.

Economics has increasingly been influenced by other disciplines such as psychology and marketing (from the field of business management). This has subsequently popularized the growing role of behavioural economics (see Chapter 6) and its impact on public policy. Behavioural economics applies the understanding of human behaviour by psychologists and marketers to economic analysis. Richard Thaler, a pioneer of behavioural economics, called this 'economics done with strong injections of good psychology'. Behavioural economists argue that people do not always behave in a rational way that is in their own best interest. Behavioural economics provides a framework to create opportunities to nudge people towards wiser, safer and healthier choices. Supporters of behavioural economics argue that it makes economics more accurate and realistic about how humans behave and their decision-making.

Economics of the twenty-first century has also seen an increasing awareness of the interdependencies that exist between the economy, society and the environment. As Kofi Annan (1938–2018), former Secretary-General of the United Nations, once said *'We need to think of the future and the planet we are going to leave to our children and their children'*. Essentially, economists are more aware that sustainable development requires consideration of economic sustainability (profit) for the economy, social sustainability (people) for the society, and ecological sustainability (planet) for the environment. By contrast, over-production can cause pollution and waste, so is therefore not sustainable.

In economics, **sustainability** is the ability to conduct economic activity indefinitely. This means whatever the economy does today does not compromise the opportunities for future generations. For example, the sustainable use of the Earth's limited natural resources means that the current level of economic activity does not jeopardize the resources available for future generations. By contrast, overfishing and deforestation are examples of unsustainable economic activities.

Theory of Knowledge

Some economists argue that economics as a social science is still in its infancy, and that with time as empirical testing methods and the quality of data improve, it will become more reliable in making accurate predictions. Do you agree with this statement?

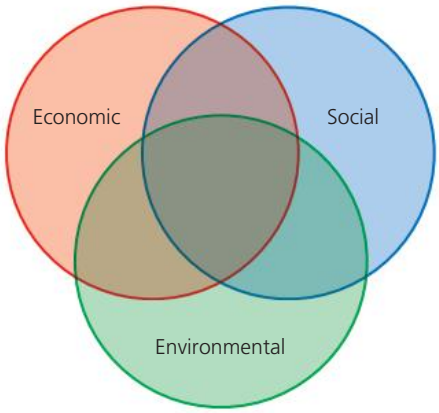


Figure 2.13 Aspects of sustainable development

Figure 2.14 Deforestation is not a sustainable activity

◆ **Sustainability** is a key concept referring to an economy achieving the economic goals that allow an increase in living standards without jeopardizing the needs and wants of future generations.

This increasing awareness of the interdependencies that exist between the economy, society and environment also led the United Nations to upgrade their eight Millennium Development Goals to the Sustainable Development Goals (see Chapter 33).

Sustainability requires firms to be more efficient and economical with the use of the planet's scarce resources, such as by reducing, reusing and recycling. It may also involve the use of green technologies, renewable energy sources and recycled materials in production processes. This has led to the compelling reasons for moving towards a circular economy, rather than the traditional 'take–make–waste' approach of the linear economy.

Figure 2.15 The circular economy

A **circular economy** describes an economic system in which raw materials, components and other resources are used sustainably to generate output. The model is underpinned by a transition to green and renewable energy sources in order to build natural, economic and social capital. It is based on the principle of manufacturing in such a way that consumption and resource usage enable recycling, reuse and regeneration in the manufacturing cycle. In the long-term, this means economic opportunities also provide both environmental and societal benefits.

◆ A **circular economy** describes an economic system in which raw materials, components and other resources are used sustainably to generate output.

PAPER 1 EXAM PRACTICE QUESTION 2.1

a Explain the significance of normative and positive economics. *[10 marks]*

b Explain the importance of the *ceteris paribus* assumption for economic models and theories. *[10 marks]*

c Explain the difference between *equity* and *equality*. *[10 marks]*

Theory of Knowledge

In the era of 'fake news' spread on the Internet, what criteria should be used for evaluating knowledge claims in economics?

Chapter summary

- Economic methodology studies the evolving processes, practices and principles that underpin the discipline as a social science.
- John Neville Keynes categorized economic methodology as either positive or normative economics.
- Positive economics is about the way the economy actually works, based on statements of 'what is', backed by empirical evidence.
- Normative economics is about how the economy should work, based on statements of 'what should be'.
- Logic is integral to the study of positive economics as it is based on rationality and reasoning, rather than emotions or value judgements.
- As part of economic methodology, economists use hypotheses, models and theories to explain real-world phenomena. These can be supported or refuted by empirical evidence.
- As economics is not an exact science, the ceteris paribus assumption is used when examining the behaviour of individuals and societies.
- Economists are interested in both equity (fairness in the distribution of scarce resources) and equality (equal recognition of people).
- Economic thought studies the evolving schools of thought of economics from a historical perspective, from classical economics of the eighteenth century to behavioural economics and emphasis of the circular economy in the twenty-first century.

WISE ChoICES

To what extent has **change** impacted on **sustainability**? You may choose to investigate this in the context of economic activity for an economy of your choice.

Theory of Knowledge

To what extent does the Persian Proverb *'Doubt is the key to knowledge'* apply to economics?

Microeconomics

3 Demand

Real-world issue

How do consumers and producers make choices in trying to meet their economic objectives?

Conceptual understandings (WISE ChoICES)

- Consumer and producer **choices** are the outcome of complex decision-making.
- Welfare is maximized if allocative **efficiency** is achieved.
- Constant **change** produces dynamic markets.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- the law of demand, that is, the relationship between price and quantity demanded (AO2)
- the income and substitution effects, and the law of diminishing marginal utility, and explain the assumptions underlying the law of demand (AO2) (HL only)
- the downward sloping demand curve diagram (AO4)
- the relationship between an individual consumer's demand and market demand (AO2)
- the non-price determinants of demand: income, tastes and preferences, future price expectations, price of related goods (in the cases of substitutes and complements), and the number of consumers (AO2)
- movements along the demand curve and shifts of the demand curve (AO2)
- a demand curve diagram that shows movements along the demand curve and shifts of the demand curve (AO4).

The law of demand (AO2)

Demand is the quantity of a good or service that customers are *willing* and *able* to buy at given prices in a particular period of time. It is sometimes referred to as **effective demand** to distinguish It from a want or a desire to buy something. Only when a person has the desire (willingness) to buy a product that is also backed up by the capability (ability) to pay, do economists speak of demand. For example, a person may want to buy a luxury sports car, but without enough real purchasing power then that person does not have effective demand.

◆ **Demand** is the quantity of a good or service that customers are both willing and able to buy at a given price, per period of time.

Price refers to the amount of money a customer pays in order to purchase a good or service. The **law of demand** states that the quantity demanded for a good or service falls as its price rises, ceteris paribus. In other words, the higher the price of a product, the lower its demand tends to be. Likewise, the quantity demanded will tend to rise at lower prices. This relationship is represented using a demand curve diagram (see Figure 3.2).

◆ The **law of demand** states that the quantity demanded of a product will fall if the price increases, and vice versa.

ATL ACTIVITY 3.1

The law of demand applies to almost all goods and services. It is rational to assume that customers will want to buy larger quantities of goods or services at lower prices. However, this might not always be the case. With a partner discuss why might customers buy *more* when price increases and *less* when the price drops.

Common mistake

'Demand is the amount of a good or service that customers want to buy at each given price level.'

Demand is the quantity of a good or service that customers are *willing* and *able* to buy at given price.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 3.1

Define the term *law of demand*. *[2 marks]*

◆ The **income effect** is one of three factors accounting for the downward sloping demand curve. As the price of a product falls, consumers' real incomes increase so they are able to buy more goods and services at lower prices.

◆ The **substitution effect** states that as the price of a product falls, more people can buy the product, so choose this over rival products, that is, it causes consumers to replace higher priced products with lower priced ones.

Assumptions underlying the law of demand (HL only)

There are three causes of the negative relationship between price and quantity demanded. These factors form the assumptions underlying the law of demand:

- The **income effect** – as the price of a product falls, the real income of customers increases. In other words, they are able to buy more products at lower prices. In other words, real income increases when prices fall, as customers are able to purchase more goods and services with the income they currently have. By contrast, an increase in the price of a good or service reduces people's real disposable incomes, which in turn reduces the quantity demanded.
- The **substitution effect** – as the price of a good or service falls, more customers are able to buy the product, so choose this over rival or substitute products that they might have previously bought. Essentially, the substitution effect causes consumers to replace higher priced products with lower priced ones.
- The law of diminishing marginal utility – as individuals consume more of a particular good or service, the utility (return or satisfaction) gained from the additional unit of consumption declines, so customers will only purchase more at lower prices. This is because the demand for a good or service is dependent on the utility derived from consuming it, which diminishes from the consumption of each successive unit of the product. At some point, consumers will not want any more, because the marginal utility of consumption drops to zero.

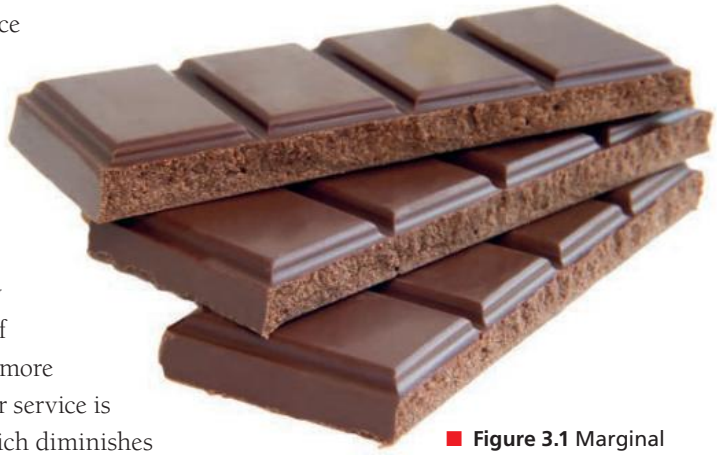


Figure 3.1 Marginal utility falls as the consumption of chocolates increases

Demand curve (AO4)

The demand curve illustrates an inverse correlation between the price and quantity demanded of a product over a certain period of time (see Figure 3.2). It shows how much of a product is demanded at different prices. The amount of a good or service demanded at each price level is called the *quantity demanded*.

When constructing a demand curve diagram, *price* is measured along the vertical (y) axis, and *quantity demanded* is measured on the horizontal (x) axis. It is important to include the correct units of measurement too, such as '$' for price and 'kilograms' for quantity demanded. In economics, it is conventional to draw the demand curve as a linear downward sloping line.

Figure 3.2 shows that as the price of carrots increases from P_1 to P_2, the quantity demanded falls from Q_1 to Q_2 as fewer people are willing and able to pay the higher price. The opposite is true for a reduction in price. In other words, the quantity demanded will increase as the price of a product falls, ceteris paribus.

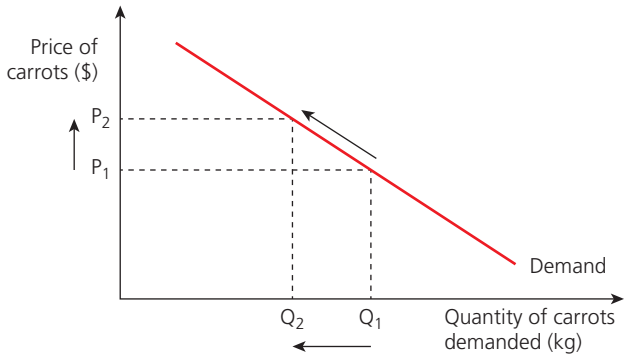


Figure 3.2 The demand curve

Common mistake

Be careful when using the terms 'demand' and 'quantity demanded' as these are often confused by candidates. *Demand* refers to an entire demand curve, whereas *quantity demanded* refers to a point on the demand curve.

Relationship between an individual consumer's demand and market demand (AO2)

A **market** is any place where transactions take place between buyers and sellers; for example, shares are traded in a stock market and currencies are traded on the foreign exchange market. Scarcity must exist for a market to occur. The **market demand curve** refers to the sum of all individual demand for a product at each price level. It is found by adding up all individual demand at each price level (see Figure 3.3). For example, if a cinema charges $15 for movie tickets and the demand from male customers is 500 and 400 from female customers per week, then the market demand (D_{m+f}) at this price is 900 cinema tickets per time period.

◆ A **market** is any place where transactions take place between buyers and sellers.

◆ The **market demand curve** refers to the sum of all individual demand for a product at each price level.

Figure 3.3 The market demand curve

Common mistake

'The market demand refers to demand for a particular good or service in a particular market.'

This is a very imprecise definition. The **market demand** refers to the sum of all individual demand for a particular product at each price level.

◆ **Non-price determinants of demand** are the various factors other than the price of the good or service that affect the demand for the product.

Non-price determinants of demand (AO2)

Although price is regarded as the key determinant of the level of demand for a good or service, it is not the only factor that affects the quantity demanded of a product. **Non-price determinants of demand** are the various factors other than the price of the good or service that affect the demand for the product. These factors change the demand (shift the demand curve) for a good or service.

◆ **Normal goods** are products that customers tend to buy more of as their real income level increases. This includes both necessities and luxury goods and services.

Non-price determinants of demand include: income, tastes and preferences, future price expectations, the price of related goods (in the cases of substitutes and complements) and the number of consumers.

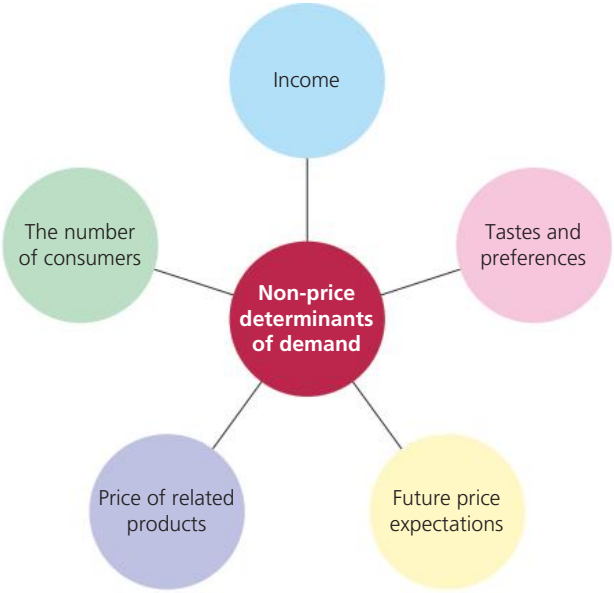


Figure 3.4 Non-price determinants of demand

Top tip!

The non-price determinants of demand can be remembered by the acronym **RIPEN**: **R**elated products (substitutes and complements), **I**ncome, **P**references and tastes, **E**xpectations of future prices, and the **N**umber of consumers.

Income

Higher levels of real income mean that customers are able and willing to buy more goods and services, owing to their greater purchasing power, ceteris paribus. For example, the average person in the USA will have a higher level of demand for goods and services compared to the average person in Algeria or Vietnam. Another example is that in economically developed societies with rising average real incomes, the demand for smartphones has risen steadily. This would be shown diagrammatically by a rightwards shift of the demand curve for smartphones. Products that have higher demand when real income increases, such as smartphones, are known as **normal goods**. In the same way, if an individual's income falls, so will demand. This would be represented by a leftwards shift of the whole demand curve.

Note that there is one exception to this rule – in the case of **inferior goods**, an increase in real income will lead to a fall in the demand. For example, individuals with low incomes are less likely to demand fresh meat and more likely to buy canned or processed meats. However, if they earn higher real incomes, demand for canned and processed meats is likely to fall, while the demand for fresh meats is likely to rise. Hence, fresh meats are considered as normal goods and canned or processed meats are classified as inferior goods.

◆ **Inferior goods** are products with a negative income elasticity of demand. In other words, the demand for such products tends to fall when consumers' real incomes rise.

Tastes and preferences

Changes in people's tastes and preferences can affect the demand for all types of goods and services. Products that become fashionable (such as the latest smartphones and smartwatches) lead to an increase in demand, whereas those that become unfashionable (such as last season's clothes) will cause a fall in the level of demand.

Tastes and preferences can change over time, so demand increases or decreases to reflect these changes. For example, smoking cigarettes is far less fashionable than it used to be. Hence, the demand for tobacco products has fallen over time. Similarly, the demand for sugary drinks has declined due to health concerns and changes to people's (healthier) lifestyle habits.

Consumers may have a preference for certain brands, so this can have a significant influence on how consumers spend their income. Branding and advertising aim to increase the emotions and desires attached to a product. Successful brands give consumers confidence and minimize the risks associated with purchasing decisions. Companies with strong brand loyalty include Apple, Nike and McDonald's. Hence, the demand for their products is relatively high as customers have a greater preference for these brands.

Theory of Knowledge

Some products, like alcohol, gambling and cigarettes, can be highly addictive. The stronger the addiction, the greater the level of demand for these products. To what extent does addiction to a product interfere with rational decision-making in economics? Do individuals forgo their own self-interest when making such decisions?

Future price expectations

Consumer expectations of prices in the future can influence the demand for many products, especially in the case of expensive and luxury goods, such as consumer electronics products, motor vehicles and housing. If people expect that the value of a good or service will increase in the foreseeable future, they will demand more of it now, ceteris paribus. So, for example, if people expect property prices to rise next year, they may buy now in anticipation that prices will rise in the future. There is greater demand because customers can expect to make a **capital gain** on the property purchase (the positive difference between the price the buyer paid for the property and its market value), and the lack of affordability should property prices increase in the future. Similarly, there will be less demand in the market for stocks and shares if people expect prices will drop in the near future.

Price of related goods (substitutes and complements)

The price of related goods means that the sales of one product affect the demand of another product. Related products can be classified as complements or substitutes.

Complements (or **complementary goods**) are products jointly demanded as they go well together. Examples include cars and petrol, indoor plants and plant pots, scarves and gloves, Doritos and salsa dip, or wedding dresses and wedding rings.

◆ **Complements** (or **complementary goods**) are products that are jointly demanded; for example, torches and batteries, or printers and ink cartridges.

As complements are in joint demand, when there is a change in the price of one product (such as a games consoles), it negatively affects the demand for the related good (the games for the console). For example, an increase in the price of cinema tickets is associated with a decline in the demand for popcorn sales.

Figure 3.5 Nuts and nut crackers are complementary goods

This relationship is shown in Figure 3.6. In the market for smartphones, an increase in supply (perhaps due to technological progress, which reduces the costs of production) results in a price falling from P_1 to P_2. The law of demand states that the quantity demanded will increase from Q_1 to Q_2. As more customers buy smartphones, the demand for downloadable applications will also increase. Note that the demand curve for downloadable apps shifts from D_1 to D_2 as this was caused by a change in a non-price factor affecting the demand for downloadable apps (it was the price of smartphones that changed). Hence, in the case of complements, a fall in the price of one product (smartphones) leads to an increase in the demand for the complementary product (downloaded applications).

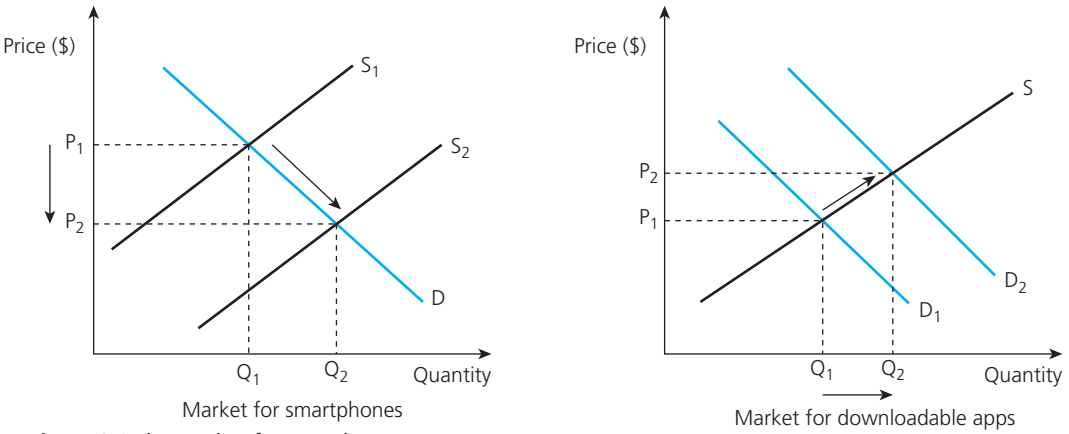


Figure 3.6 The market for complements

Top tip!

When providing examples of complementary goods, remember that it is possible to have more than just two products that are jointly demanded. For example, when purchasing tickets for the cinema or a theatre show, it is quite common for customers to also buy popcorn, hotdogs, sweets (candy), drinks and other snacks.

Figure 3.7 Popcorn and drinks are complementary goods at the cinema

Substitutes are products that are in competitive demand. This means that they are rival products that can be used as alternatives to each other. Examples of goods in competitive demand include Apple iPhone and Samsung Galaxy smartphones, Nike and Adidas sports apparel, Coca-Cola and Pepsi Cola, or BMW, Mercedes Benz and Audi motor vehicles. Firms compete for customers who would ordinarily buy from only one of the suppliers.

◆ **Substitutes** are products that are in competitive demand because they can be used in place of each other; for example, butter and margarine, tea and coffee, or Honda and Toyota cars.

Figure 3.8 Coffee and tea are substitute products

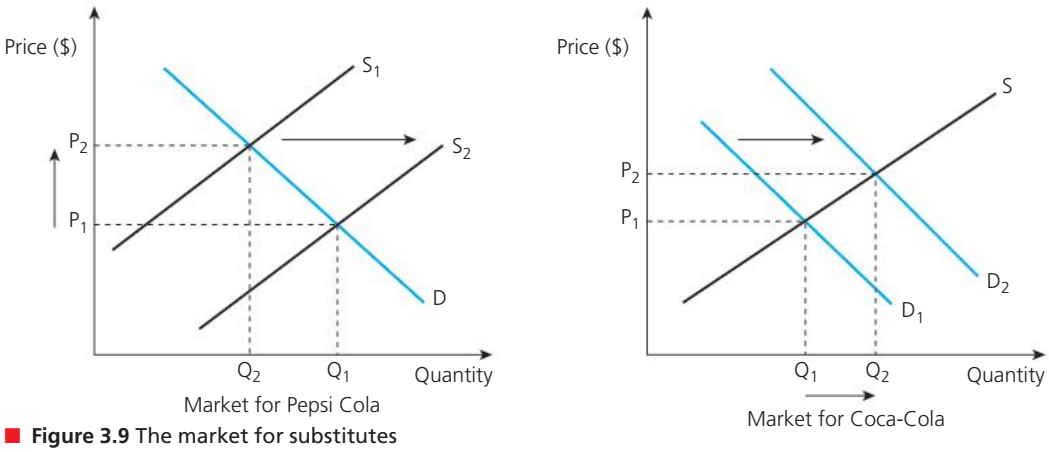


Figure 3.9 The market for substitutes

As substitutes are in competitive demand, when there is a change in the price of one product (such as Pepsi Cola), it positively affects the demand for the substitute good (Coca-Cola). For example, an increase in the price of Pepsi Cola (perhaps due to higher production costs) causes an increase in the demand for Coca-Cola, as customers switch to the relatively cheaper brand, ceteris paribus.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 3.2

Explain the likely relationship between the demand for private and public transportation. *[4 marks]*

PAPER 1 EXAM PRACTICE QUESTION 3.3

Distinguish between *substitutes* and *complementary goods*. *[10 marks]*

Number of consumers

A change in the total number of consumers of a product causes the entire demand curve to shift. As more buyers enter the market, or existing customers buy more, the demand curve shifts to the right. This means there is an increase in demand. For example, the markets for smartphones, tablet computers and electric cars have increased significantly. This is the case even if prices do not change. Similarly, the growth in population size is an important factor influencing the increase in the number of consumers in the market for any given product. For example, the total number of buyers and renters in the property markets in major world cities has expanded due to the growing population, which increases the demand for housing in these locations. Hence, the greater the number of consumers of a good or service, the larger the market demand for the product.

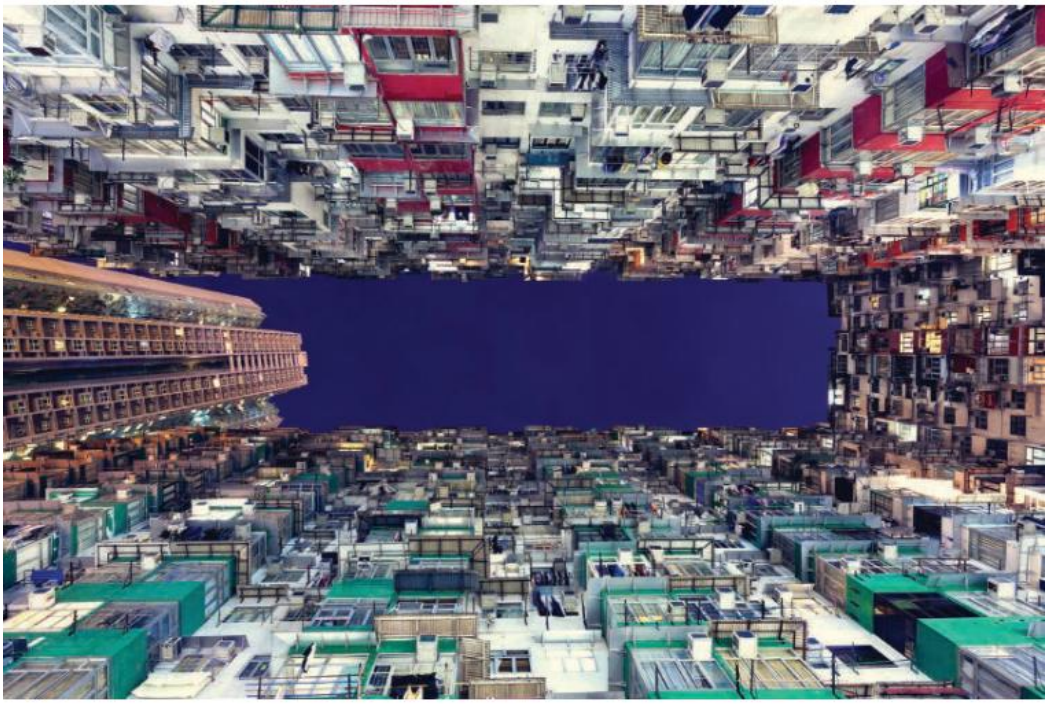


Figure 3.10 The growing demand for housing in Hong Kong increases property prices

Top tip!

Although the IB Economics syllabus states five non-price determinants of demand, other factors can also influence the level of demand, depending on the context of the good or service being considered. For example, important non-price determinants of demand could include the weather and demographics (such as the age, gender, ethnicity or religious beliefs of customers). Although these are not explicitly examined, it can be useful to be aware of this fact to enhance your critical thinking in economics.

WISE ChoICES

Examine how the concepts of **change** and **economic well-being** impact the demand for a product of your choice.

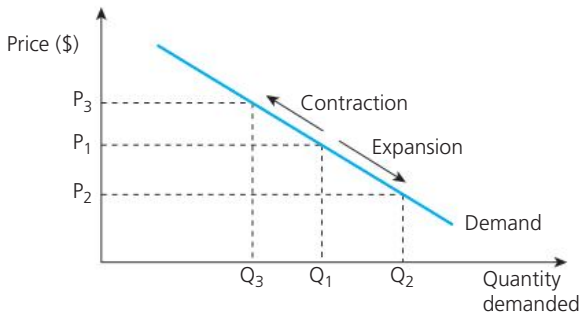
Movements along the demand curve and shifts of the demand curve (AO2, AO4)

Recall that the law of demand states there is an inverse relationship between the price of a product and its quantity demanded. A change in the price of a good or service causes a **movement along the demand curve**. An increase in price causes a **contraction** in the quantity demanded. By contrast, a fall in price causes an **expansion** (or an **extension**) in the quantity demanded (see Figure 3.11).

◆ **Movement along a demand curve** is caused by price changes only. A fall in price causes quantity demanded to expand while an increase in price causes quantity demanded to contract.

In Figure 3.11, a fall in price from P_1 to P_2 causes quantity demanded to expand from Q_1 to Q_2, as per the law of demand. By contrast, an increase in price from P_1 to P_3 will cause the quantity demanded to contract from Q_1 to Q_3 units of output.

While a movement along a demand curve is caused by price changes only, a change in all other (non-price) factors that affect demand will cause a **shift** of the demand curve. For example, BMW recorded higher than expected profits in 2019 owing to increasing demand for its cars in Europe, the USA and Asia. For the first time ever, the German carmaker enjoyed sales of more than 500,000 vehicles in mainland China during the first three quarters of the year. Hence, the demand for BMW's cars was higher at all prices.



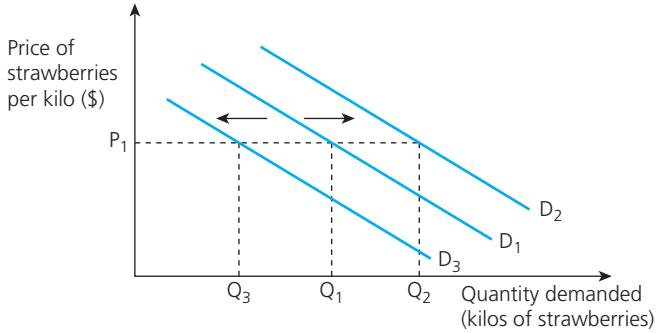
■ **Figure 3.11** Movements along a demand curve

■ **Figure 3.12** China is BMW's largest growth market

◆ A **contraction in demand** occurs when there is a fall in the quantity demanded for a product following an increase in its price.

◆ An **expansion in demand** occurs when there is an increase in the quantity demanded for a product following a fall in its price.

An **increase in demand** (rather than an *increase in the quantity demanded*) for a product is illustrated by a rightwards or outwards shift of the entire demand curve, caused by favourable changes in non-price factors that affect demand. For example, in Figure 3.13, successful advertising campaigns promoting the consumption of strawberries will shift the demand curve from D_1 to D_2. This results in more strawberries being demanded at all prices (in this case, the demand for strawberries at a price of P_1 increases from Q_1 to Q_2 kilograms of strawberries). *Note:* in reality, an increase or decrease in the demand will cause the price to change, rather than to remain at P_1.



■ **Figure 3.13** Shifts in the demand curve for strawberries

By contrast, a **decrease in demand** (rather than a *decrease in the quantity demanded*) is shown by shifting the entire demand curve to the left, from D_1 to D_3, resulting in less quantity demanded at all prices. In Figure 3.13, demand was Q_1 at a price of P_1, but has now fallen to Q_3 owing to an adverse change in a non-price factor affecting demand. For example, a fall in real income is likely to reduce the demand for strawberries, causing the quantity demanded at a price of P_1 to fall from Q_1 to Q3.

◆ An **increase in demand** refers to a rightwards shift of the entire demand curve for a product, caused by favourable changes in non-price factors that affect demand.

◆ A **decrease in demand** refers to a leftwards shift of the entire demand curve for a product, caused by unfavourable changes in non-price factors that affect demand.

◆ A **shift** of the demand curve occurs when there is a change in any non-price factor that affects the demand for a product. The entire demand curve shifts to the left if there are unfavourable changes affecting demand, and vice versa.

Top tip!

Remember the difference between *changes in demand* and *changes in the quantity demanded*. A movement along a demand curve is caused by changes in the price of the product, resulting in *changes in the quantity demanded*. A **shift** in demand is caused by changes in non-price factors that affect demand, resulting in *changes in demand*.

Common mistake

'When the price falls, people have higher real income so there is an increase in demand for a product, ceteris paribus.'

The wording of this statement suggests a shift in demand (caused by changes in non-price factors). The student should have referred to the reduced price causing an increase in the quantity demanded, not an increase in demand.

Common mistake

'A fall in price causes demand to rise.'

This statement also suggests a shift in the demand curve. A fall in price will cause an extension (movement) along the demand curve, ceteris paribus, rather than an increase in demand.

WISE ChoICES

Choose any good or service that you can purchase from a supermarket. Consider how the concepts of **scarcity** and **choice** affect the demand for this product.

ATL ACTIVITY 3.2

Using a relevant diagram for each scenario below, explain the impact on the market for tuna fish. Be prepared to share your responses with the rest of the class.

a An increase in the average household's real income.
b Eating fresh fish becomes increasingly popular as part of a healthy diet.
c The price of fish is expected to fall in the coming weeks.
d The price of trout and salmon falls by about 10 per cent.
e The price of salads, which go well with tuna dishes, falls by around 10 per cent.
f Net migration leads to a noticeable increase in the population size.

Chapter summary

- Demand refers to the *willingness* and *ability* of customers to pay a given price to purchase a good or service.
- The amount of a good or service demanded at each price level is called the *quantity demanded*.
- The *law of demand* states there is an inverse relationship between the price of a product and its quantity. This is caused by the income effect, the substitution effect and the law of diminishing marginal utility.
- The *market demand* refers to the sum of all individual demand for a product at each price level.
- The demand curve is drawn downward sloping from left to right, showing the inverse relationship between price and quantity demanded.
- The non-price determinants of demand can be remembered by the acronym **RIPEN**: related products (substitutes and complements), income, preferences and tastes, expectations of future prices, and the number of consumers.
- Price fluctuations cause *movements* along an existing demand curve.
- A rise in price results in a *contraction* in quantity demanded, whereas a fall in price causes an *expansion* (or *extension*) in the quantity demanded.
- A change in all other (non-price) factors that affect demand will cause a *shift* in demand.
- Diagrammatically, higher levels of demand are shown by a rightwards (outwards) shift of the entire demand curve. By contrast, a leftwards (inwards) shift of the demand curve occurs when there is less demand for the product at all given prices.

4 Supply

Real-world issue

How do consumers and producers make choices in trying to meet their economic objectives?

Conceptual understandings (WISE ChoICES)

- Consumer and producer **choices** are the outcome of complex decision-making.
- Welfare is maximized if allocative **efficiency** is achieved.
- Constant **change** produces dynamic markets.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- the law of supply, that is, the relationship between price and quantity supplied (AO2)
- the assumptions underlying the law of supply: the law of diminishing marginal returns, and increasing marginal costs (AO2) (HL only)
- an upward sloping supply curve diagram (AO4)
- the relationship between an individual producer's supply and market supply (AO2)
- the non-price determinants of supply: changes in costs of factors of production, prices of related goods, indirect taxes and subsidies, future price expectations, changes in technology, and the number of firms (AO2)
- movements along and shifts of the supply curve (AO2)
- a supply curve diagram that shows movements along and shifts of the supply curve (AO4).

◆ **Supply** is the amount of a good or service that firms are *willing* and *able* to provide at any particular price, per time period.

◆ The **law of supply** states that there is a direct (or positive) relationship between quantity supplied and price, ceteris paribus.

The law of supply (AO2)

Supply refers to the quantity of goods or services that firms are *willing* and *able* to sell at any given price, per time period.

The **law of supply** states that there is a positive correlation between the quantity supplied of a product and its price, ceteris paribus. Higher prices will tend to increase the quantity supplied, because more firms are able and willing to supply their products owing to higher profit margins. In other words, the potential for greater profitability occurs with an increase in price. The law of supply only holds if price changes while all other factors affecting supply remain unchanged.

Figure 4.1 A fresh fruits supplier

The opposite exists if price falls. In other words, there is less quantity supplied when price falls because profit margins fall. Diagrammatically, the law of supply is shown by an upward sloping supply curve (see Figure 4.3).

There are two reasons for the positive relationship between price and quantity supplied:

- Existing firms in the market can earn higher profit margins if they supply more when the price increases. Hence, higher prices create an incentive for existing firms to supply more.
- New firms enter the market, attracted by the higher prices and profit margins, which enable them to cover their costs of production and the potential to earn profits.

Hence, as the price of a product increases, firms are likely to expand production so that there is an increase in the quantity supplied, ceteris paribus.

Assumptions underlying the law of supply (HL only) (AO2)

The law of supply suggests a higher price creates greater incentives for existing firms to increase output and for new firms to enter the market, thus increasing the market supply of a product, ceteris paribus. There are two key assumptions behind this law (or general rule): the law of diminishing marginal returns (DMR) and increasing marginal costs.

The law of diminishing marginal returns

The **law of diminishing marginal returns** (DMR) describes how output is affected when a firm uses more variable inputs (factors of production) while maintaining at least one factor of production as fixed in the **short run**. For example, a supermarket can hire more labour (the variable input) while the number of checkouts and the physical retail store remains the same (the fixed factor). The law of DMR states that by employing additional variable factors of production (such as labour), the marginal returns (additional output) for each additional unit of input will eventually decline. The law of DMR applies only in the short run because all factors of production are variable in the **long run**.

Diminishing marginal returns occur in the short run, when at least one factor of production is fixed. It is usually assumed that capital is fixed in the short run. This is because a firm's real or physical capital (such as buildings, machinery, equipment and vehicles) does not usually change in a short period of time, so is kept fixed when demonstrating this law (or general principle).

At first, the addition of extra workers creates benefits for the firm as it becomes teamwork and the division of labour become possible, thereby improving efficiency and productivity. Hence, initially, there are increasing marginal returns to each additional worker. However, the marginal returns will eventually diminish because each additional worker will gradually contribute less than the one before as work cannot be efficiently divided so well. Also, the additional workers will eventually get in each other's way thereby diminishing the productivity and effectiveness of each extra worker. The diminishing marginal returns from each successive worker leads to higher production costs, meaning that firms require a higher price to create incentives to produce.

Top tip!

Students sometimes confuse theories, tools and models used by economists, such as the law of diminishing marginal returns (DMR) and diminishing marginal utility (DMU). A useful way for you to remember this distinction is that DMR is used when referring to production (or output), while DMU is used when referring to consumption.

ATL ACTIVITY 4.1

The law of supply applies to almost all goods and services. It is rational to assume that firms will want to produce and sell greater quantities of whichever product earns them the most profit. With a partner discuss when might this *not* be the cause, that is, as price increases, individuals or firms supply less.

◆ The **law of diminishing marginal returns** explains what happens to the output of products when a firm uses more variable inputs while keeping a least one factor of production fixed. It states that by employing additional variable factors of production, the marginal returns will eventually decline.

◆ The **short run** is a period of time when at least one factor of production is fixed (such as capital), whereas other factor inputs are variable (such as labour).

◆ The **long run** is a period of time when no factors of production are fixed, hence, all factor inputs are variable.

Theory of Knowledge

Can laws in economics, such as the law of diminishing marginal returns, have the same status as laws in the natural sciences?

Top tip!

The law of DMR applies to your revision too! Revising over extended periods of time, and very late at night (to keep you sleep deprived) can be counterproductive. If you revise economics for consecutive 30-minute periods, you will improve your knowledge quite a bit in the first few sessions. However, if you continue to revise for additional 30-minute blocks, eventually the amount that you learn will increase by only a very small amount (if at all) because you become too tired.

Figure 4.2 Studying for prolonged periods of time can be counterproductive

Increasing marginal costs

Decisions to supply a good or service are largely determined by the marginal cost of production. **Marginal cost** refers to the cost of producing an additional unit of output. In general, marginal costs rise with each successive unit produced owing to diminishing marginal returns. Hence, firms are willing and able to increase production only if they receive a higher price for the additional units of output. Therefore, the principle of increasing marginal costs underpins the law of supply (the direct relationship between price and quantity supplied). Essentially, the supply curve slopes upwards because higher prices are needed to cover the higher marginal costs of production when a firm increases its output. In the same way, if the market price falls, the costs of producing the additional unit of output will be greater than the price received by firms, meaning they will cut back supply.

◆ **Marginal cost** refers to the cost of producing an additional unit of output.

The supply curve (AO4)

The supply curve for a product is shown diagrammatically by an upward sloping linear line (see Figure 4.3). As with the demand curve, the price of the product is shown on the *y*-axis, with the unit of measurement being a currency (usually $), and the *x*-axis measuring the quantity supplied.

In Figure 4.3, an increase in the price of tuna fish from P_1 to P_2 creates an incentive for firms to spend more time and effort to catch or farm tuna fish. Hence, the supply of tuna increases from Q_1 to Q_2 tonnes.

Similarly, in the labour market, people are more willing to offer (supply) their labour or services at higher wage rates. Hence, as wages increase the supply of labour will also increase due to individuals willing to spend more time working and more workers wanting to join the profession.

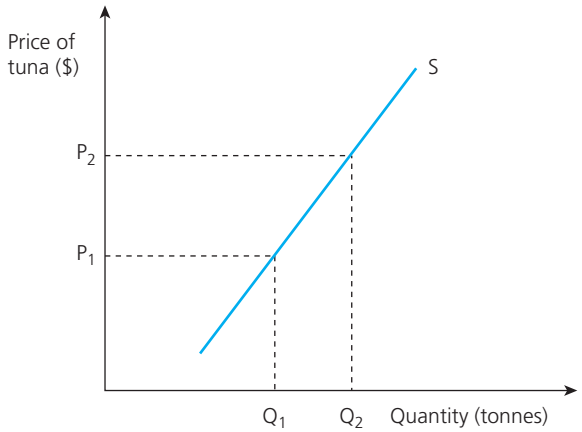


Figure 4.3 The supply curve

Relationship between an individual producer's supply and market supply (AO2)

The **market supply** curve refers to the sum of all individual supply of a product at each price level. It is found by adding up all individual supply of producers at each price level (see Figure 4.4). For example, if the market price of organic grapes is $10 per kilogram and Firm 1 is able and willing to supply 200 kilograms while Firm 2 supplies 300 kilograms, the market supply (S_m) is then 500 kilograms of organic grapes per time period.

◆ **Market supply** refers to the sum of all individual supply of producers at each price level for a given product.

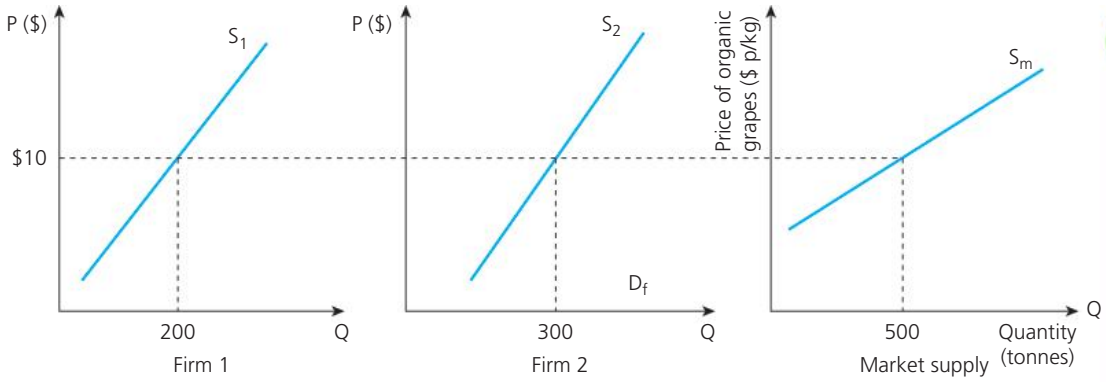


Figure 4.4 The market supply curve

Top tip!

The market supply curve is flatter than the supply curves for individual firms. This is because a change in price causes a more than proportionate change in the market supply curve, whereas individual firms have less flexibility to respond to price changes in the short run.

Non-price determinants of supply (AO2)

Although price is regarded as the key determinant of the level of supply for a good or service, it is not the only factor that affects the quantity supplied of a product. **Non-price determinants of supply** are the various factors other than the price of a good or service that affect the supply of the product. These factors change the supply (shift the supply curve) of a good or service. The non-price determinants of supply include: changes in costs of factors of production, prices of related goods (joint and competitive supply), indirect taxes and subsidies, future price expectations, changes in technology, and the number of firms in the industry.

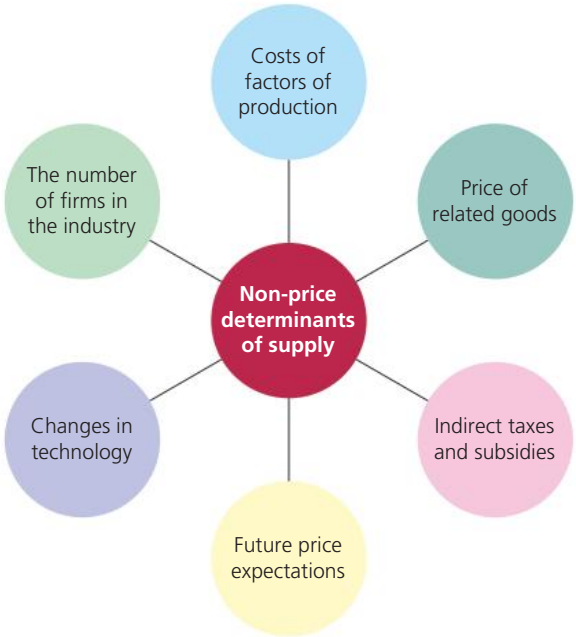


Figure 4.5 Non-price determinants of supply

Top tip!

A useful way to remember some of the key non-price determinants of supply is to use the acronym **CISTERN**: **C**osts of production of factors of production; **I**ndirect taxes; **S**ubsidies; **T**echnological change; **E**xpectations of future prices; prices of **R**elated products (joint supply or competitive supply); and the **N**umber of firms in the industry.

Changes in costs of factors of production

Any change in the costs of production or a change in the costs of factors of production (FOPs) will shift the supply curve of the product being considered. If the cost of the FOPs increases (perhaps due to a rise in wages and/or an increase in rent costs), the supply curve will shift to the left, ceteris paribus. This is because the higher costs of production mean that existing firms cannot produce the same quantity as before and fewer firms are willing and able to supply output. Hence, less will be supplied at each price level, as costs of production rise. The opposite is true if there is a fall in the costs of FOPs.

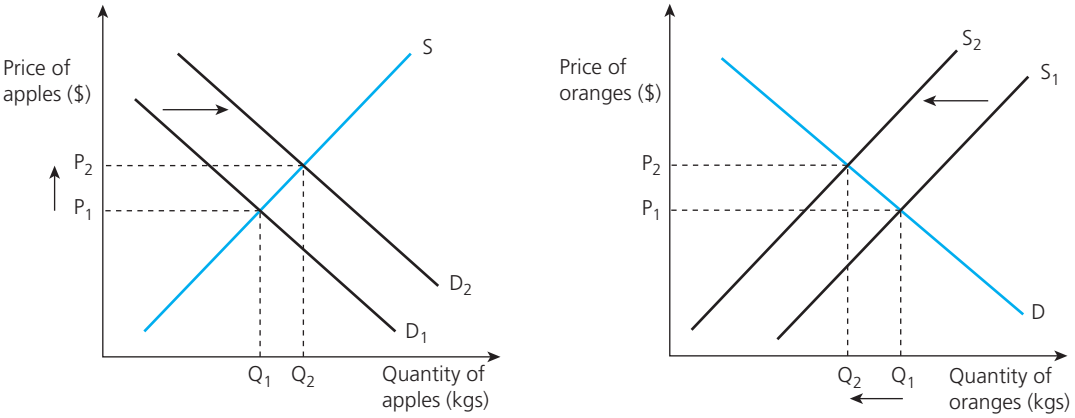
◆ **Non-price determinants of supply** are the various factors other than the price of a good or service that affect the supply of the product.

Prices of related goods – joint and competitive supply

The price of related goods also affects an individual producer's supply. Some products are in competitive supply while others may be in joint supply. In the case of **competitive supply**, the output of one product (such as apples) prevents or limits the output of alternative products (such as oranges). This is due to competing resources, as producers have limited resources such as land and labour, so cannot supply more of one product without producing less of the other. The relative price

◆ **Competitive supply** means the output of one product prevents or limits the production of alternative products, due to competing resources.

and profitability of the two products determines the level of supply for each product. For example, if the price of apples increases by 10 per cent while the price of oranges remains the same, farmers may allocate more land and other resources to harvesting apples than oranges. Hence, the increase in the price of apples reduces the supply of oranges, ceteris paribus (see Figure 4.6).



■ **Figure 4.6** Competitive supply

In Figure 4.6, higher demand for apples (as shown by the shift of the demand curve from D_1 to D_2) causes an increase in the price of apples from P_1 to P_2. This causes an extension along the supply curve as more producers supply apples, increasing the quantity traded from Q_1 to Q_2. The relatively higher price and improved profitability of supplying apples causes a leftwards shift of the supply curve for oranges from S_1 to S_2 as these and other producers switch to supplying apples instead.

In the case of **joint supply**, an increase in the production of one product automatically increases the supply of at least another (joint) product, such as cows, milk and leather or lamb and wool. The less important product in joint supply is called the **by-product**, such as mutton being the main product from sheep and the by-products being wool, milk and cheese. When the supply of the main product changes, the supply of all its by-products is affected in the same way (see Figure 4.7). For example, if the supply of chickens drops dramatically, the supply of chicken eggs would also fall sharply.

◆ **Joint supply** refers to the supply of a product that results in the output of at least one by-product.

■ **Figure 4.7** Chickens and eggs are in joint supply

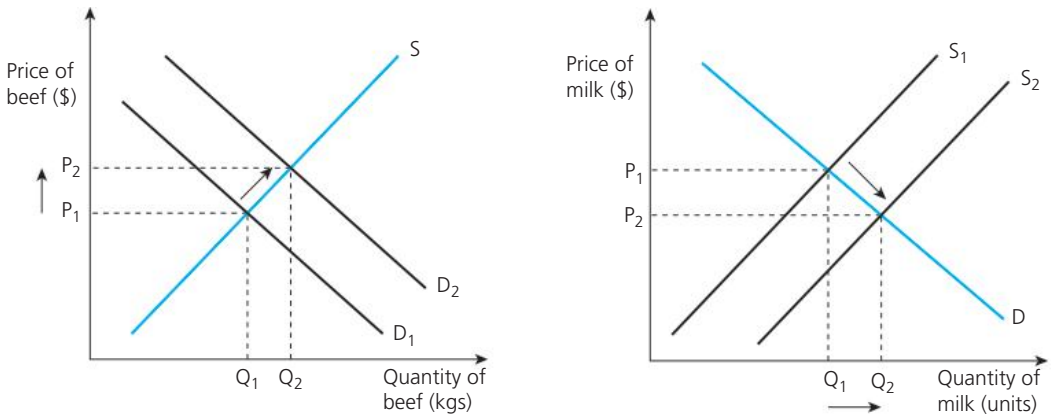


Figure 4.8 Joint supply

In Figure 4.8, increased demand for beef products (shown by the shift in demand from D_1 to D_2) has led to an expansion in the supply of cows, from Q_1 to Q_2. As the quantity of cows has increased, this also causes a corresponding increase in the supply of milk.

Common mistake

'Rising demand for a product will cause an increase in the supply of the product, which will probably mean lower prices.'

Rising demand would cause price to increase, which could lead to an expansion in supply, an *increase in the quantity supplied* (not an *increase in supply*). This does not lead to lower prices (as it was the higher demand that caused prices to rise).

Indirect taxes and subsidies

Indirect taxes are government levies or charges on expenditure, rather than on incomes. They are imposed on goods and services, so inevitably increase the price paid by customers. Indirect taxes are imposed on the supplier of a product, which basically adds to the costs of production, even if the producer can pass on some of the tax to consumers in the form of higher prices. The indirect taxes reduce the profitability of firms, so tend to reduce market supply. This is illustrated by an inwards (leftwards) shift of the supply curve at each price level.

◆ **Indirect taxes** are government levies on expenditure, rather than on incomes. Essentially, these taxes add to the producer's costs of production.

Common mistake

'An increase in taxes causes a fall in supply, ceteris paribus.'

This statement is not entirely true, as it is a change in indirect taxes that causes a fall in quantity due to higher costs of production. A change in direct taxes imposed on wages and salaries, for example, would shift the *demand* curve as households would have less disposable income.

Subsidies are a form of financial assistance from the government to help encourage output or supply, by reducing the costs of production. Hence, subsidies shift the supply curve of a product to the right at each price level. Subsidies are usually given to producers or providers of goods or services deemed to be beneficial to society, such as the production of agricultural products or the provision of education, training and healthcare services.

Figure 4.9 Subsidies are provided to schools and universities to encourage greater supply of educational services

Future price expectations

Price acts as a signal to producers to allocate their resources to the production of provision of goods and services with greater profitability. Thus, expectations of price changes in the foreseeable future can shift the supply curve. If sellers expect the demand for their product to increase sharply in the near future (thus causing higher prices), then the firms may choose to increase production in the current time period. For example, if property developers anticipate future price increases, they will seek to increase the supply of housing. Similarly, if producers are concerned about future market conditions deteriorating and causing prices to fall, production would fall. In other words, there would be an inwards (leftwards) shift of the supply curve.

Changes in technology

Advances in technology mean that there can be increased supply of goods and services at every given price level. This is because production becomes more efficient, so the average cost of production almost certainly falls. Hence, the supply curve will shift to the right. For example, e-commerce enables businesses to reach customers from across the world, without the need for a physical outlet. Instead, e-commerce firms can operate 24/7 with reduced costs, such as lower rents and reduced labour costs.

Number of firms

The number of firms in an industry is usually dictated by the size of the market. If the market for a good or service increases, then it is likely that there will be an increase in the number of firms in the industry. For example, the market supply of IB Economics textbooks is much larger than that for IB Philosophy or IB Geography, but much smaller than the market supply of IB History or IB Mathematics. This is because there is a significantly larger market for popular or compulsory

subjects, so there are more textbook publishers who supply resources to these markets. Hence, an increase in the number of firms in a market would shift the supply curve outwards to the right. Similarly, as the market size for certain products falls (such as CDs, DVDs, newspapers and diesel cars), the number of firms supplying these products also falls.

Table 4.1 Causes of a change in market supply

Increase in supply	Decrease in supply
Lower production costs	Higher production costs
Introduction or addition of subsidies	Removal or reduction of subsidies
Lower indirect taxes	Higher indirect taxes
Improved production technologies	Obsolete technologies
Technological advances	Technological breakdowns
Expectations of higher prices	Expectations of lower prices
Increased joint supply	Lower joint supply
Lower price of goods in competitive supply	Higher price of goods in competitive supply
Larger number of firms in the industry	Smaller number of firms in the industry

Top tip!

Although the IB Economics syllabus specifies six non-price determinants of supply, other factors can also influence the level of supply, depending on the context of the good or service being considered. For example, important non-price determinants of supply could include time and the weather. Supply tends to be lower in the short run but can increase over time; for example, it is difficult for farmers to increase their supply of crops within a short period of time. The output of some products also depends on the weather, for example, agricultural output. Although these factors are not explicitly examined, it can be useful to be aware that other non-price determinants of supply exist, in order to enhance your critical thinking in economics.

Figure 4.10 Adverse weather conditions can affect supply

ATL ACTIVITY 4. 2

Investigate the factors that affect the supply of coffee beans.

Figure 4.11 What are the main factors that affect the supply of coffee?

ATL ACTIVITY 4.3

Investigate how firms such as Apple, Amazon or Google are using technologies such as artificial intelligence (AI) to improve productivity and supply. Be prepared to share your findings in class.

◆ **Expansion** in the quantity supplied is caused by a higher price for the product, thus causing an upward movement along the supply curve.

◆ A **contraction** in the quantity supplied is caused by a lower price for the product, thus causing a downward movement along the supply curve.

Movements along and shifts of the supply curve (AO2, AO4)

A **movement** along a supply curve occurs if, and only if, the price of the product changes. This causes a *change in the quantity supplied* (see Figure 4.13). As with the demand curve (see Chapter 3), there are two types of movements along the supply curve: expansions and contractions.

There is an **expansion** in supply if price increases (owing to the law of supply). An increase in the price of a product from P_1 to P_3 causes the quantity supplied to expand from Q_1 to Q_3. By contrast, there is a **contraction** in supply if the price of a good or service falls. The fall in price from P_1 to P_2 causes the quantity supplied to contract from Q_1 to Q_2.

By comparison, a **shift of the supply curve** for a good or service is caused by changes in any non-price factor that affects the supply. In other words, there is a *change in supply* (rather than a *change in the quantity supplied*). In Figure 4.14, a rightwards shift of the supply curve from S_1 to S_2 represents an **increase in supply**. This can be caused by any change in a non-price factor that increases supply, such as lower production costs caused by advances in technology or from government subsidies. By contrast, a leftwards shift of the supply curve from S_1 to S_3 shows a **decrease in supply**. This can be caused by changes in non-price factors that reduce supply, such as adverse weather conditions or government taxes imposed on businesses.

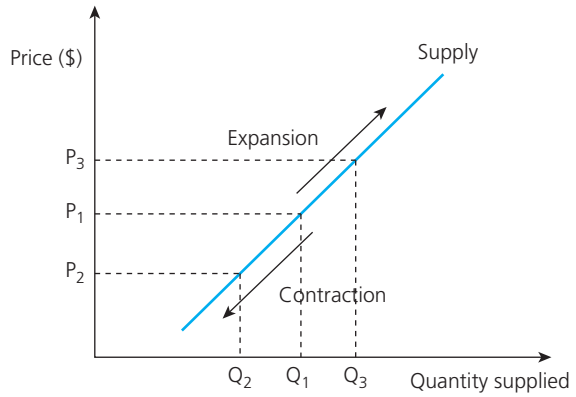


Figure 4.12 Movements along the supply curve

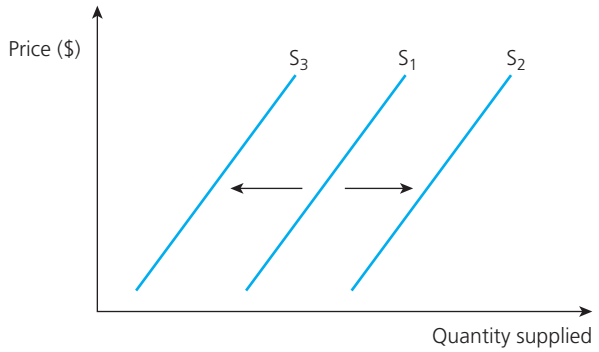


Figure 4.13 Shifts of the supply curve

Top tip!

Make sure you can distinguish between shifts and movements. A *shift* of the supply curve is caused by changes in non-price factors that affect supply (changes in any of the CISTERN factors). A *movement* along a supply curve is caused only by changes in the price of the product.

◆ **Shifts of the supply curve** occur when there is a change in non-price factors that affect supply, resulting in changes in supply. This is represented by a shift of the entire supply curve of the product.

Common mistake

'Supply theory states that higher prices lead to more supply and that lower prices lead to less supply.'

Remember the difference between *changes in supply* and *changes in the quantity supplied*. A movement along a supply curve is caused by changes in the price of the product, resulting in *changes in the quantity supplied.* A shift in the supply curve is caused by changes in non-price factors that affect demand, resulting in *changes in supply*. Therefore, higher prices lead to an increase in the quantity supplied, rather than an increase in supply (or more supply).

Common mistake

Students often claim that the imposition of a *sales tax* on a product shifts its demand curve to the left. This is incorrect as the tax is imposed on firms, thereby raising their costs of production. The tax increases the price of the product so therefore reduces the quantity demanded. In other words, it causes a movement along the demand curve.

PAPER 1 EXAM PRACTICE QUESTION 4.1

Explain the difference between a *shift* of the supply curve for a product and a *movement* along the supply of the product. *[10 marks]*

PAPER 1 EXAM PRACTICE QUESTION 4.2

Explain **two** factors which can cause a firm's supply curve to shift to the left. [*10 marks*]

ATL ACTIVITY 4.4

For each of the questions below, use an appropriate and fully labelled diagram to outline the impact on the market for laptop computers:

a The approval of a per unit subsidy on laptops bought for educational purposes.
b A new government directive requiring new software to be installed on all new laptops to combat piracy and to prevent illegal online activities.
c A 15 per cent government tax imposed on the sale of all foreign laptops.
d An increase in the average wage rate of those working in factories that manufacture laptops.
e An increase in interest rates (think carefully about this one!).

WISE ChoICES

Examine how the concepts of **scarcity** and **sustainability** affect the supply of a product of your choice.

ATL ACTIVITY 4.5

Use an appropriate and fully labelled diagram to explain how the following changes affect the supply or the quantity supplied in each of the following scenarios:

a Droughts across regions of India destroy vegetable crops.
b Investment in new technologies at Tata Motors.
c The Indian government subsidizes the output of hybrid vehicles.
d India increases the tax on cigarettes by 40 per cent.

Chapter summary

- The *law of supply* states that as the price of a good or service increases, firms will supply more of the product, ceteris paribus.
- The supply curve for a product is shown diagrammatically by an upward sloping line, depicting the law of supply. In other words it has a positive slope owing to the positive correlation between price and quantity supplied.
- A higher price creates greater incentives for existing firms to increase output and for new firms to enter the market.
- There are two key assumptions behind the law of supply: the law of diminishing marginal returns (DMR) and increasing marginal costs. (HL only)
- The law of diminishing marginal returns (DMR) states that employing extra variable factors of production (such as labour) eventually causes the marginal returns (additional output) to diminish. (HL only)
- Diminishing marginal returns occur in the short run, when at least one factor of production is fixed. (HL only)
- Higher production costs, caused by diminishing marginal returns, mean that firms require a higher price to have an incentive to produce more. (HL only)
- Firms need to sell their additional output at a higher price in order to cover the higher marginal costs of production, due to diminishing marginal returns. (HL only)
- The supply curve for a product is shown diagrammatically by an upward sloping linear line, with price on the *y*-axis and the quantity supplied on the *x*-axis.
- The market supply curve is the sum of all individual supply of a product at each price level.
- Non-price factors that affect supply can be remembered by CISTERN: changes in costs of production; indirect taxes; subsidies; technology; expectations of future prices; related products (joint supply and competitive supply); and the number of firms.
- A *movement along* a supply curve occurs if the price changes, causing a change in the *quantity supplied*. There is an *expansion* in supply if price of a good or service increases (owing to the law of supply), and a *contraction* in supply if the price falls.
- A *change in supply* occurs when the supply curve shifts, caused by a change in a non-price factor affecting supply. An *increase in supply* in shown by an outwards shift of the supply curve, whereas a *decrease in supply* is shown by an inwards shift.

5 Competitive market equilibrium

Real-world issue

How do consumers and producers make choices in trying to meet their economic objectives?

Conceptual understandings (WISE ChoICES)

- Consumer and producer **choices** are the outcome of complex decision-making.
- Welfare is maximized if allocative **efficiency** is achieved.
- Constant **change** produces dynamic markets.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- a diagram that shows the demand and supply curves forming a market equilibrium (AO4)
- how shifting the demand and supply curves produce a new market equilibrium, with reference to excess demand (shortage) and excess supply (surplus) (AO2)
- a diagram that shows changes in equilibrium/the role of the price mechanism (AO4)
- the functions of the price mechanism: resource allocation (signalling and incentive functions) and rationing (AO2)
- consumer and producer surplus (AO2)
- social/community surplus (AO2)
- a demand and supply diagram that shows consumer surplus and producer surplus being maximized at competitive market equilibrium (AO4)
- allocative efficiency at the competitive market equilibrium (AO2)
- social/community surplus is maximized at equilibrium, where marginal benefit equals marginal cost (AO4)
- consumer surplus and producer surplus from a diagram (AO4) (HL only).

◆ **Market equilibrium** occurs when the quantity demanded for a product is equal to the quantity supplied of the product, so there are no shortages or surpluses.

◆ **Equilibrium** is the condition that holds when a market is cleared of any shortages or surplus. Hence, it occurs at the price where the quantity demanded for a product is equal to the quantity supplied.

Demand and supply curves forming a market equilibrium (AO4)

Market equilibrium is the condition that holds when a market is cleared of any shortage or surplus. Hence, it occurs at the price where the quantity demanded for a product is equal to the quantity supplied. Diagrammatically, this occurs at the point where demand equals supply (D = S), thereby establishing the equilibrium price and quantity traded (see Figure 5.1).

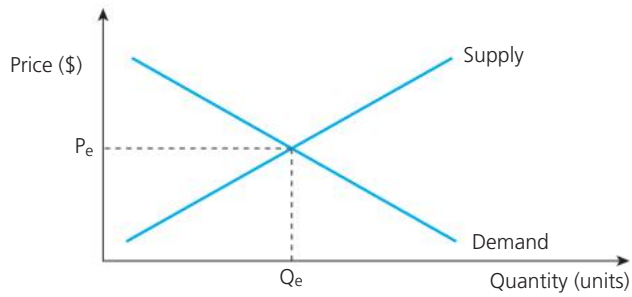


Figure 5.1 Market equilibrium

The **equilibrium price** (or **market clearing price**) is established at the point where the demand for a product matches the supply of the product in the market at a given point in time. Hence, at the equilibrium price, there is neither excess quantity demanded, nor excess quantity supplied. In Figure 5.1, the equilibrium position occurs at point A, where the demand equals the supply of the product on the market, thereby establishing the equilibrium price (P_e) and the equilibrium quantity traded (Q_e).

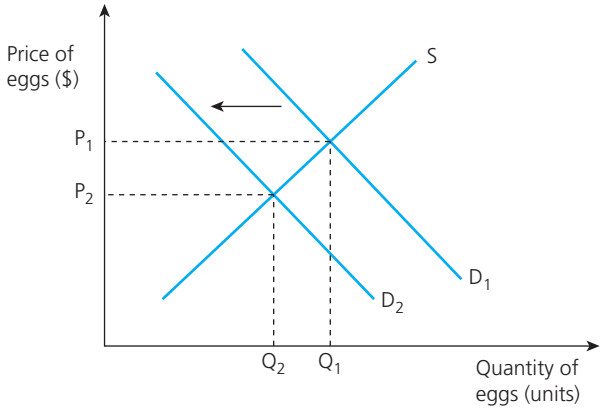
Shifting the demand and supply curves (AO2, AO4)

Recall that changes in any of the non-price determinants of supply or demand will cause a shift in the curve. Hence, a change in any of the non-price factors that affect the demand for or supply of the product will cause a change in the equilibrium price and quantity traded. For example, in Figure 5.2, a health scare linking Salmonella bacteria in eggs shifts the demand curve for eggs to the left from D_1 to D_2, and reduces the quantity traded from Q_1 to Q_2.

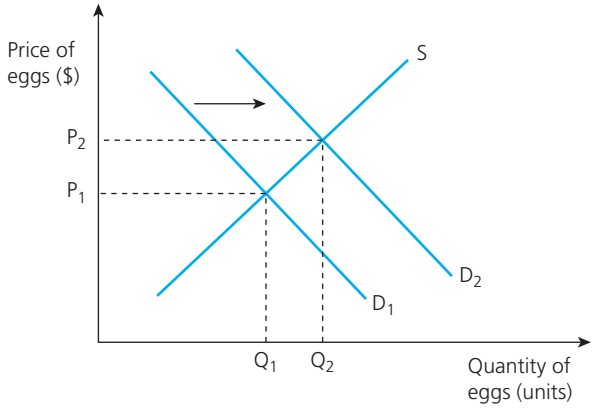
By contrast, successful marketing campaigns that promote eggs as part of a healthy diet will shift the demand curve to the right, as shown in Figure 5.3. In this case, the equilibrium price increases from P_1 to P_2, caused by the higher level of demand for eggs.

Market prices can also change due to changes in non-price factors that shift the supply curve. For example, in Figure 5.4, severe flooding or droughts will shift the supply of agricultural output to the left from S_1 to S_2, thereby increasing the equilibrium price (from P_1 to P_2) and reducing the quantity traded (from Q_1 to Q_2).

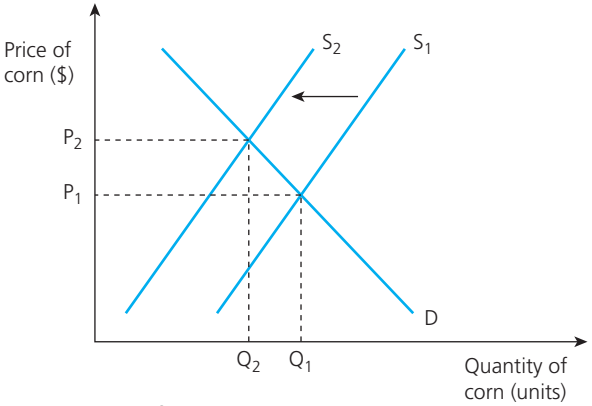
By contrast, technological progress or favourable weather conditions will increase the supply of agricultural output. As shown in Figure 5.5, this would lead to a rightward shift of the supply curve (from S_1 to S_2), thereby reducing the market price (from P_1 to P_2) and expanding the quantity demanded (from Q_1 to Q_2).



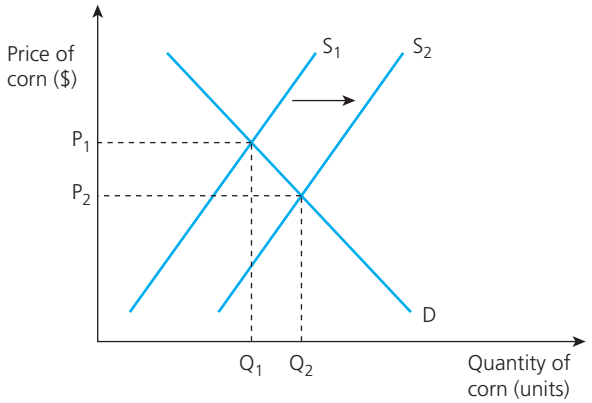
■ **Figure 5.2** A fall in demand



■ **Figure 5.3** An increase in demand



■ **Figure 5.4** A fall in supply



■ **Figure 5.5** An increase in supply

Market disequilibrium occurs when the quantity demanded for a product is either higher or lower than the quantity supplied. Disequilibrium is inefficient as it means there is either a shortage (excess demand) or surplus (excess supply) in the market.

Excess supply occurs when the price is above the market equilibrium price. A **surplus** exists because at this higher price, supply exceeds the quantity demanded – there is an incentive for firms to supply more at higher prices, but less of an incentive for customers to demand the product. For example, suppose a baker sells bread at a price of $3 per loaf. At this price, the baker is willing and able to supply 200 loaves per day, but consumers demand only

◆ **Market disequilibrium** occurs when the quantity demanded for a product is either higher or lower than the quantity supplied in the market, that is, there is either a shortage (excess demand) or surplus (excess supply).

◆ **Excess supply** refers to a disequilibrium situation where the price of a product is set above the equilibrium price, thus creating a surplus in the market.

◆ A **surplus** is created when the supply of a product exceeds its demand because the price is set higher than the market equilibrium price.

150 loaves per day at this price. Hence, the quantity supplied and quantity demanded (at a price of $3) is expressed as:

$$S(\$3) = 200 \text{ and } D(\$3) = 150$$

Excess supply is then given as:

$$ES(\$3) = S(\$3) - D(\$3) = 200 - 150 = 50$$

At this price, there is a surplus of 50 loaves of bread per day.

Excess supply is shown by the shaded area in Figure 5.6. At a price of P_1 (which is above the equilibrium price of P_e), firms supply Q_s units of output whereas customers only demand Q_d units of output. Since $Q_s > Q_d$ at a price of P_1, there is excess supply in the market.

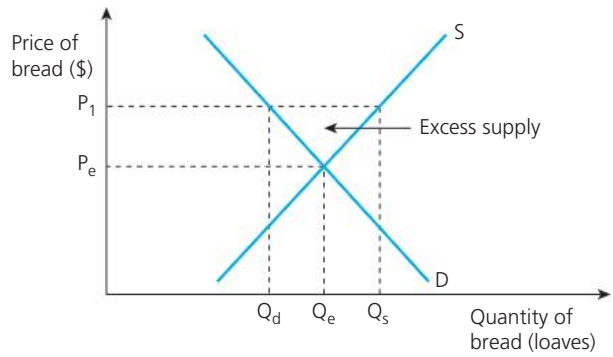


Figure 5.6 Excess supply (surplus)

If a firm has a surplus, the price can be reduced to encourage an expansion in the quantity demanded (a movement along the demand curve) and a contraction in the supply of the product. In order for firms to get rid of the excess supply (shown by the horizontal distance between Q_s and Q_d at each price level above the equilibrium price), the firm needs to reduce price (from P_1 to P_e). For example, any unsold end-of-season fashion items in a retail store are likely to be reduced in price in order to get rid of the excess supply.

Figure 5.7 Retailers use seasonal sales to get rid of excess supplies

Excess demand occurs when the price is set below the equilibrium. In other words, a shortage exists as shown in Figure 5.8. The shortage exists because at a price below the equilibrium price, demand exceeds the quantity supplied – consumers are more willing and able to buy more at lower prices, but there is less of an incentive for firms to supply the product. For example, suppose a café sells carrot cakes at a price of $4 per piece. At this price, consumers demand 50 pieces of carrot cake per day but the café is willing and able to supply only 30 pieces of carrot cake per day at this price. Hence, the quantity demanded and quantity supplied (at a price of $4) is expressed as:

◆ **Excess demand** occurs when the price is set below the equilibrium, that is, a **shortage** exists.

$$D(\$4) = 50 \text{ and } S(\$4) = 30$$

Excess demand is then given as:

$$ED(\$4) = D(\$4) - S(\$4) = 50 - 30 = 20$$

At this price, there is a shortage of 20 pieces of carrot cake per day.

Excess demand is shown by the horizontal distance between Q_s and Q_d at each price level below the equilibrium price, as shown in Figure 5.8. At a price of P_1 (which is below the equilibrium price of P_e), customers demand Q_d units of output whereas firms supply only Q_s units of output. Since $Q_d > Q_s$ at a price of P_1, there is excess demand (a shortage of the product) in the market.

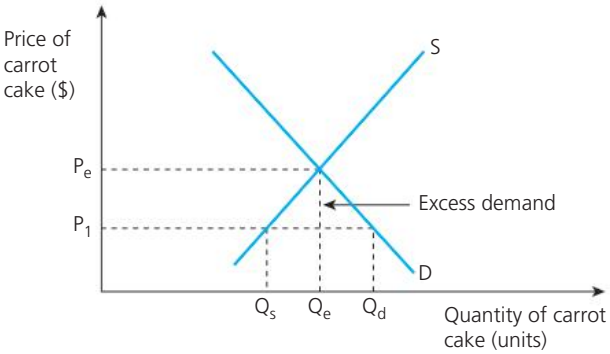


Figure 5.8 Excess demand (shortage)

If the market faces a shortage, the price will be increased to encourage an expansion in the quantity supplied (that is, a movement along the supply curve) and a contraction in the quantity demanded of the product. For example, a shortage of housing in the property market will tend to force prices up. There is also excess demand for air travel and hotel accommodation during the school holidays, so airlines and hotel operators increase their prices accordingly.

Common mistake

'Firms have a surplus when they have excess stock due to producing too much or being unable to sell their goods.'

While there is some truth in this statement, the reason for there being excess stock is that the price is too high, that is, above the market equilibrium. Hence, at this (higher) price, quantity supplied exceeds the quantity demanded.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 5.1

The diagram below represents the market for textbooks. Explain **two** possible reasons for the shift in the supply curve. *[4 marks]*

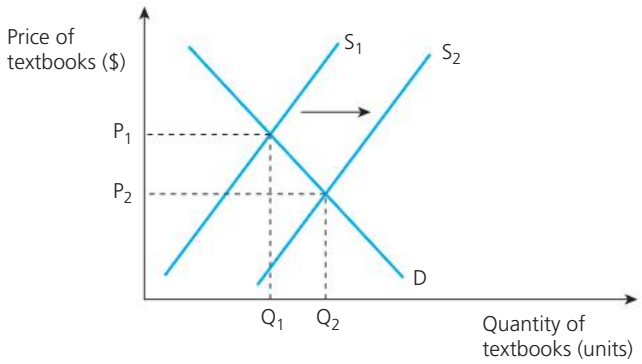


Figure 5.9 The market for textbooks

PAPER 2 AND 3 EXAM PRACTICE QUESTION 5.2

With reference to the diagram below, calculate the change in total revenue if the market demand curve for grass seeds shifts from D_1 to D_2. *[2 marks]*

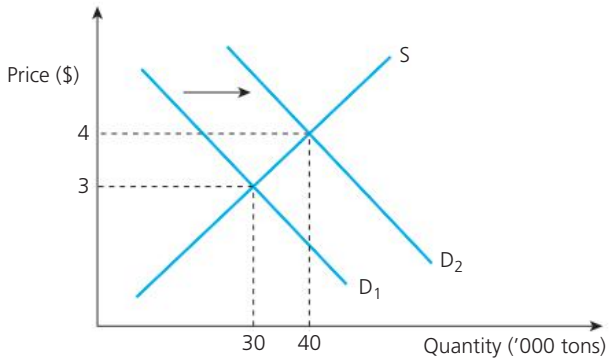


Figure 5.10 The market for grass seeds

PAPER 2 AND 3 EXAM PRACTICE QUESTION 5.3

The table below shows the demand and supply schedules for Product X.

P ($)	Q_d ('000)	Q_s ('000)
0	80	0
10	60	20
20	40	40
30	20	60
40	0	80

a Identify the equilibrium price. *[1 mark]*

b Calculate and identify the value of excess demand or supply if the price is $10. *[2 marks]*

c Calculate and identify the value of excess demand or supply if the price is $40. *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 5.4

Mexico is the world's biggest producer of avocados, supplying about 45 per cent of the planet's avocados. In 2017, avocado prices reached a record high due to reduced harvests from major producers in Mexico, Peru and California. Prices also surged due to global demand in places like the US, Europe and China. The industry has become very profitable for Mexican farmers, with avocados being nicknamed Mexico's green gold.

Using a demand and supply diagram, explain why reduced harvests caused avocado prices to reach a record high in 2017. *[4 marks]*

Functions of the price mechanism (AO2)

The **price mechanism** refers to the means by which the forces of demand and supply determine the allocation of the economy's scarce resources between competing uses. It is the buyers (demand) and sellers (suppliers) who interact to determine the price of a good or service. The price mechanism is a key model used to explain the allocation of resources, addressing the basic economic problem of scarce resources, and needs and infinite wants of the economy. Relative prices and changes in market prices reflect the forces of demand and supply in a market, thereby determining what, how and for whom production should take place (see Chapter 1).

◆ The **price mechanism** refers the interactions between buyers and sellers (or the forces of demand and supply) in the free market in order to allocate resources, thereby determining production and consumption choices.

Resources are finite in supply yet wants (desires, or goods and services that people would like to have) are infinite. This creates a situation of scarcity, which subsequently imposes production and consumption choices. Owing to limited resources, consumers must also make choices. This results in an opportunity cost. For example, choosing to buy a smartphone may come at the opportunity cost of buying a games console or going on vacation.

There are two main functions of the price mechanism – **resource allocation** (signalling and incentives) and rationing (of scarce resources).

Resource allocation (the signalling and incentives functions)

In a free market economy (see Chapter 1), price has both a **signalling function** and an **incentive function**. The price mechanism performs a signalling function because the forces of demand and supply signify where resources are required (in markets where prices increase) and where they are not (in markets where prices fall). A rise in the market price of smartphones, for example, sends a signal to potential manufacturers to enter the industry. Higher prices in the market also act as an incentive for existing suppliers of smartphones to raise output because they can earn more profit.

For example, in Figure 5.11, if the demand for onions increases, shifting the demand curve from D_1 to D_2, the equilibrium price will increase from P_1 to P_2, ceteris paribus. This *signals* to firms to raise their supply as there is an *incentive* for them to do so (greater levels of profits).

Rising prices give a signal to consumers to reduce their demand or to withdraw from the market if prices are too high. Long-term changes in price also inform producers about whether to enter or leave a market. An incentive refers to anything that motivates producers or consumers to follow a particular course of action or to change their behaviour. For instance, higher prices provide an incentive for existing producers to supply more as they can earn more sales revenue and higher profits.

Fluctuating market prices reflect the relative scarcities and surpluses in different markets. Other examples of the signalling and incentive functions of the price mechanism include:

- If prices are rising due to an increase in demand for a product, this is a signal to suppliers to expand production in order to meet the higher level of demand.
- If there is a surplus in the market, the price mechanism eliminates the excess supply by adjusting the market price downwards.
- Higher salaries (which represent the price of labour services) in a certain profession incentivises more people to enter the industry by acquiring the relevant qualifications, skills and experiences.
- During an economic recession when demand is lower, supply contracts as producers reduce their output.

◆ The **signalling function** is an aspect of the price mechanism in allocating resources by providing information to producers and consumers where resources are required (in markets where prices increase) and where they are not (in markets where prices fall).

◆ The **incentive function** is an aspect of the price mechanism in allocating resources as price changes provide an incentive (motivation) for producers and consumers to change their behaviour in order to maximize their benefits.

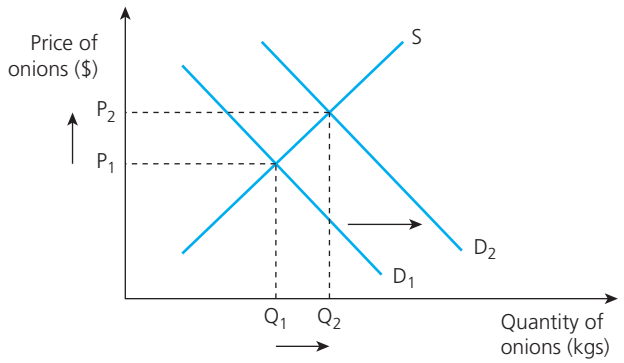


Figure 5.11 The signalling and incentive functions of the price mechanism

Figure 5.12 A financial crisis would cause demand for goods and services to decline

Top tip!

The signalling function is associated with shifts in the demand or supply curve of a product. The incentive function of a price change is associated with a movement (an extension or contraction) of the demand or supply curve of a product.

Essentially, the signalling and incentive functions of the price mechanism result in the reallocation of resources if the price of a good or service changes due to changes in any of the factors that affect the demand for or supply of a product.

PAPER 1 EXAM PRACTICE QUESTION 5.5

Using an appropriate diagram(s), explain the signalling and incentive functions of price. *[10 marks]*

Common mistake

'Government intervention is required to restore equilibrium if there is excess demand or excess supply in a market.'

The market system tends to restore equilibrium owing to the price mechanism (see tip below). The market system will tend to restore equilibrium owing to the market forces of demand and supply. In a situation of excess demand (shortage) the price will rise, thereby reducing the quantity demanded and increasing the quantity supplied. The opposite occurs if there is excess supply (surplus).

◆ The **rationing function** of the price mechanism deters some consumers from buying a product or resource owing to higher prices, thereby rationing (preserving) it. It serves to ration scarce resources when the demand for a product exceeds its supply.

The rationing function

The **rationing function** of the price mechanism serves to limit or preserve resources. Essentially, the higher the price of the product, the lower the quantity demanded will tend to be (owing to the law of demand), thereby helping to preserve (ration) the good or service. This is important and necessary when demand outstrips supply, that is, if there is excess demand or a shortage. The effect of such an increase in price is a fall in the quantity demanded, thereby helping to conserve scarce resources and to spread out their use over time. For example, as oil slowly runs out across the world, its market price will rise, so this causes demand to contract thereby conserving oil supplies. Therefore, the greater the scarcity of a resource, the higher its price so the more the resource or product is rationed.

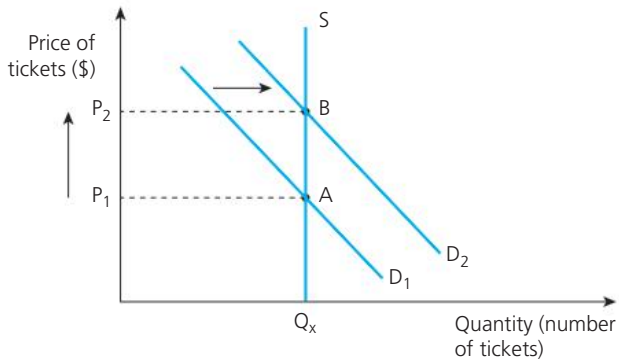


Figure 5.13 The rationing function of the price mechanism

If there is a **shortage** in the market, the price mechanism will force the equilibrium price up, thereby reducing the number of people willing and able to pay to purchase the product. For example, tickets for a major sporting event or music concert are fixed in supply so high demand for these events would cause the price mechanism to act as a rationing tool to equate demand with supply (see Figure 5.13).

In Figure 5.13, the supply of tickets for a World Cup football final is fixed at point Q_x on the *x*-axis. Higher demand for these tickets only pushes prices up from P_1 to P_2 as the supply cannot increase (owing to the fixed seating capacity in the stadium). This causes a change in the market equilibrium from point A to point B, rationing the sales of the tickets to only those who are able and willing to pay the higher price to secure tickets for the sporting event.

Another common example of the rationing function of the price mechanism is the use of auctions to allocate resources, with the highest bidder securing the purchase of the product. Examples of products sold via auctions include cars, jewellery, houses, artwork and antiques.

Figure 5.14 Auctions are a common method of rationing property sales

Top tip!

The rationing function of the price mechanism is associated with a contraction or expansion along the demand curve. An increase in price causes a contraction in the quantity demanded, whereas a reduction in price causes an expansion in quantity demanded.

WISE ChoICES

Consider real-world examples of when the price mechanism fails to ration the **scarcity** of resources. To what extent does government **intervention** alleviate this problem?

◆ **Consumer surplus** refers to the gain or benefit to buyers who can purchase a product at a price lower than that which they are willing and able to pay.

◆ **Producer surplus** refers to the gain or benefit to firms that receive a price higher than the price they are willing and able to supply at.

Consumer and producer surplus (AO2, AO4)

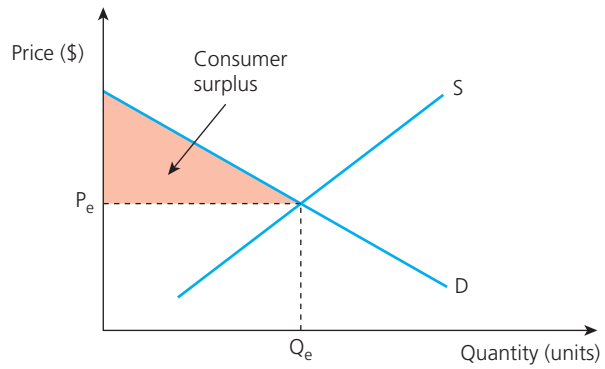
Consumer surplus refers to the gain or benefit to buyers who can purchase a product at a price lower than what they are willing and able to pay for the product. It is the difference between the price that customers actually pay and the price they were willing and able to pay. Consumer surplus (CS) is calculated using the formula:

$$CS = WTP - P$$

(where WTP is the willingness to pay and P is the market price).

For example, if customers are willing and able to pay $800 for a laptop but the price is $750, they gain $50 of consumer surplus. Hence, the consumer's marginal utility of consumption is greater than the market price.

Figure 5.15 shows that at the equilibrium price of P_e, there are customers who are willing and able to pay a higher price, as shown by all prices above the equilibrium price. Thus, the shaded area represents the total consumer surplus (the sum of the vertical distances above the equilibrium price level but below the demand curve, up to Q_e units of output).



■ **Figure 5.15** Consumer surplus

Producer surplus refers to the gain or benefit to firms who receive a price that is higher than that which they are willing and able to supply. It is the difference between the price that firms actually receive and the price they were willing and able to supply at. Producer surplus (PS) is calculated using the formula:

$$PS = P - WTS$$

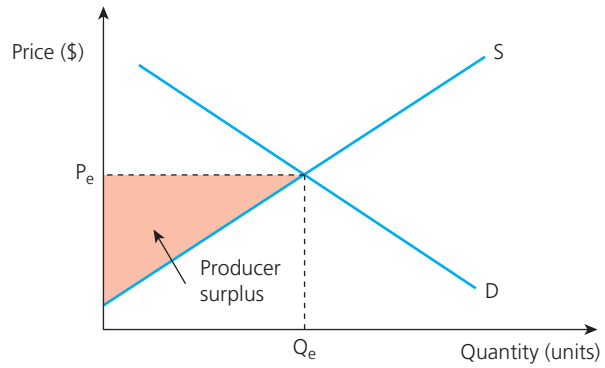
(where P is the market price and WTS is the willingness to supply).

For example, if farmers are willing and able to supply potatoes at $3 per kilogram but the market price is $3.80 per kilogram, they earn producer surplus of $0.80 per kilogram. This surplus is often referred to as *supernormal* or *abnormal profit* (see Chapter 14).

In Figure 5.16, there are some firms that are willing and able to supply at a price lower than the equilibrium price of P_e. This is shown by the shaded area, which represents the total producer surplus (the sum of the vertical distances above the supply curve but below the equilibrium price level up to Q_e units of output.

Theory of Knowledge

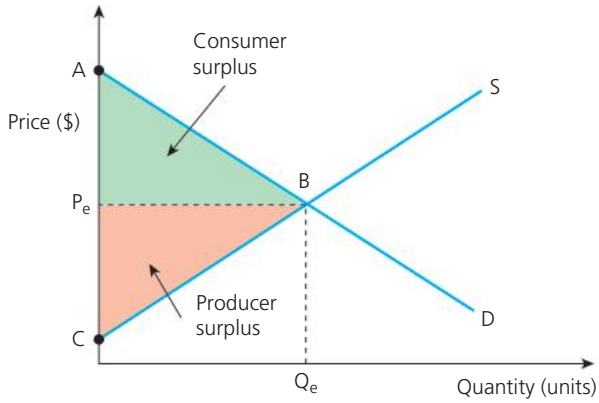
How is it possible, if at all, for economists to know how much customers are genuinely willing and able to pay in order to purchase a good or service?



■ **Figure 5.16** Producer surplus

Social/community surplus (AO2, AO4)

Community surplus (also known as **social surplus**) is the sum of consumer and producer surplus at a given market price and output. It represents the total benefit available to society from an economic transaction or activity. In Figure 5.17, community surplus is represented by the area ABC, which is made up of both consumer and producer surplus. Community surplus is maximized when the price mechanism clears the market of any excess demand or supply. As there are no shortages or surpluses at the market clearing price (P_e) and equilibrium quantity (Q_e), resources are allocated efficiently (at point B) and maximize economic welfare.



■ **Figure 5.17** Allocative efficiency and social surplus

Allocative efficiency is a situation where resources are allocated in an optimal way such that changing the price would result in consumers and/or producers being worse off. This means that consumers cannot be better off (they cannot increase their level of utility) by purchasing different quantities of goods and services, nor can firms gain more sales revenue or profits from the current combination of goods and services being supplied and sold. For example, in Figure 5.17, if the economy was operating at a quantity below Q_e, then both consumer and producer surplus could be increased by demanding and supplying a greater quantity, thereby improving economic welfare.

◆ **Community surplus** (also known as **social surplus**) is the sum of consumer and producer surplus at a given market price and output, thereby maximizing economic welfare.

Community surplus and allocative efficiency are important economic tools to consider when discussing whether economic decisions by individuals, firms and governments increase or decrease economic welfare.

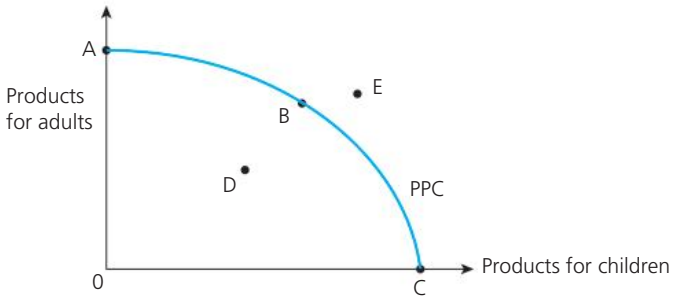
Top tip!

As a Higher Level student, you need to be able to calculate consumer surplus and producer surplus from a diagram. Have a go at some of the practice questions below to develop your understanding of consumer and producer surplus.

Allocative efficiency at the competitive market equilibrium (AO2, AO4)

Allocative efficiency is the socially optimal situation that occurs when resources are distributed in such a way that consumers and producers get the maximum possible benefit. This means that no one can be made better off without making someone else worse off. This occurs when community or social surplus is maximized (see Figure 5.17) above. It can also be shown using a PPC diagram (see Figure 5.18).

In Figure 5.18, any point on the PPC represents allocative efficiency because any reallocation of resources will make others worse off. For example, if the economy reallocates production from point A to point B, then children benefit while adults lose. Similarly, if resources are reallocated so that the economy moves from point B to point A, children will lose while adults gain. Point D represents allocative inefficiency as resources are underutilized. The economy could move from point D to point B, thereby benefiting both adults and children. Point E is not attainable given the quantity and quality of the current resources of the economy.



■ **Figure 5.18** Allocative efficiency and the production possibility curve (PPC)

◆ **Allocative efficiency** is the socially optimal situation that occurs when resources are distributed in such a way that consumers and producers get the maximum possible benefit. In other words, no one can be made better off without making someone else worse off.

Allocative efficiency can be increased if producing or buying more of something results in a greater marginal benefit to society than marginal cost. This means that allocative efficiency is achieved at the price where the marginal benefit and marginal

cost of an economic transaction are equal. In other words, community surplus is maximized in competitive markets at an equilibrium output when price is equal to marginal cost, that is, P = MC.

Note: the price mechanism is not always efficient in allocating scarce resources. **Market failure** (see Chapters 11 and 12) occurs when the signalling, incentive and rationing functions of the price mechanism fail to operate optimally, which leads to a loss in economic welfare. For example, the market may fail to take into account the external costs (such as congestion and pollution) and external benefits (such as provision of public and merit goods) arising from production and consumption.

◆ Market failure
Market failure refers to any situation when the price mechanism (the free market forces of demand and supply) allocates scarce resources in an inefficient way.

PAPER 3 EXAM PRACTICE QUESTION 5.6 (HL ONLY)

a Use the data in the demand and supply schedules below to plot a demand and supply diagram. *[4 marks]*

P ($)	Qd	Qs
0	80	0
10	60	20
20	40	40
30	20	60
40	0	80

b Use your graph to calculate the total value of the consumer surplus at the equilibrium price. *[2 marks]*

c Calculate the total value of the producer surplus if the price falls to $10. *[2 marks]*

d Calculate the quantity of excess demand at a price of $10 per unit. *[2 marks]*

e Calculate the quantity of excess supply at a price of $30 per unit. *[2 marks]*

Theory of Knowledge

How can we know when an economic problem truly justifies the need for government intervention? Refer to one of the six real-world issues (page v) in your response.

PAPER 3 EXAM PRACTICE QUESTION 5.7 (HL ONLY)

Explain the view that the best allocation of resources occurs when consumer surplus and producer surplus are maximized. *[10 marks]*

Chapter summary

- The interaction of the demand and supply curves for a product forms a market equilibrium, which clears the market of any excess demand or supply.
- Market disequilibrium occurs when the quantity demanded for a product is unequal to the quantity supplied of the product, so there are shortages or surpluses.
- A change in any non-price factor that affects the demand or supply curves produces a new market equilibrium, due to the shift in one of the two market forces.
- If *excess supply* (surplus) exists in the market, there is a tendency for the price to fall in order to remove the surplus.
- If *excess demand* (shortage) exists in the market, the price will tend to rise in order to restore market equilibrium.
- The price mechanism is a key model used to explain the allocation of resources, addressing the basic economic problem of what, how and for whom to produce. Owing to limited resources, economic agents must make choices.
- The functions of the price mechanism are resource allocation (the signalling and incentive functions) and rationing of scarce resources.
- *Consumer surplus* is the positive difference between what consumers are willing and able to pay for a product and the amount they actually pay.
- *Producer surplus* occurs when firms are able to charge a higher price for a product than what they are willing and able to supply it for.
- *Social surplus* (or *community surplus*) is the sum of producer surplus and consumer surplus.
- In a competitive market economy, allocative efficiency occurs at the market equilibrium because community surplus is maximized at this point.
- Allocative efficiency at the competitive market equilibrium generates the maximum social surplus as the marginal benefit (MB) of an economic transaction equals the marginal cost (MC).

6 Critique of the maximizing behaviour of consumers and producers (HL only)

Real-world issue

How do consumers and producers make choices in trying to meet their economic objectives?

Conceptual understandings (WISE ChoICES)

- Consumer and producer **choices** are the outcome of complex decision-making.
- Welfare is maximized if allocative **efficiency** is achieved.
- Constant **change** produces dynamic markets.

SYLLABUS CONTENT

By the end of the chapter, as a Higher Level student, you should be able to understand:

- the assumptions behind rational consumer choice: consumer rationality, utility maximization and perfect information (AO3)
- the limitations of the assumptions of rational consumer choice in terms of behavioural economics: biases, bounded rationality, bounded self-control, bounded selfishness and imperfect information (AO3)
- behavioural economics in action: choice architecture (default, restricted and mandated choices) and nudge theory (AO3)
- different business objectives: profit maximization and alternative objectives (corporate social responsibility, market share, satisficing and growth) (AO3).

Rational consumer choice (AO3)

Rational choice theory states that individuals use logical and sensible reasons to determine the right choice associated with an individual's best self-interest. **Rational consumer choice** refers to the decision-making process based on people making choices that result in the optimal or maximum level of benefits or utility for an individual. This means that rational consumer choice is expected to result in outcomes that provide individuals with the greatest benefit (or satisfaction), given the choices available to them.

◆ **Rational consumer choice** or **consumer rationality** refers to the decision-making process based on the assumption that people make choices that result in the optimal level of benefits (utility). It is the basis of most introductory and mainstream economic theories.

Many of the tools and theories in introductory and mainstream economics assume that economic agents behave rationally, in order to maximize their own benefit. For example, firms are assumed to be profit maximizers whereas consumers aim to maximize their utility from the consumption of goods and services. It may also be rational to assume that:

- consumers buy more of a product if the price falls (the law of demand)
- firms buy products from the supplier that offers the best (lowest) price
- shareholders buy stocks with the greatest expected rates of return
- investors aim to minimize their level of risk
- people are happier if they have a higher standard of living
- the government wants to improve the economic and social well-being of citizens
- IB students who work hard are more likely to pass their exams and secure a place at university.

Figure 6.1 Rational students work hard to pass their IB exams

However, behavioural economics challenges the assumption of pure and consistent rationality in people's decision-making. Behavioural economists question whether we always act rationally and whether all firms aim to maximize their profits. So, while IB students may 'know' that they need to work hard in order to pass the Diploma, they may lack the self-discipline to do so at times. There are also many distractions in life that may take priority for these students.

Assumptions

Traditional economic theories suggest that choice is a good thing, and that people should relish having many choices. These theories make three key assumptions: consumer rationality, utility maximization and perfect information.

Consumer rationality

It is assumed that individuals use rational calculations or considerations to make sensible choices that align with their own best interests. Rational choice theory assumes that people are driven by self-interest and incentives (the invisible hand theory).

Utility maximization

The preferences of individuals are often described by economists by the level of utility (satisfaction) gained from the decision or choice. Economic theory assumes that people decide on the option that gives them the maximum level of utility. So, if individuals would prefer to watch a film at the cinema over going ice skating, it is assumed that they choose the movies as this yields the greatest level of satisfaction.

Perfect information

Perfect information refers to equal and easy access of information about alternative products available on the market. In order to act rationally, economic agents must have easy access to full (or perfect) information, such as prices, product specifications and information about alternative products available on the market. Rationality assumes that people use all the available information to make their decision.

Ironically, however, having too many choices can make people less likely to make a rational decision and can be both physically and mentally exhausting.

ATL ACTIVITY 6.1

With a partner, discuss the assumption of rationality and come up with **five** other real-world examples of rational behaviour. Be prepared to share your responses with the rest of the class.

ATL ACTIVITY 6.2

Most standard economic theories assume that every economic agent takes part in an action or activity in a rational way. Discuss how rational consumer choice can also refer to individuals aiming to minimize their losses with certain decisions.

Behavioural economics – limitations of the assumptions of rational consumer choice (AO3)

Typically in economics, it is assumed that people aim to maximize their own welfare (economic well-being) when making a decision. However, in reality, when making day-to-day decisions, people rarely behave in a well-informed and fully rational way. For example, some people choose to drink too many sugary drinks, eat too much junk food, take too little exercise, smoke too much, or do not save enough money for their retirement.

Behavioural economics is the study of actual economic decision-making, with emphasis on human behaviour being less rational than traditional economic theory suggests. Behavioural economics acknowledges the power of social conformity on decision-making, because humans are social beings whose behaviour is clearly influenced by the behaviour of others in society, including peer pressure and the desire to conform to social norms. Contrary to traditional economic theory, behavioural economics does not assume that individuals make choices in isolation or simply to serve their own interest.

Limitations of the assumptions of rational consumer choice include the following five interrelated factors: biases, bounded rationality, bounded self-control, bounded selfishness and imperfect information.

Biases

When making decisions, people's cognitive responses are influenced by biases including common sense, intuition, emotions, and both personal and social norms. Biases influence how people process complex information as part of the decision-making process. These biases make it extremely difficult, if not impossible, for social beings to make a pure and instantaneous rational decision. Biases can be categorized as rule of thumb, anchoring and framing, and availability.

Rule of thumb – This phrase refers a practical and approximate way of doing something, without having to be exact. It is about judging a situation based on experience which affects choices in trying to meet economic objectives, such as consumers aiming to maximize their utility. As a rule of thumb, people tend to stick to their default choice. For example, children tend to stick to the same choice of food when visiting fast food restaurants such as McDonald's. Drivers often go into 'auto pilot' mode and drive a particular route to and from their workplace. Some savvy shoppers might have a rule of thumb of only buying supermarket goods that have a special offer. However, using rules of thumb can lead to irrational decision-making. For example, it is normal to assume that discounted prices are cheaper than the average price of the same product over a period of time, when in fact this is often not the case.

Anchoring and framing – Anchors and framing create a bias in favour of a particular choice. For example, what people see when they pass by or first enter a retail store becomes the anchor from which future choices are assessed. So, if a clothes retailer uses a window display showing that clothing items are on sale with a discount of up to 70 per cent, that becomes the first anchor. Anything less than 70 per cent off seems less attractive but helps to convince individuals that they are still getting a

Figure 6.2 Anchoring and framing – is this really a closing down sale?

Theory of Knowledge

Can the use of empirical evidence, even without any bias, genuinely allow us to arrive at the truth about real-world issues?

◆ **Rule of thumb** is a type of bias referring to making decisions in a practical and approximate way without having to be exact. It is about judging a situation based on experience and sticking to a default choice.

◆ **Anchoring** is a cognitive bias that influences how people view a product by comparing it to something else.

bargain. Without the first anchor, some customers may not have even walked into the store. Real estate agents also use anchors for property prices – showing the offer prices for houses on sale acts as an anchor for potential home buyers. Essentially, anchoring is a cognitive bias that influences how people view a product by comparing it to something else.

Similarly, framing is about presenting information in such a way that it creates a bias in favour of a particular decision. Choices can be presented to highlight the positive or negative aspects of the same decision, which influences their relative attractiveness to decision makers. Hence, anchoring and framing also provide a constraint on people's tendency to exercise rational decision-making.

CASE STUDY

Tipping NYC taxi drivers

A famous example of anchoring is New York City's introduction of credit card payment systems back in 2009, which automatically suggested tips of either 30 per cent, 25 per cent or 20 per cent. The lowest suggested amount becomes the anchor, causing passengers (especially tourists) to believe that 20 per cent is the lowest tip that is acceptable, whereas the previous average was only around 8–10 per cent! The Taxi and Limousine Commission reported that since the installation of the credit card systems in the city's taxis, average tips increased by 22 per cent.

◆ **Availability bias** refers to the ease with which an idea or event can be recalled from memory. This causes people to overestimate the likelihood or probability of the event or idea ever happening, thereby distorting rational decision-making.

Availability bias – The availability bias occurs when people overestimate the likelihood of an event, or the frequency of its occurrence, either because something similar has happened recently or because they feel very emotional about a similar event. The result is that people tend to misjudge things simply because certain examples come easily to their minds. One of the consequences of this is that people become poor at estimating probability and risks, such as overestimating the likelihood of shark attacks or aircraft accidents. People often blame climate change and global warming for seasonal variances in the temperature. More people catch the bus after reports of a train accident.

◆ A **choice architect** is an individual or organization who organizes the context in which people make decisions, based on the simplified design of different choices being presented to them.

Behavioural economists study how choices can be influenced or manipulated by policymakers in order to meet economic objectives and improve economic well-being (see section below on **choice architecture** and nudge theory).

Bounded rationality

◆ **Bounded rationality theory** states that rationality is limited. This means that people's cognitive decision-making capacity is limited owing to barriers such as information failure or the time available to make the decision.

In behavioural economics, **bounded rationality theory** argues that people frequently choose options without the ability to gather all relevant information they need to make the most rational (or best) possible decision in a timely manner. Coined by Nobel laureate Herbert Simon, the theory of bounded rationality states that decisions are often based on incomplete information. This is because rationality is bounded (restricted) by people's thinking capacity, the availability of information and time. Hence, irrationality causes a loss of social welfare.

People do not always like choice, especially when there are too many. As social beings, they have habits and natural tendencies. Thinking about choices in a rational way and making sensible and considered decisions can stress people out. So, people often think automatically rather than rationally.

Bounded self-control

◆ **Bounded self-control** is a limitation of the assumptions of rational consumer, as people may lack the self-control to think for themselves and to make a rational choice. Instead, they conform to social norms and group preferences and pressures, which may not match their own preferences.

Humans are social beings. As such, they often let their family, friends and acquaintances influence their thinking. They lack the self-control to think for themselves and to make a rational choice. Instead, people conform to social norms and group preferences, which may not match their own preferences.

Post-purchase, consumers may reflect and ask themselves whether they really needed to buy that product. For example:

- Did I really need to buy those expensive shoes or clothes?
- Did I really need to order so much wine?
- Did I really need to have dessert knowing how much time I would need to exercise just to burn off all those calories?
- Did I really need to buy that just because I felt pressurised to do so by my friends (or the salesperson)?

The rational consumer might decide otherwise, but bounded self-control means emotions take over and we might do something irrational by justifying that we are worth it right now. Indeed, marketers thrive on boundless self-control by creating strong marketing messages to appeal to their target audience and finding the right emotional appeal to attract customers. For example, marketing campaigns such as special offers or discounted prices 'for a limited time period only' or 'while stocks last' can make consumers react irrationally due to the fear of missing out on a bargain.

ATL ACTIVITY 6.3

Watch this version of The Marshmallow Test, which is an experiment that involves giving children the choice of either having one marshmallow to eat immediately or waiting a little bit longer to receive two marshmallows: **https://youtu.be/QX_oy9614HQ**

How do the ideas of bounded self-control and bounded-selfishness apply to these children as decision makers?

Bounded selfishness

Selfishness means economic agents care only about their own well-being. Traditional economic theories and economic models assume economic agents make the best decision, based on the available information, which will maximize their personal benefit. However, behavioural economics suggests that humans do not act only in self-interest, but often behave reciprocally and altruistically rather than pursue their own pure self-interest. This is known as **bounded selfishness**. As Martin Luther King Jr said, 'Every man must decide whether he will walk in the light of creative altruism or in the darkness of destructive selfishness.'

◆ **Bounded selfishness** means that people are not necessarily selfish and only act in self-interest, but often behave reciprocally and altruistically. People do not always pursue their own pure self-interest.

Examples of people's bounded selfishness (doing things for other people without getting any recognition or financial benefit from doing so) include:

- helping strangers without expecting anything in return
- donating money to charity
- giving food and water to homeless people
- random acts of kindness towards others
- organ donations, such as kidney transplant, to another person
- doing voluntary work, such as beach clean ups.

Imperfect information

In reality, information is rarely perfect. In some industries, trade secrets exist, and intellectual property rights (patents, copyrights and trademarks) act as barriers for firms to enter some industries. The lack of equal and easy access to information makes rational decision-making more challenging. For example, people do not necessarily know the long-term costs and benefits of their decisions, such as:

- the potential long-term risks of using tanning salons or disposal of plastic waste
- the true odds against gamblers winning any money in casinos or lotteries

Figure 6.3 Beach clean ups are example of bounded selfishness

- subscription packages to mobile phone networks or TV streaming services such as Netflix
- complexities and the true level of risks associated with financial services such as motor and health insurance, pensions schemes and financial investments.

◆ **Imperfect information** occurs when people have inaccurate, incomplete or unreliable information, so decision-making is not optimal.

Imperfect information (or **information failure**) occurs when people have inaccurate, incomplete or unreliable information, so decision-making is not optimal. Sellers often have more detailed information than buyers about a good or service because they are more familiar with it, such as the product features, quality and profit margin. This means buyers, with their limited or asymmetric information (see Chapter 13) can make potentially 'wrong' choices. For example, buyers of second-hand cars have far less accurate information about the products than the sellers of these cars.

Imperfect information creates bounded rationality, thus distorts rational decision-making. For example, when Apple launched its iPhone at a price of $600, but then quickly reduced it to $400, the company was able to convince people to think that $400 was a real bargain for a smartphone. If Apple had set the initial price at $400, customers would not have had a basis for comparison, so the price tag may have seemed very expensive.

Figure 6.4 Imperfect information exists in the market for second-hand cars

Behavioural economics in action (AO3)

Economic models tend to assume that economic agents are rational decision makers. This means that consumers will buy at a price that maximizes their level of utility, while producers will sell at a price that enables them to maximize profits. However, market failure and behavioural economics have seriously challenged this core assumption. For example:

- People may know that it is important to save for their retirement but may lack the willpower to do so.
- Experience tells people that most of the time they spend at an amusement (theme) park is wasted by waiting around doing nothing while queuing for the rides or shows; yet they still choose to go to these amusement parks.
- If people are informed and educated about the harmful effects of smoking and drinking alcohol, then why do they continue to consume these demerit goods (products that are considered to be damaging to society as a whole)?
- Similarly, most gamblers know the odds are highly against them when gambling in casinos and better in other places, yet they continue betting.

Figure 6.5 Given the odds against winning anything, is it irrational to gamble?

In reality, decisions are more often than not made instinctively and unconsciously rather than logically and consciously (by weighing up all the costs and benefits of a particular decision). Therefore, given that people are not necessarily rational in their decision-making, choice architects and nudges are used to overcome the pitfalls to reasonable and rational decision-making. This applies whether referring to consumers making purchasing decisions or employees trying to meet performance targets.

Choice architecture

Choice architecture refers to the deliberate design of different ways of presenting choices to members of society, and the impact of these methods on decision-making. Examples of the way choices are presented include the layout of supermarkets and calorie counts on restaurant menus. In modern societies, people have a limited attention span due to the ever-growing complexities in life, so prefer simplicity (just consider how often people go beyond page 1 of a search on Google). Changing the context in which people make choices can help create new and desired social and economic behaviour. For example, marketers often use phrases such as 'most popular', 'our top pick', or 'best seller' to reduce choice overload. By simplifying the choices, it makes the process of selecting a product easier. The term choice architecture was coined by Richard Thaler and Cass Sunstein in their book *Nudge: Improving Decisions about Health, Wealth, and Happiness* (2008).

A **choice architect** is an individual or organization responsible for shaping the context in which people make decisions. The choice architect could be a parent, teacher, colleague, friend, employer, the media, or a government official who organizes the context in which people make choices based on the simplified design of different options being presented to them.

Choice architects help to deal with the problem of choice overload. For example, restaurant owners design menus to create nudges for customers to choose certain items – noticeably, there may be one item (which the restaurateur doesn't expect many customers to buy) that is significantly more expensive than the others, but the second or third most expensive items may seem like a bargain when customers compare relative prices on the menu. In the context of supermarkets, choice architects position fresh fruits and vegetables at the entrance of the retail store. The products are placed at eye level as a nudge to encourage and facilitate consumer choice.

Choice architects and nudges are used owing to the contradiction between rational human behaviour in theory and irrational decision-making in reality, such as the over-consumption of junk food and sugary products, which has led to problems of childhood obesity in many parts of the world. Perception nudges are used by choice architects in response to people reacting differently to the same information depending on how it is presented, for example, food products described as 99 per cent fat free are perceived more favourably than food items described as 1 per cent fat.

Choice architecture falls into three categories: default, restricted and mandated choices.

Default choices – A default choice occurs when a person is automatically signed up into a system, such as employees being enrolled on a pension scheme to make financial contributions to their retirement savings accounts, or parents being automatically registered to contribute to the parent–teacher association (PTA) at their children's school. Similarly, some restaurants have made salad the default side option instead of chips (fries). E-commerce businesses usually have a default shipping option, although other alternatives are available. Default choices minimize the costs of choosing and deciding what to do. Hence, a default choice is the given decision of the economic agent if no action is taken. The default means that people are more likely to participate when they are automatically enrolled into the scheme, especially as this is the simplest or easiest option. Research shows that people rarely change their default course of action. However, if the default option or setting is nudged, social behaviour will also change.

Top tip!

Behavioural economics aligns well with the (six) real-world issues in the IB Economics syllabus, namely *'How do consumers and producers make choices in trying to meet their economic objectives?'* (page 5). Relevant key concepts in this topic include economic well-being and choice. Ensure that you can provide real-world examples of behavioural economics, including choice architecture and nudge theory.

◆ **Choice architecture** refers to the way choices are presented to members of society and how different designs affect the choices made. Effective choice architecture helps people avoid making poor choices and irrational decisions.

Restricted choices – As the name suggests, this type of choice architecture limits the choices available to people, such as placing healthier food items on a menu and restricting the choice of unhealthier foods. For example, a growing number of restaurants serve only diet or low sugar versions of soft drinks, such as Diet Coke or Coke Zero. Other eateries charge a premium price for drinks with high sugar content, owing to a sugar tax. Either way, the choice for consumers is restricted, but it is still there. The nudge is created to encourage consumers to opt for the less unhealthy drink. Hence, restricting the number of available choices can help consumers to make a more rational or sensible decision.

Mandated choices – A mandated choice is when people are required to make an advanced decision and declare whether they wish to participate in a particular activity. Policymakers use mandated choices to ensure people make a choice on matters that are deemed to be important to society and the economy. For example, when people apply for a new driving licence or renew their driving licence, they may be asked to declare whether they wish to donate their organs in the event of death. Some governments give a mandated choice to individuals who must register to vote in person (at a polling station or venue) or by post; either way, they must declare how they wish to vote. Essentially, they are required by law to make that choice.

Top tip!

In reality, it can be tricky to categorize choices and nudges as mandated, default or restricted. For example, cafés that default to using 'skinny' (low-fat or skimmed) milk also restrict consumer choice. The nudge could also have been mandated by the owners or government guidelines and regulations.

Nudge theory

Nudge theory is the practice of influencing the choices that people make. Nudges are created by choice architects using small prompts or tweaks to alter social and economic behaviour, but without taking away the power for people to choose. Nudge theory was coined by Richard Thaler, who went on to win the Nobel Prize in Economics in 2017 for his influential research in behavioural economics.

◆ **Nudge theory** is the practice of influencing the choices that people make. Nudges are created by choice architects using small prompts or tweaks to alter social and economic behaviour, but without taking away the power for people to choose.

Thaler and other behavioural economists have shown time and time again that in reality people make decisions that are far from rational. Instead, they make decisions quickly under pressure, often based on intuition and sometimes unconsciously but guided by personal, psychological and social biases. Thaler argues that firms should use nudges in a responsible way, that is, subtle interventions that guide choices without restricting them. He writes (2008) '*To count as a mere nudge, the intervention must be easy and cheap to avoid. Nudges are not mandates. Putting the fruit at eye level counts as a nudge. Banning junk food does not.*'

Theory of Knowledge

To what extent do others know what we want, when we don't even know this ourselves?

Some examples of nudges include:

- fly (insect) stickers placed on urinals in public toilets, to reduce spillages
- warning systems used in motor vehicles for safety reasons, such as automatic LED lighting, low fuel notification and audio alerts when seatbelts are not used
- highly visible speed cameras to reduce speed in strategic places, such as busy road junctions and roads near schools
- membership schemes being opt out rather than opt in by default, such as parent–teacher associations (PTAs) in schools, workplace pension schemes, or trade union affiliation – the perceived effort needed to opt out of the system means membership rates are much higher
- some countries offering free vaccinations to school children, funded by the government, and only if parents opt out can this be avoided
- food items displayed in cafeterias, where healthy food products (such as fresh fruits) are placed at the beginning of the line and/or at eye level within easy reach to encourage healthier choices
- different coloured bins to encourage recycling of waste by category, for example, paper, plastics, food, or general waste

- businesses informing customers of the monthly payment for a good or service, rather than the aggregate yearly amount, to encourage consumption as it seems more affordable
- in countries such as Australia and the UK, cigarettes and other tobacco products are hidden from view, meaning smokers must put in extra effort to buy the (demerit) goods
- labels on food products that show how much time people will need to exercise in order to burn off the calories can help to promote healthier eating choices
- governments educating the general public about the dangers of excess consumption of products with high sugar content, to discourage the consumption of goods such as soft fizzy drinks
- encouraging workers to save into a pension scheme (for their retirement) rather than to spend all their money now. Note that with nudges there are no specific penalties for workers who choose not to do this.

Figure 6.6 Placing books at eye-level in public libraries or in bookstores encourages certain choices

Figure 6.7 Governments use nudge theory to encourage people to pay their taxes on time

The UK government has used nudge theory to inform and influence wealthy households failing to pay their taxes (the correct amount and/or on time). The government sent letters to wealthy people explaining how their taxes would help to improve local services, highlighting what would disappear without funding, that is, the opportunity cost of not paying their taxes. The BBC reported that these tweaks led to £210 million ($276m) in overdue taxes being paid.

Nudge theory describes situations where choice architects create nudges to improve the social and economic well-being of individuals and societies. For example, Dr David Halpern from the UK's Behavioural Insights Team argues that financial incentives do not necessarily increase labour productivity in the workplace. Instead, he argues that giving a financial incentive with a twist works better, such as spending this reward on someone else in the firm. This has several benefits:

- It improves communication in the workplace as employees need to find out more about the person they are spending the financial gift on.
- It feels good to be able to give to someone else.
- It improves well-being in the workplace, which has a positive and significant impact on productivity.

Halpern emphasizes the EAST framework to positively affect human behaviour to nudge decision makers such as consumers and employees:

- **Easy** – reducing or removing friction and simplifying things makes things more straightforward (as easy as possible) in order to influence what people do.
- **Attractive** – breaking through people's attention, such as personalizing a message or product. Choice architects often base nudges on loss aversion; the fact that people do not like to miss out on opportunities and bargains.

- **Social** – humans are influenced by what other people do, more so than what the rules and regulations might be; for example, if most people pay their utility bills on time, then we are also more likely to do so.
- **Timely** – determining when people are most responsive; behavioural economists consider when behaviour is most likely to be disruptive in order to encourage people to make decisions.

An example of this in practice is seen in supermarkets, when people line up at the checkouts. Supermarket operators use attractive point of sale displays to make it easy for people to pick up small items, such as batteries and chewing gum, just in time before they need to pay for their shopping.

Figure 6.8 Supermarket operators are experts at using nudges

Coffee retailers such as Starbucks also use nudge theory to encourage people to buy larger sizes of their drinks. Research has shown that when people are presented with three choices of different sized cups of coffee, they tend to go for the middle one (what is known as the Goldilocks effect). This is because, given the three choices, people tend to think the largest size of coffee is just too large for them. However, by removing the smallest size of drink and offering an even larger one at the other end, people are then nudged to buy what used to be the largest drink.

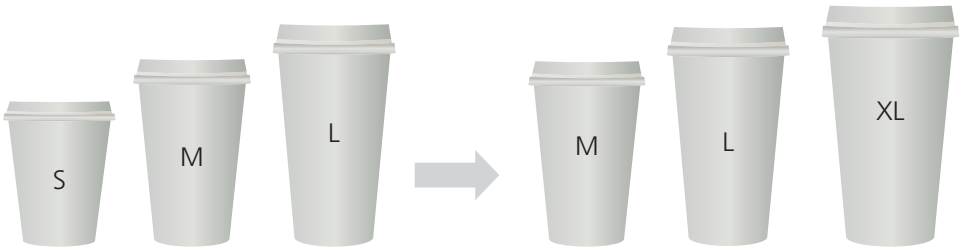


Figure 6.9 Nudge theory is used in upselling of coffee drinks

Critics of nudge theory argue that policymakers and marketers use choice architecture to simply manipulate decision-making. Nevertheless, nudge theory at least enables us to better understand what and why policymakers and firms do this. For businesses, effective nudging seems to be increasingly crucial for maintaining competitiveness.

CASE STUDY

Piala Inc.

In September 2017, Japanese marketing company Piala Inc (established in 2004) introduced a fringe benefit to all its non-smoking employees by giving them an extra six days of paid holiday per year. The idea came from an employee who complained that colleagues who smoke take more regular breaks throughout the day. The CEO agreed, but the nudge encouraged many workers to quit smoking in order to take advantage of the perk. The nudge was used as an incentive, rather than a penalty or coercion. According to the WHO, about 130,000 people die every year in Japan of smoking-related diseases, and an extra 15,000 people die of second-hand smoke-related conditions.

ATL ACTIVITY 6.4

Investigate how nudges are used in the real world by reflecting on how **you** have personally been influenced by nudges (including those that you may not even have been aware of at the time of purchase).

WISE ChoICES

Discuss how nudges affect the concepts of **economic well-being** and **choice**.

ATL ACTIVITY 6.5

With a partner, discuss the extent to which nudges are manipulative and unethical in influencing economic behaviour.

Business objectives (AO3)

Profit is an important objective for most private sector firms. It is the reward for entrepreneurs for taking risks in economic activity. It is generally assumed that firms are profit maximizers, seeking to produce at the level of output that generates the highest level of profits. **Profit maximization** occurs when there is the greatest positive difference between total revenue and total costs. This is achieved by strategies to increase total revenue and/or strategies to minimize costs of production. Total revenue is the total value of a firm's sales revenue, while total costs are the sum of all expenses incurred by the firm in the production process.

Profit maximization occurs at the output level where the **marginal cost** (MC) equals the **marginal revenue** (MR), that is, MC = MR. This is because:

- If MR > MC, profits will rise if firms increase their output because the marginal (additional) revenue outweighs the marginal cost, that is, profits increase when MR > MC.
- If MC > MR, firms will reduce their output because the marginal cost outweighs the marginal revenue, that is, profits decrease when MC > MR.
- Hence, profits are maximized when MC = MR. This is because each extra unit of output creates no extra loss or extra revenue.

Profit maximization is important to many firms because:

- It provides the financial reward for entrepreneurs who take risks in the production process.
- Profits are an important source of finance, especially as it saves firms having to incur interest payments on bank loans to fund their operations.
- Higher profits enable firms to pay higher wages to their employees, which is important for staff motivation, loyalty and productivity.
- The shareholders or owners of most private sector firms aim to maximize the return from their investments. Shareholders demand an attractive dividend payment (a share of the company's annual profits distributed to shareholders), so it is important to be profitable. The interests of the shareholders or owners must be met for firms to operate sustainably.
- Profits also help firms to be in a better financial position to survive, especially during a recession (economic downturn).

While profit maximization has dominated traditional theory of the firm, behavioural economics suggests profit is only one motivating factor for entrepreneurs and other key stakeholders. Alternative business objectives include corporate social responsibility (CSR), market share, satisficing and growth.

Corporate social responsibility (CSR)

Firms are increasingly setting ethical objectives, rather than aiming solely or primarily for profit maximization. **Corporate social responsibility** (CSR) refers to committing to ethical objectives to benefit both internal and external stakeholders, not just to the owners or shareholders of the business. CSR is about a firm committing to the values, decisions and actions that impact society in a positive way.

Theory of Knowledge

To what extent is *emotion* more important than *reason* when studying behavioural economics?

◆ **Profit maximization** is the assumed fundamental goal of private sector firms. It occurs when there is the greatest positive difference between total revenue and total costs, or at the level of output where MC = MR.

◆ **Marginal cost** (MC) is the cost of producing an extra unit of output. It is calculated by dividing the change in total costs (ΔTC) by the change in the level of output (ΔQ), that is, MC = ΔTC / ΔQ.

◆ **Marginal revenue** (MR) is the extra revenue received from the sale of an extra unit of output, that is, MR = ΔTR / ΔQ.

◆ **Corporate social responsibility** (CSR) is about businesses considering the impact of their operations on the society as a whole welfare in a positive and ethical way.

Essentially, CSR practices are about what is deemed by society to be right, such as:

- reducing carbon footprints to combat the problems related to climate change
- engaging in fair trade practices
- participating in charitable and volunteer work in local communities
- improving the well-being of workers
- making use of socially responsible marketing and advertising campaigns
- minimizing waste
- implementing green technologies in the workplace.

Market share

Market share refers to a firm's portion of the total value of sales revenue in a particular industry. It is a measure of a firm's relative size by expressing its sales revenue as a percentage of the total market sales revenue. Market share is calculated by using the formula:

$$\frac{\text{Firm's total sales revenue}}{\text{Industry's total sales revenue}} \times 100$$

For example, firms with the largest market share in the sports shoes industry include Nike, Adidas, New Balance and Puma. These firms are **market leaders**. Market leaders are firms with the largest market share in a given industry. Market share data help to reveal the extent to which an industry is competitive by aggregating the market share of the largest few firms (the market leaders) in the industry. Being able to compare the relative size of businesses enables entrepreneurs to take appropriate actions to remain competitive.

Firms may aim to increase and protect their market share. This is because having a large market share can give the firm competitive advantages, such as:

- brand loyalty, which leads to ongoing or repeat sales and customers being less sensitive to higher prices being charged
- Economies of scale – these are cost-saving benefits due to the firm's larger scale operations (see page 209). and
- higher profits as the firm accounts for more of the market's overall sales.

Market leaders may also enjoy a good corporate reputation, which can help to attract more investors and better-quality employees. In addition, retailers and distributors are more likely to stock the products of market leaders, further helping to improve their sales and profits. These benefits (economies of scale, customer loyalty and higher profits) are all forms of competitive advantages for market leaders so act as barriers to entry for other firms wanting to compete in or enter the industry.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 6.1

Heritage Pots has a 16.4 per cent market share of the garden pots market in Wallington. The market is worth $14,000,000 per year. Calculate the annual sales of Heritage Pots. *[2 marks]*

Satisficing

Neo-classical economic theory assumes that firms aim to maximize profits. However, Herbert Simon introduced the idea of satisficing as an alternative objective in 1956. **Satisficing** is about aiming for a satisfactory or adequate level of profit, rather than the maximum profit. While the

ATL ACTIVITY 6.6

Conduct a critical investigation into the CSR practices of **two** different companies of your choice.

ATL ACTIVITY 6.7

Watch this documentary about the Deep Water Horizon disaster:

https://www.youtube.com/watch?v=szwdgughGLE

Discuss the consequences of firms such as BP neglecting their corporate social responsibilities.

◆ **Market share** refers to a firm's portion of the total value of sales revenue in a particular industry.

WISE ChoICES

Discuss the importance of corporate social responsibility for **economic well-being** and **sustainability**.

◆ **Satisficing** is a business objective that aims for a satisfactory or adequate level or profit, owing to the sacrifices needed if profit maximization were to be pursued.

shareholders or owners of a company might want to maximize profits, the workers and managers might lack the incentive to do so; they want sufficient profit to ensure there is job security but do not want to overwork.

In reality, it is difficult for a firm to be exact about its profit maximizing level of output, as it cannot accurately determine marginal revenue and marginal cost. In addition, profit maximization might require significant expenditure of time, effort and financial resources. The conventional theory of the firm assumes that businesses have enough information, market power and willingness to set prices in order to maximize profits. This assumption has been criticized for ignoring that modern-day businesses have a range of organizational objectives such as corporate social responsibility.

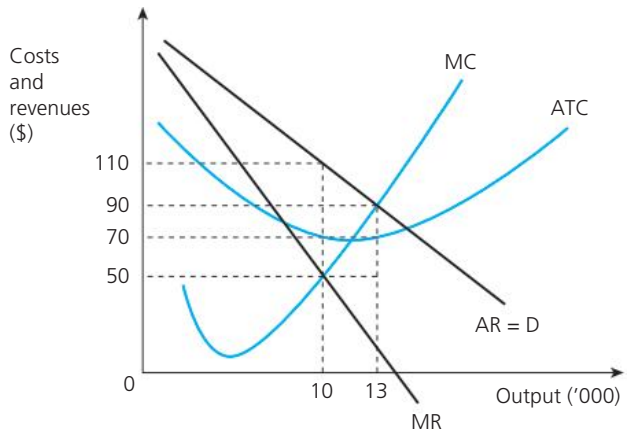


Figure 6.10 Profit satisficing

With reference to Figure 6.10, the following applies to the profit-maximizing firm:

- Output = 10,000 units as this is the output level where MC = MR
- Price (or average revenue) = $110
- Average cost = $70
- Hence, the profit for the firm = TR – TC = ($110 × 10,000) – ($70 × 10,000) = $400,000.

By contrast, the profit satisficing firm can charge a price between $90 (the free market equilibrium price) and $110 (the profit-maximizing price). Suppose the profit-satisficing firm chose to charge $90:

- Price = $90
- Average cost = $70
- Output = 13,000 units
- Hence, the profit for the firm = TR – TC = ($90 × 13,000) – ($70 × 13,000) = $260,000.

	Profit-maximizer	Profit-satisficer
Price	$110	$90
Output	10,000	13,000
Average cost	$70	$70
Profit (P × Q) – TC	$400,000	$260,000

In this case, the profit sacrificing firm settles for a lower acceptable level of profit ($260,000 instead of $400,000) by charging a price (between $90 and $110) and supplying a larger quantity (between 10,000 and 13,000).

Growth

Growth refers to increasing the size and scale of operations of a firm. While new firms tend to aim to break-even and survive, established businesses might aim for growth and to expand their operations. For example, firms such as Coca-Cola, Subway and McDonald's have expanded throughout most parts of the world. Growth can be measured in numerous ways, such as an increase in the firm's:

- sales revenue and/or profits
- product range (the different products sold by a business)
- number of employees
- number of stores/outlets
- volume of output and/or production capacity
- value of its assets or the capital invested.

Firms may aim to grow for numerous reasons, such as:

- To take advantage of economies of scale. By operating on a larger scale, firms can gain from cost-saving benefits such as financial, risk bearing and purchasing economies of scale (see Chapter 14). In the long run, this would lower their average costs of production, which makes them more profitable.
- Greater market power and the prestige associated with being a large firm.
- Being more competitive, as larger firms tend to be better positioned to withstand competitive pressures from rival firms. For example, Walmart and Amazon are market leaders in their industries, so they are in a strong position to continue dominating the market.
- Similarly, larger firms are more likely to survive during difficult trading periods, perhaps due to seasonal fluctuations in demand or an economic recession.

Firms can grow in several ways, including:

- Mergers and acquisitions, such as taking over another company (see Tiffany & Co case study below). This is a relatively quick method of growth but is highly expensive and is likely to create a change in corporate objectives and strategies.
- Market penetration growth strategies, which require firms to focus on increasing the sale of existing products in existing markets to existing customers. This means focusing on and developing what the firm knows and does well. It is a low-risk strategy so requires little, if any, investment in new market research as the firm concentrates on increasing the sales revenue or market share of its existing products, for example, by using more competitive pricing strategies or introducing customer loyalty schemes.
- Market development growth strategies, where a firm sells its existing products in new markets. In other words, the product remains the same, but it is sold to a new group of customers. For example, a car manufacturer might choose to export its cars to overseas markets or invest in production plants in other countries. This growth strategy carries some risks as customer tastes may vary in different regions and countries, and there are the added costs of market research. Nevertheless, the firm knows its products well so should be familiar with customer needs. This helps to minimize some of the risks involved.
- Product development growth strategies, where a firm introduces new products in existing markets, so it targets new products at existing customers. This is a common method of growth for many firms, such as car manufacturers creating and developing new cars to replace their existing models. These new cars are then marketed to existing customers. Another example is Apple's introduction of the iPhone, iPad, iTunes, Apple TV and Apple Watch. Product development is a medium-risk strategy as it often involves significant investment in research and development (R&D).
- Diversification growth strategies, which involve firms producing and selling completely new products to new customers. It is a high-risk growth strategy as firms enter markets in which they have little or no experience. Related diversification occurs when firms remain in a market or industry that they are familiar with. For example, Coca-Cola produced its first energy drinks in 2019. Unrelated diversification involves businesses entering new industries, in which they have no previous market experience. For example, Honda launched its HondaJet (aircraft) division in 2015. Lenovo introduced smartphones, watches and sports shoes(!) in 2015.

CASE STUDY

LVMH

In 2020, French luxury group LVMH bought American jeweller Tiffany & Co for $16.2 billion. This is the largest luxury brand acquisition in history, which adds Tiffany & Co to LVMH's huge product range including brands such as Louis Vuitton, Moët & Chandon, Christian Dior, Marc Jacobs, TAG Heuer and Sephora. Tiffany & Co was founded in 1837 in New York City by Charles Lewis Tiffany but failed to adapt to market changes prior to its takeover by LVMH.

PAPER 1 EXAM PRACTICE QUESTION 6.2

a Describe **two** cognitive biases that influence consumers when making consumption choices. *[10 marks]*

b Using real-world examples, evaluate the effectiveness of nudge theory in helping consumers to make better choices. *[15 marks]*

Chapter summary

- Rational choice theory states that individuals use logical and sensible reasons to determine the right choice associated with an individual's best self-interest.
- People might like to believe that they are always capable of making rational decisions, but biases (rule of thumb, anchoring and framing, availability) and other factors (such as imperfect information) prevent them from doing so.
- Behavioural economics is the study of actual economic decision-making, with emphasis on human behaviour being less rational than traditional economic theory suggests.
- Imperfect information occurs when people have inaccurate, incomplete or unreliable information, so decision-making is not optimal.
- Choice architecture is used to deliberately influence decision-making by designing and presenting different choices, and shaping the context in which people make decisions.
- Nudges are easy and low-cost techniques used by choice architects to change the social and economic behaviour of people without taking away or reducing their freedom of choice.
- Nudges can prevent decision makers (such as consumers and employees) from making irrational (illogical) or poor choices, thereby improving their welfare or well-being.
- The main business objective is profit maximization. Alternative business objectives include: corporate social responsibility (CSR), market share, satisficing and growth.

Top tip!

While firms may have alternative objectives, they are not mutually exclusive. For example, CSR practices can help to improve a firm's corporate image, thereby enabling it to grow and increase its market share. Essentially, this can help the firm to maximize long-term profits.

Theory of Knowledge

How can economists know whether a firm is genuinely acting in a socially responsible way or whether it is only truly complying with social and cultural norms?

7 Elasticity of demand – price elasticity of demand (PED)

Real-world issue

How do consumers and producers make choices in trying to meet their economic objectives?

Conceptual understandings (WISE ChoICES)

- Consumer and producer **choices** are the outcome of complex decision-making.
- Welfare is maximized if allocative **efficiency** is achieved.
- Constant **change** produces dynamic markets.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

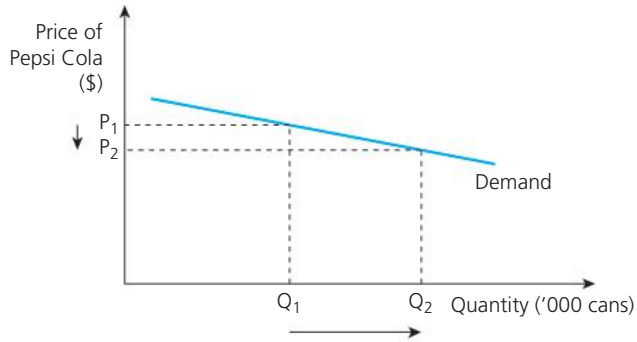
- the concept of elasticity (AO1)
- price elasticity of demand, including the formula (AO2)
- the degrees of PED, that is, the theoretical range of values for PED (AO2)
- diagrams to show relatively price elastic and relatively price inelastic demand curves (AO4)
- relevant diagrams to show constant PED: perfectly elastic, perfectly inelastic and unitary PED along a demand curve (AO4)
- the changing value of PED along a linear downward sloping demand curve (AO3) (HL only)
- a diagram that shows how the value of PED changes along a linear demand curve (AO4) (HL only)
- the determinants of PED (AO2)
- the relationship between PED and total revenue (AO2)
- a diagram that shows changes in revenue as a result of price changes when demand is price elastic and price inelastic (AO4)
- PED, changes in price, quantity demanded or total revenue from data provided (AO4)
- the importance of PED for firms and government decision-making (AO3)
- the reasons why the PED for primary commodities is generally lower than the PED for manufactured products (AO2) (HL only).

Concept of elasticity (AO1, AO4)

Elasticity of demand refers to the responsiveness of quantity demanded for a good due to a change in a factor that affects demand, such as the price of the product or a change in consumers' real disposable income (the money available after the impact of inflation, or higher prices, and taxes have been accounted for). Demand is **price elastic** or income elastic if a slight change in the price or income leads to a large change in the demand for the product. Hence, demand is relatively elastic if it is sensitive to changes in price. By contrast, demand is **price inelastic** if a change in price or income has little impact on the demand for a good or service. In this case, demand is relatively inelastic if it is comparatively unresponsive to changes in price.

◆ **Elasticity of demand** is a measure of how the quantity demanded of a product changes owing to a change in a factor that affects demand, such as price or real disposable income.

Price elastic demand occurs when a relatively small change in the price of a product causes a larger percentage change in the quantity demanded. This means that customers are highly responsive to the change in price (see Figure 7.1). This is because the increase in price gives consumers less of an incentive to buy the product (as per the law of demand) and go without or switch to substitute products instead. Examples of products with price elastic demand include motor vehicles, overseas holidays, fast food, fruits and vegetables, supermarket own-brand goods and chocolate bars.



■ **Figure 7.1** The price elastic demand curve

In Figure 7.1, as the price of Pepsi Cola falls from P_1 to P_2, the quantity demanded rises by a greater proportion from Q_1 to Q_2 as existing customers are more willing and able to buy the product and other customers switch from rival soft drinks brands such as Coca-Cola, Fanta and 7-Up.

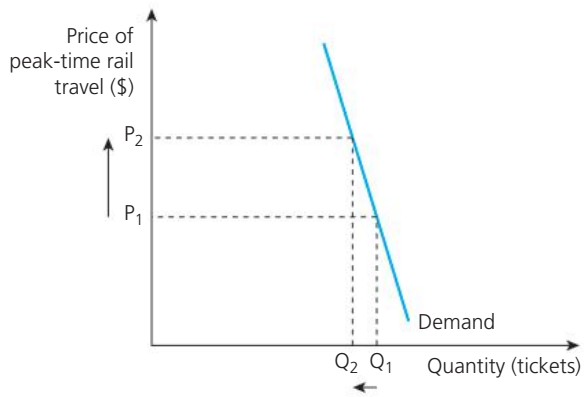
■ **Figure 7.2** The demand for soft drinks is price elastic

◆ **Price elastic demand** describes demand for a product that is relatively responsive to changes in the product's price, usually due to close substitutes being widely available. In this case, the value of PED > 1.

◆ **Price inelastic demand** describes demand for a product that is relatively unresponsive to changes in the product's price, mainly due to the lack of close substitutes being available. In this case, the value of PED < 1.

By contrast, **price inelastic demand** occurs when a small change in the price of a product causes a smaller proportionate change in the quantity demanded. In other words, customers are not very responsive to the change in price (see Figure 7.3). This is mainly because there is a lack of substitute products for consumers to switch to. Examples of products with price inelastic demand include salt, safety matches, petrol (gas), medications, alcohol, electricity, examination registration fees, cigarettes and nail clippers.

In Figure 7.3, as the price of peak-time rail travel rises from P_1 to P_2, the quantity demanded falls by a smaller proportion from Q_1 to Q_2, as commuters are somewhat unresponsive to the change in the price as they need to get to work. For the small number of commuters who do switch, this may be owing to alternatives being available such as taking the bus, cycling or walking to work.



■ **Figure 7.3** The price inelastic demand curve

Common mistake

Students often state that 'price elastic' means that as the price of a good or service increases, its demand falls. This describes the *law of demand*. In other words, as price goes up the quantity demanded will fall, irrespective of whether demand is price elastic or price inelastic. Instead, if demand is price elastic the percentage fall in demand is greater than the percentage increase in price.

Top tip!

It may be helpful to remember that **in**elastic demand occurs when customers are **in**sensitive (unresponsive) to small changes in price.

PAPER 1 EXAM PRACTICE QUESTION 7.1

Explain why the price elasticity of demand (PED) for concert tickets is likely to differ from the PED for airline tickets. *[10 marks]*

Price elasticity of demand (AO2, AO4)

Price elasticity of demand (PED) measures the degree of responsiveness of quantity demanded for a product following a change in its own price. For any product, demand can be price elastic or price inelastic, depending on the percentage change in the quantity demanded after a change in the price of the product.

◆ **Price elasticity of demand** (PED) measures the degree of responsiveness of quantity demanded for a product following a change in its own price.

Price elasticity of demand is calculated using the formula:

$$\frac{\text{Percentage change in quantity demanded}}{\text{Percentage change in price}}$$

Or in annotation:

$$\text{PED} = \frac{\%\Delta Q_D}{\%\Delta P}$$

◆ A **change in quantity demanded** refers to a movement along a demand curve, that is, a change in quantity demanded when price changes.

The mathematical value of PED is negative owing to the law of demand. In other words, an increase in the price of a product tends to cause a fall in the quantity demanded, and vice versa. However, it is conventional to ignore the negative coefficient when interpreting the value of PED. The value of PED for a product depends on the degree of customers' ability and willingness to continue buying the product if the price changes. For example, a rise in the price of a necessity product will have a minimal impact on the quantity demanded.

Top tip!

There is no need to consider the negative coefficient when interpreting the value of PED – the negative sign is simply the result of the law of demand. For example, when price increases, the quantity demanded will fall, ceteris paribus.

Top tip!

To calculate the percentage change between two numbers, simply use the formula:

$$\frac{\text{New figure} - \text{Old figure}}{\text{Old figure}} \times 100$$

If PED > 1 then demand is **price elastic**. In other words, customers are responsive to changes in price (see Figure 7.1). This is because the percentage change in price causes a larger proportionate change in the quantity demanded.

If PED < 1 then demand is **price inelastic**. In other words, customers are relatively unresponsive to changes in price (see Figure 7.3). This occurs when customers find it difficult to react swiftly to changes in price, perhaps owing to the lack of available substitutes.

The range of values for PED (or the degree of PED), ignoring the minus coefficient, is outlined in Table 7.1.

Table 7.1 Degrees of PED values

PED value	Degree of PED	Diagram (PED along a demand curve)
PED > 1	Price elastic demand	Relatively elastic demand
PED < 1	Price inelastic demand	Relatively inelastic demand

PAPER 2 AND 3 EXAM PRACTICE QUESTION 7.2

A retailer reduces the price of designer label sunglasses from $300 to $250 each, resulting in an increase in the quantity demanded from 85 to 95 pairs of sunglasses per month. Calculate the value of PED and comment on the result. *[4 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 7.3

A jeweller reduces the price of platinum earrings from $400 to $350 per pair, causing an increase in the quantity demanded from 25 units to 30 pairs per month.

a Calculate the value of the PED for the earrings and comment on the result. *[10 marks]*

b Examine the significance of price elasticity of demand (PED) for the decision-making of firms and governments. *[4 marks]*

Degrees of PED

The theoretical range of values for constant PED along a demand curve, ignoring the minus coefficient, is outlined in Table 7.2.

Table 7.2 Degrees of PED values

PED value	Degree of PED	Diagram (PED along a demand curve)
PED = 0	Perfectly price inelastic	Constant PED – perfectly inelastic
PED = ∞	Perfectly price elastic	Constant PED – perfectly elastic
PED = 1	**Unitary price elastic demand**	Constant PED – unitary PES

◆ **Unitary price elastic demand** occurs when a given price change of a product leads to the same percentage change in the quantity demanded. Hence, the PED value = 1

Perfectly price elastic demand

If PED = ∞ (infinity) then demand is *perfectly price elastic*. In other words, demand exists at one price only, and any change in price leads to zero demand. This theoretical extreme suggests that customers will switch to buying other substitutes that are readily available if firms increase their price (see Figure 7.4).

In Figure 7.4, demand exists only at a price of P_e. An increase in price leads to an infinite change in the quantity demanded. This hypothetical situation exists only if there are perfect substitutes readily available on the market that customers can easily access.

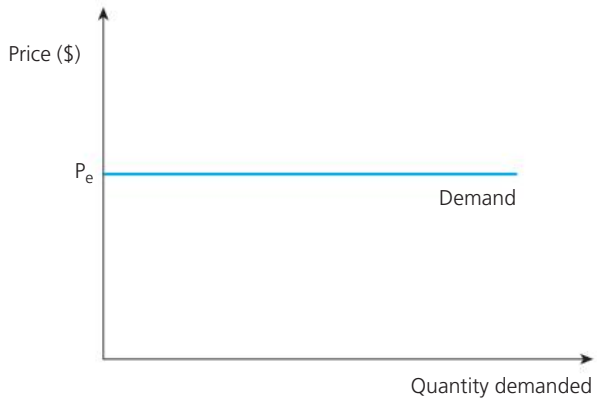


Figure 7.4 The perfectly price elastic demand curve

Perfectly price inelastic demand

When the value of PED is zero, demand is perfectly price inelastic. In other words, a change in price has no impact on the quantity demanded (see Figure 7.5). This theoretical extreme suggests that there are no substitutes for the product.

In Figure 7.5, as the price increases from P_1 to P_2, the quantity demanded remains unchanged at Q_e. In reality, demand will not exist irrespective of all price levels, but the demand for prescription drugs, anti-venom, clean water or veterinary services would be close to being perfectly price inelastic.

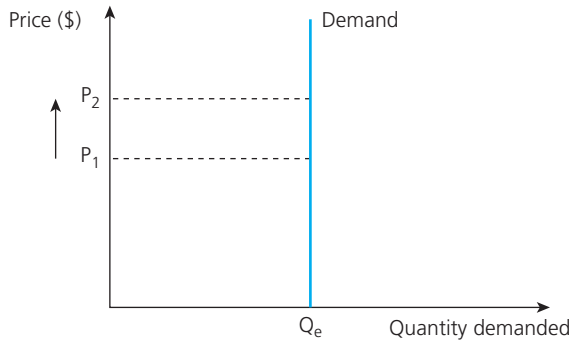


Figure 7.5 The perfectly price inelastic demand curve

Figure 7.6 The demand for veterinary services is price inelastic

Unit elastic demand

When the PED value = 1.0, demand is said to be of *unit elastic demand*. This occurs when a given price change of a product leads to the same percentage change in the quantity demanded (see Figure 7.7). This means that total spending by customers on the product will remain the same, at all given prices.

In Figure 7.7, as the price increases from P_1 to P_2, the quantity demanded changes by the same proportion, reducing quantity demanded from Q_1 to Q_2. As the percentage change in quantity demanded is the same as the percentage change in price, there is no change to the firm's sales revenue. Before the price change, total revenue was represented by the area $0P_1BQ_1$, but changes to area $0P_2AQ_2$ after the price rise – both areas are of the same value.

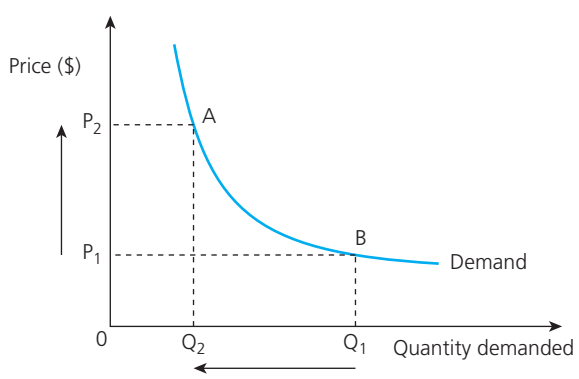


Figure 7.7 The unit price elastic demand curve

Top tip!

Whether the demand for a product is price elastic or price inelastic really depends on the breadth of definition of the product. A broadly defined good or service (such as food rather than fruit, meat, apples or salmon) will be more price inelastic. There is clearly no real substitute for food! However, it is perhaps more useful to measure the PED for specific brands or products, such as carbonated soft drinks or Canadian beef.

PAPER 1 EXAM PRACTICE QUESTION 7.4

Explain why the price elasticity of demand for many primary commodities (such as crude oil or coffee beans) has a relatively low value, while the demand for manufactured products (such as laptops) has a relatively high PED value. *[10 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 7.5

Suppose the demand for football match tickets at $50 is 50,000 per week. If the football club increases the price to $60 per ticket and quantity demanded subsequently falls to 45,000 per week, calculate the value of the price elasticity of demand and comment on the result. *[4 marks]*

Changing PED along a straight-line downward sloping demand curve (HL only)

The value of price elasticity of demand (PED) will change along a normal downward sloping linear demand curve. The value of PED increases as the price of the product rises (and vice versa), as illustrated in Figure 7.8. This occurs because customers will be more responsive to changes in prices at higher levels (when the price accounts for a greater proportion of consumers' income).

PAPER 1 EXAM PRACTICE QUESTION 7.6

Explain why price elasticity of demand (PED) varies along the length of a straight-line demand curve. [*10 marks*]

Mathematically, the value of PED at the mid-point along a linear demand curve is equal to 1.0. This is shown in Figure 7.8, where the price is P_e and the quantity demanded is Q_e. At all prices above this point, PED is greater than 1 (that is, demand is price elastic) as customers become more sensitive to increases in price. This is because the higher prices account for a greater percentage of their income. When price becomes too high, the quantity demanded drops to zero (where the demand curve intersects the *y*-axis), so the PED value = ∞ (infinity). By contrast, at all prices below P_e, the PED value will be below 1.0 as customers are far less sensitive to reductions in price. The lower prices account for a smaller proportional of household income, so customers become less sensitive to the price changes.

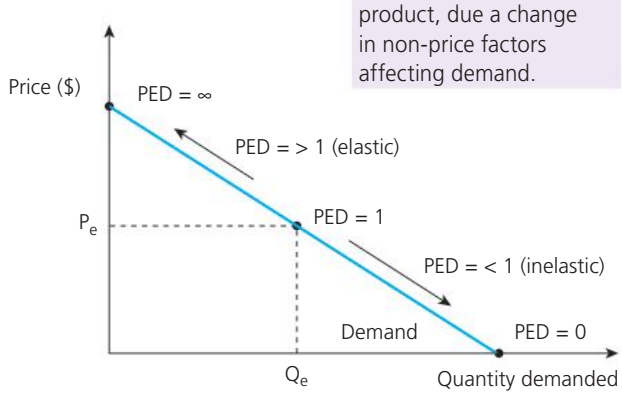


Figure 7.8 PED along a linear demand curve

Common mistake

It is technically incorrect to say that a product is 'elastic' or 'inelastic'. Instead, use the correct economics terminology by saying that the product is either 'price elastic' or 'price inelastic' to signify that PED refers to the responsiveness of demand to changes in the price of the product.

Top tip!

Despite the gradient of the linear demand curve being the same, the percentage **change in demand** is greater at higher price levels. When calculating PED, it is important to remember that economists use percentage changes, not absolute changes, in price and quantity demanded.

◆ A **change in demand** refers to a shift in the demand curve for a product, due a change in non-price factors affecting demand.

Determinants of price elasticity of demand (AO2)

There are four main determinants of price elasticity of demand (PED): the number and closeness of substitutes, the degree of necessity, the proportion of income spent on the product, and the time period under consideration.

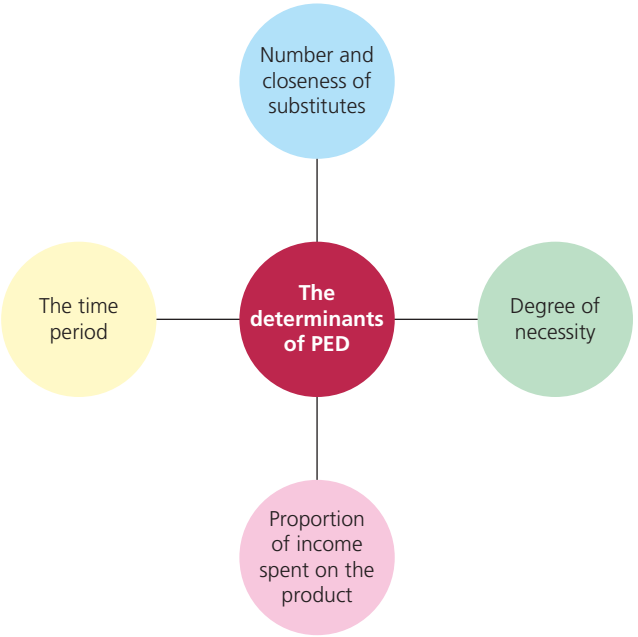


Figure 7.9 The determinants of price elasticity of demand (PED)

Top tip!

The determinants of price elasticity of demand can be remembered by the acronym TINS: **t**ime period, **i**ncome (the proportion of income spent on the good or service), **n**ecessity (the degree of necessity of the product) and **s**ubstitution (the number and closeness of available substitutes).

Number and closeness of substitutes

The greater the number and availability of close substitutes there are for a good or service, the higher the value of its PED will tend to be. This is because consumers can readily switch to different substitute products or brands if there is a small change in the price of the product. For example, if the price of Fanta goes up, people can easily switch other soda brands such as Tango, Orangina, 7-Up, Dr Pepper, Pepsi and so forth. By contrast, products with no or few substitutes (such as private education or prescribed medicines) have price inelastic demand. For these products, a small change in the price will cause a smaller percentage fall in the quantity demanded as customers do not have much, if any, choice.

The number and closeness of substitutes can be reduced by firms deliberately attempting to increase the opportunity costs of switching to rival brands. If there are high costs involved for customers to switch between brands or products, then demand tends to be price inelastic. For example, mobile (cell) phone subscribers are bound by lengthy contracts, thus switching between rival services is made less easy. Customer loyalty programmes also make the opportunity cost of switching more difficult, thereby reducing the value of the PED for the firm's products.

Figure 7.10 Customer loyalty schemes can make switching to rival brands more difficult

Degree of necessity

Products that are regarded as essential (such as food, fuel, drinking water and housing) tend to be price inelastic as there is unlikely to be a significant drop in the quantity demanded even if prices rise. Hence, the demand for necessities is relatively unresponsive to changes in price. By contrast, the demand for luxury products (such as diamonds, chandeliers and large OLED televisions) is relatively price elastic. These are not necessities, so customers can do without, or choose from cheaper alternatives, if the price increases. Hence, an increase in the price of luxuries will tend to cause a more significant fall in the quantity demanded.

The degree of necessity depends on the definition of what is deemed to be a necessity for individuals. For example, rice is a staple food product for many people, but not for everyone. If a product is addictive (such as tobacco), highly fashionable (such as smartphones), or an essential part of a hobby (such as helmets and saddles for horse-riding), its PED tends to be relatively price inelastic. This is because the consumer has little choice but to pay the market price, even if it is increased. For example, if the price of cigarettes goes up, it is unlikely to have a large effect on the quantity demanded.

Proportion of income spent on the product

The greater the proportion of consumers' real income spent on a good or service, the more price elastic demand will be, ceteris paribus. This is because consumers are highly sensitive to changes in the price of products that account for a larger proportion of their income (such as jewellery, sports cars and overseas holidays). By contrast, if products account for only a very small proportion of people's income, then any increase in price will have a minimal impact on their spending. For example, even if the price of toothpicks or safety matches went up by 50 per cent, it is unlikely to cause people to reduce their purchases by that much (see Figure 7.11).

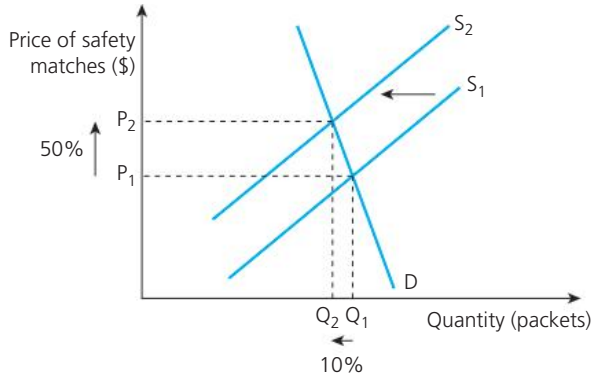


Figure 7.11 Price inelastic demand curve for safety matches

In Figure 7.11, a 50 per cent increase in the price of safety matches, caused by the leftwards shift of the supply curve from S_1 to S_2, only causes quantity demand to fall by 10 per cent. Hence, the PED value is 0.2, price inelastic. This is because the higher price of safety matches still accounts for only a small proportion of people's income, so they are relatively unresponsive to the change in the price.

Time

The time period under consideration is a key determinant of the value of PED. Over time, a good tends to become more price elastic because consumers and firms have more time to find alternatives or substitutes. People also need time to change their habits and preferences. For example, diesel fuel is more expensive than leaded fuel, but has greater fuel efficiency (miles or kilometres per litre). However, with time, the higher tax and price of diesel fuel has meant an increasing number of drivers switch to unleaded fuel. Over an even longer period of time, these drivers are likely to switch to electric cars as the price per mile or kilometre is lower.

Time also plays a role in when customers change their purchases or buying habits. The more durable a product is, such as furniture and home appliances, the more price elastic the demand tends to be. This is because higher prices of durable products mean there is no urgency for consumers to replace their existing furniture and home appliances if prices are increased.

Top tip!

The key determinant of the value of PED for a good or service is the degree of substitution. This means, the extent to which consumers have access to close alternatives for the product. The greater the availability and closeness of substitutes, the more price elastic demand will be.

Relationship between PED and total revenue (AO2, AO4)

The value of price elasticity of demand shows changes in a firm's revenue as a result of price changes, depending on whether demand is price elastic or price inelastic. A firm that has price inelastic demand for its product can increase its prices to earn more sales revenue (see Figure 7.12).

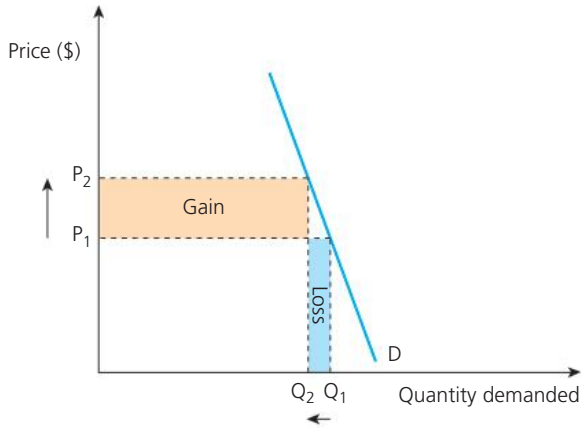


Figure 7.12 Price inelastic demand and sales revenue

In Figure 7.12, the total revenue before the increase in price is the area $P_1 \times Q_1$. After the increase in price, total revenue changes to $P_2 \times Q_2$. As demand is price inelastic, consumers are largely unresponsive to the change in the price. This means increasing the price from P_1 to P_2 results in a net gain in total revenue because the higher price per unit is greater than the percentage fall in quantity demanded. In this case, the gain in total revenue (shown by the orange shaded area) is greater than the loss (the blue shaded area). Essentially, a firm's total revenue increases by increasing price when demand is price inelastic.

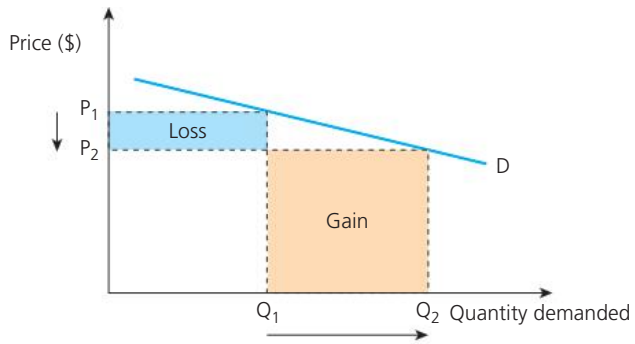


Figure 7.13 Price elastic demand and sales revenue

By contrast, a firm that has price elastic demand for its product can earn more sales revenue by reducing its price (see Figure 7.13).

In Figure 7.13, the total revenue before the increase in price is the area $P_1 \times Q_1$. After the reduction in price, total revenue changes to $P_2 \times Q_2$. As demand is price elastic, consumers are highly responsive to the change in the price. This means reducing the price from P_1 to P_2 results in a net gain in total revenue because the lower price leads to a greater percentage increase in the quantity demanded. In this case, the gain in total revenue (shown by the orange shaded area) is greater than the loss (the blue shaded area). Logically, if demand is highly price elastic and the firm *increased* price, customers would simply switch to other close substitutes that are widely available, thereby generating a net loss in total revenue. Essentially, a firm's total revenue increases by reducing price when demand is price elastic.

Table 7.3 The relationship between PED and sales revenue

Price change	Inelastic	Unitary	Elastic
Increased price	Total revenue increases	No change in revenue	Total revenue falls
Reduced price	Total revenue falls	No change in revenue	Total revenue increases

Common mistake

Consider the example below:

The demand for theatre tickets at $50 is 5,000 per week. If the theatre raises its price to $60 per ticket and demand subsequently falls to 4,500 per week, calculate the value of PED. *[2 marks]*

Too often, students place the calculation of the percentage change in price as the numerator, rather than the denominator in the formula for PED. This may be because they calculate the price change first, since this is what causes the quantity demanded to change. In this case:

- $\%\Delta P = (\$60 - \$50) \div 50 \times 100\% = +20\%$
- $\%\Delta Q_D = (4{,}500 \div 5{,}000) \div 5{,}000 \times 100\% = -10\%$
- Therefore, PED = +20% ÷ –10% = –2.0 (price elastic).

This is a classic error, where the student does the right calculations, but then puts the figures the wrong way around in the PED formula: $\%\Delta Q_D \div \%\Delta P$. Instead, the correct answer should be:

PED = –10% ÷ +20% = –0.5 (price inelastic).

Common mistake

'A lower price means quantity demanded will rise, so the firm earns more revenue.'

While a lower price generally increases the quantity demanded (the law of demand), the statement above is true only if demand is price elastic. Quantity demanded increases if the price falls, ceteris paribus, but revenue actually falls if demand is relatively unresponsive to the lower price.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 7.7

A firm has a price elasticity of demand of –3.0 for its product. At a price of $20, the firm sells 200 units of output. Calculate the effect on the firm's sales revenue of increasing the price to $22. *[3 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 7.8

A firm sells 200 units of its product each day at a price of $4, with a known price elasticity of demand of –2.0.

a Calculate Juke Ltd's sales revenue. *[2 marks]*

b Calculate the new sales revenue if Juke Ltd increases its price by 20%. *[3 marks]*

c Explain whether raising price was a good decision for Juke Ltd. *[2 marks]*

Importance of PED for firms and government decision-making (AO3)

Price elasticity of demand (PED) is important for firms as it informs business decision makers about the effect of price changes on sales revenue (price multiplied by quantity traded, or P × Q). When considering whether to change price, a firm needs to determine the resulting impact on quantity demanded and therefore the sales revenue.

Firms with different PED values for their products can charge different customers different prices for essentially the same product (a strategy known as *price discrimination*). For example, cinemas and theme parks charge adults higher prices and offer discounts to children, students and the elderly in order to increase their revenues.

Admission prices

	Peak	Off-peak
Adult	$5.25	$3.85
Child	$3.85	$3.05
Family of 4	$16.75	$12.50
Students	$4.25	$3.25

Figure 7.14 Price discrimination is used by firms to increase sales revenue

Similarly, airlines use *surge pricing* (also known as *dynamic pricing*) to hike prices depending on the time of the year; for example, airline prices soar during the summer and Christmas holiday seasons when there is a peak in the number of travellers. Taxi services such as Uber and Grab also use surge pricing when demand spikes, during peak travelling periods. Essentially, firms with the ability to do so will charge higher prices when demand is price inelastic (and vice versa).

Governments use PED values to determine taxation policies. For example, imposing heavy taxes on demerit goods (see Chapter 11) such as cigarettes and alcohol, knowing that the demand for such products is highly price inelastic (see Figure 7.15). This means that the tax has to be sufficiently high to discourage consumption, while generating a lot of tax revenue for the government.

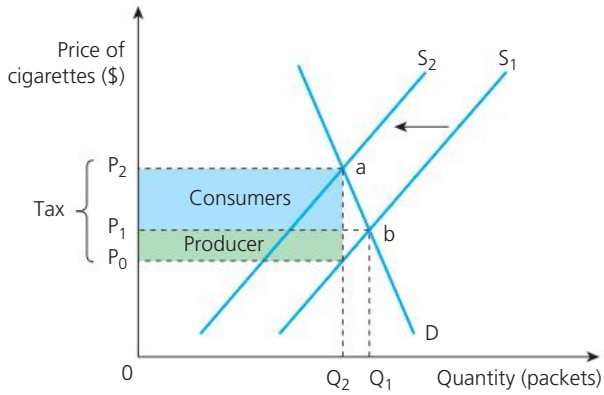


Figure 7.15 Price inelastic demand curve for cigarettes

In Figure 7.15, the government imposes an indirect tax (a tax imposed on expenditure, rather than on income) on the purchase of cigarettes, a price inelastic product. The government imposes a per unit tax, shown by the vertical distance between P_2 and P_0. This forces the equilibrium price up from P_1 to P_2 (point b to point a). Hence, consumers pay most of the tax burden, shown by the shaded blue area. Producers absorb the remainder of the tax burden, shown by the shaded green area. These combined areas show the tax revenue collected by the government. The huge increase in price does little to reduce the consumption of cigarettes but collects a lot of tax revenue for the government. This knowledge of PED also informs the government that the significant increase in price does little harm to the cigarette manufacturing firms and jobs in the industry.

In the above situation, firms are also aware that they can pass on most of the incidence of the indirect tax to the customers. This is because cigarettes are highly price inelastic. The same applies to other price inelastic products that are deemed as demerit goods, such as alcohol, petrol (gas) and gambling.

Reasons why the PED for primary commodities is generally lower than the PED for manufactured products (AO2) (HL only)

The value of PED for primary commodities (raw materials) such as crude oil and iron ores is relatively low. This is due to the main determinants of PED for primary products:

- Primary commodities such as rice lack close substitutes. For example, there are few alternatives for coal, crude oil, gold, metal ores and rice.
- They are essential (necessities) for production, so their demand is relatively price inelastic. In other words, relatively unresponsive to changes in price.
- The proportion of income spent on primary products is relatively low. The price of primary commodities used for production tends to account for a smaller proportion of overall costs of production, as labour and capital costs tend to be higher.
- The time needed to grow and harvest or extract primary commodities is a lot longer than the process to manufacture goods. The shorter the time period under consideration, the lower the value of PED.

By contrast, the demand for manufactured products (such as motor vehicles, laptops, watches or furniture) is relatively price elastic. The factors affecting PED can also be used to explain the relatively high PED value:

- Most manufactured products have many substitutes; for example, different makes and models of cars, laptops and watches.
- The degree of necessity is lower, as customers might be able to use their existing manufactured products (such as furniture) for longer in response to higher market prices or can switch to cheaper alternatives.
- Manufactured products such as motor vehicles take up a larger proportion of consumers' income and therefore demand is relatively price elastic.
- Manufactured products such as laptops can be used continuously over a long period of time, so PED tends to be higher in value.

Figure 7.16 The demand for mass manufactured cars is relatively price elastic

ALT ACTIVITY 7.1

To what extent do businesses in the real world make use of price elasticity of demand? To help answer this question, investigate how knowledge of price elasticity of demand helps a certain business (such as a private tuition centre) or an industry (such as airline companies) by informing their pricing decisions.

Theory of Knowledge

To what extent do emotions rather than reasoning play a role in the purchasing decisions of consumers?

Theory of Knowledge

When evaluating the merits of economic theories, such as elasticity of demand, how important is it to consider cultural differences and local contexts?

WISE ChoICES

Evaluate the role of government **intervention** in order to influence **choice**.

In general, the demand for primary commodities is relatively price inelastic (so has a PED value of less than 1), while the demand for manufactured products will tend to be price elastic (so PED > 1), due to the main factors affecting PED. For example, the PED for cattle hide (a primary commodity) to produce leather products is relatively lower than the PED for leather belts or leather shoes (manufactured products). This is because there are very few substitutes for cattle leather as a raw material for producing leather belts or shoes. However, leather manufactured belts or shoes can have many substitutes in the form of artificial leathers, bonded leathers, suede, or alternative leather sources (such as leather from sheep, goat, kangaroo and alligator).

WISE ChoICES

To what extent can government **intervention** lead to positive **change** in the economy?

Chapter summary

- Price elasticity of demand (PED) measures the degree of responsiveness of quantity demanded for a product due to a change in its price.
- PED is calculated using the formula: $\%\Delta Q_D \div \%\Delta P$.
- Demand is *price elastic* if there is a relatively large change in the quantity demanded for a product following a relatively small change in its price. In other words, customers are responsive to changes in price.
- Demand is *price inelastic* if a price change causes a relatively small or insignificant change in the quantity demanded. In other words, buyers are rather unresponsive to changes in price.
- Determinants of PED include: the number and closeness of substitutes, the degree of necessity, the proportion of income spent on the good or service, and the time to seek alternatives and/or adjust to the use of substitutes.
- The key determinant of PED is the degree of substitution – products with only a few or no close substitutes have price inelastic demand.
- The greater the proportion of consumers' income spent on a good or service, the more price elastic its demand tends to be.
- A firm that faces price inelastic demand should raise prices in order to earn more sales revenue. By contrast, a firm that has price elastic demand should reduce prices in order to earn higher sales revenue.
- Knowledge of PED enables producers to determine their pricing strategies (to maximize sales revenues) and for governments to determine their tax policies (to maximize tax revenues).
- Knowledge of PED also enables firms to determine their pricing strategies in order to increase sales revenue and profits.
- On a normal (downward sloping) linear demand curve, the value of PED increases as the price of the product rises, and vice versa. (HL only)
- The PED for primary commodities is relatively low (price inelastic) because they have very few substitutes, whereas manufactured products have a relatively high PED (price elastic) due to the existence of many substitutes. (HL only)

8 Elasticity of demand – income elasticity of demand (YED)

Real-world issue

How do consumers and producers make choices in trying to meet their economic objectives?

Conceptual understandings (WISE ChoICES)

- Consumer and producer **choices** are the outcome of complex decision-making.
- Welfare is maximized if allocative **efficiency** is achieved.
- Constant **change** produces dynamic markets.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- income elasticity of demand (YED), including the formula (AO2)
- income elastic demand and income inelastic demand (AO2)
- the significance of YED sign for normal goods (necessities and luxury goods), and inferior goods (AO2)
- Engel curve diagrams to show income elastic and income inelastic demand, including for inferior goods (AO4)
- the value of YED, changes in income and quantity demanded from data provided (AO4)
- the importance of YED for firms (AO3) (HL only)
- the importance of YED in explaining changes in the sectoral structure of the economy (AO3) (HL only).

Income elasticity of demand (AO2, AO4)

Income elasticity of demand (YED) measures the degree of responsiveness of demand following a change in income. YED is calculated using the formula:

$$\frac{\text{Percentage change in quantity demanded}}{\text{Percentage change in income}}$$

Or in annotation:

$$\text{YED} = \frac{\%\Delta Q_D}{\%\Delta Y}$$

Real disposable income is a key determinant of the demand for goods and services as this determines the purchasing power of consumers. Individuals with higher levels of real disposable incomes spend more on consumer goods and services as compared to lower income earners. Income elasticity of demand looks at how changes in real income affect the demand curve for a particular good or service.

◆ **Income elasticity of demand (YED)** measures the degree of responsiveness of quantity demanded following a change in the real income of consumers.

◆ **Income elastic demand** occurs when the percentage change in the quantity demanded of a product is greater than the percentage change in consumers' real income, that is, PED > 1. This tends to apply to the demand for services and luxury products.

Income elastic demand and income inelastic demand

When demand is relatively responsive to a change in consumers' real income, there is said to be **income elastic demand**. The value of PED > 1 as the percentage change in quantity demanded is greater than the percentage change in real income. This tends to apply to the demand for **services**

◆ **Services** are intangible (non-physical) products such as haircuts, bus journeys, telephone calls and internet access.

and **luxury products**. For example, if average incomes increase by 5 per cent, the consumer may spend proportionately more on designer clothing, dining out in restaurants or taking vacations abroad.

◆ **Luxury products** are those with income elastic demand, that is goods and services with a YED value greater than 1.

Figure 8.1 The demand for overseas vacations is income elastic

◆ **Income inelastic demand** occurs when the percentage change in the quantity demanded of a product is less than the percentage change in consumers' real income, that is, PED < 1. This applies to the demand for necessities.

When demand is relatively unresponsive to a change in consumers' real income, there is said to be **income inelastic demand**. The value of PED < 1 as the percentage change in quantity demanded is less than the percentage change in real income. This applies to the case of goods deemed to be **necessities**, such as staple food products, toothpaste, toilet tissue and household electricity.

◆ **Necessities** are essential products that consumers buy in order to meet their needs. They have a positive PED value that is less than 1, that is, the proportionate change in quantity demanded is less than the proportionate change in real incomes.

Significance of the YED sign

While it is the norm to ignore the negative sign when interpreting PED, it is vital to consider the positive or negative sign when looking at the value of YED. The negative coefficient of YED suggests the product is an inferior good. In other words, less is bought as real income levels rise. By contrast, **normal goods** have a positive YED value because the quantity demanded increases along with an increase in real income.

◆ **Normal goods** are products that customers tend to buy more of as their real income level increases. This includes both necessities and luxury goods and services.

Normal goods are products that customers tend to buy more of as their real income level increases, ceteris paribus. Hence, normal goods have a positive YED. Normal goods include both necessities (such as staple food products) and luxuries (such as cars, jewellery and overseas holidays).

Necessities are normal goods with a positive YED value of less than one. This means that demand is income inelastic. In other words, the proportionate change in quantity demanded is less than the proportionate change in real incomes. This is because customers will continue to buy such essential products to meet their needs, even if the price of necessities increases. They do not buy significantly more because of an increase in real incomes. For example, if people buy 5 per cent more rice following a 10 per cent increase in their real income, then the YED is +0.5 (because 5% ÷ 10% = 0.5, that is, income inelastic).

Luxuries are normal goods with a positive YED value of greater than one. This means that an increase in real consumer income leads to a proportionally greater increase in the demand for

◆ **Luxuries** are superior goods and services with a YED value of greater than one, that is, their demand is highly income elastic. An increase in real consumer incomes leads to a proportionally greater increase in the demand for luxuries.

■ **Figure 8.2** Glasses (spectacles) are a necessity good for many people

Top tip!

All luxury goods are normal goods, but not all normal goods are necessarily luxury goods.

■ **Figure 8.3** Diamond rings are luxury goods

luxuries. This means that customers are relatively responsive to changes in real income levels when buying luxury goods and services (superior products), such as diamond rings; hence, the demand for luxuries is income elastic. For example, a 6 per cent increase in real income might lead to a 10 per cent rise in the demand for diamond rings. The income elasticity of demand in this example would be +1.67 (because 10% ÷ 6% = 1.67, that is, income elastic).

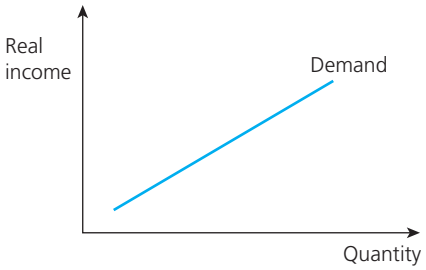
In economics, Engel curves are used to demonstrate the relationship between income and the quantity demanded. The **Engel curve**, named after the German statistician Ernst Engel (1821–96), shows the positive or negative relationship between the quantity demanded for a good or service (on the *x*-axis) and the real income of consumers (on the *y*-axis). If the slope of the Engel curve is positive, the product is classified as a normal good, but if it has a negative slope, then the product is an inferior good.

◆ The **Engel curve** is used to demonstrate the relationship between income and the quantity demanded. Normal goods (services and luxuries) have a positive slope, whereas inferior goods have a negative slope along an Engel curve.

Figure 8.4 illustrates the positive YED value for normal goods (necessities and luxuries). The demand curve is upward sloping (note the *y*-axis label is labelled 'income' not 'price') because people will tend to buy more of these products as real incomes rise, ceteris paribus. Hence, in the case of normal goods, there is a direct (or positive) relationship between income changes and quantity demanded.

The value of YED also reveals what is likely to happen if the real income level falls. For example, if the income elasticity of demand for fresh flowers is +2.1, then a 10 per cent fall in consumer incomes will lead to a 21 per cent fall in the demand for fresh flowers. Luxury goods and services are the most affected products during an economic downturn (a recession). In other words, when real GDP in the economy declines.

For normal goods, income elasticity of demand can be price elastic or price inelastic. The YED is >1 (income elastic) for luxury products such as sports cars, overseas holidays, premium economy and business class air travel, and expensive consumer durables (such as laptops and smartphones). This means any given change in real income will lead to a greater percentage change in quantity demanded. In Figure 8.5, the 8 per cent increase in real income causes a 20 per cent increase in the quantity demanded for fine wines. Hence, the YED in this case is +2.5, that is, income elastic.



■ **Figure 8.4** The Engel curve for normal goods

The YED is < 1 (income inelastic) for necessity products such as staple foods (for example, rice, bread and vegetables), public transport, medicines, utility bills (for power and water), ordinary home appliances (for example, a kettle or microwave

oven), and mobile network providers (for telephone and Wi-Fi services). This means any given change in real income will lead to a smaller percentage change in quantity demanded. In Figure 8.6, an 8 per cent increase in real income causes only a 2 per cent increase in the quantity demanded for rice. Hence, the YED in this case is +0.25, that is, income inelastic.

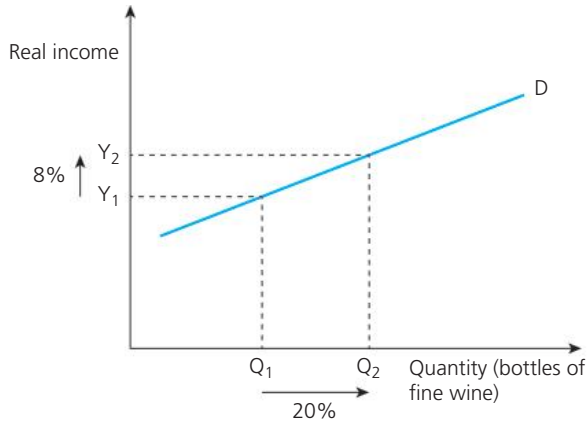


Figure 8.5 Income elastic demand

ATL ACTIVITY 8.1

Economists use empirical data such as YED values to determine the goods and services deemed to be necessities. However, the term 'necessity' is somewhat subjective. Investigate the top 10 necessities of modern life in a country of your choice – what is it that you simply cannot live without? Be prepared to share your findings and justifications.

By contrast, **inferior goods** are products that customers tend to buy less of as their real income level rises. Hence, the income elasticity of demand for inferior goods and services has a negative coefficient. This is because consumers substitute inferior or less desirable products for more superior ones. For example, people might choose to buy fresh meats instead of frozen meats when their real income levels increase. Note that in economics, inferior goods do not mean substandard quality goods but instead relates to the affordability of the goods. With an increase in consumers' real incomes, with prices held constant, normal goods are more affordable, so people may choose to switch to better quality alternatives.

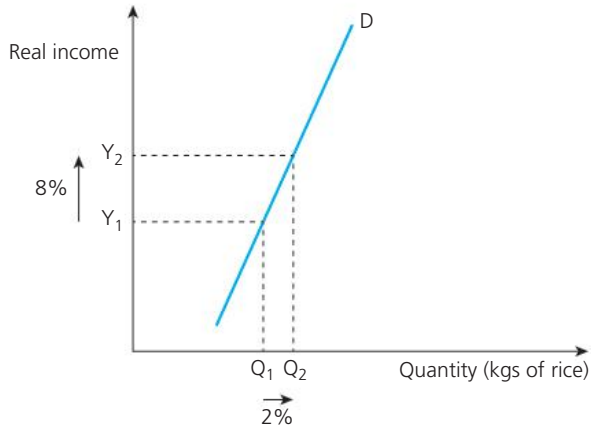


Figure 8.6 Income inelastic demand

Top tip!

It is important to remember that the term 'inferior goods' refers to affordability rather than quality, even though some inferior goods may be of lower quality.

◆ **Inferior goods** are products with a negative income elasticity of demand. In other words, the demand for such products tends to fall when consumers' real incomes rise.

Figure 8.7 illustrates the negative YED value for inferior goods and services. The downward sloping Engel curve (again, note the *y*-axis label is 'real income' not 'price') shows that people will tend to buy fewer inferior products as real incomes rise, ceteris paribus. Instead, customers will choose to buy more superior or luxury products as their real incomes have increased. Hence, in the case of inferior goods, there is an indirect (or negative) relationship between income changes and quantity demanded. Possible examples of inferior goods include low-priced own-label food products in supermarkets, tinned meats, canned vegetables, instant coffee, fast food, frozen ready meals and public transportation (such as bus travel). Consumers may buy the cheaper products when their incomes are lower and make the switch to superior products when their incomes increase.

When demand is income inelastic, a change in income causes a proportionally smaller change in the demand for a product. Figure 8.7 shows that an increase in real income causes a fall in the quantity demanded for inferior goods such as supermarket own-branded products. An increase in income from Y_1 to Y_2 causes a contraction in quantity demanded from Q_1 to Q_2 as consumers switch to more superior or premium branded products.

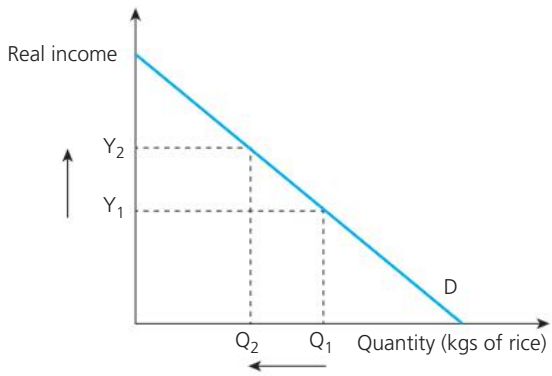


Figure 8.7 The Engel curve for inferior goods

Top tip!

Examples of inferior goods can vary from one country to another. For example, in some countries, a bicycle or public transport might not be regarded as inferior goods, but they might be classified as normal goods in many other places. It is therefore important that you provide and justify your arguments and reasoning.

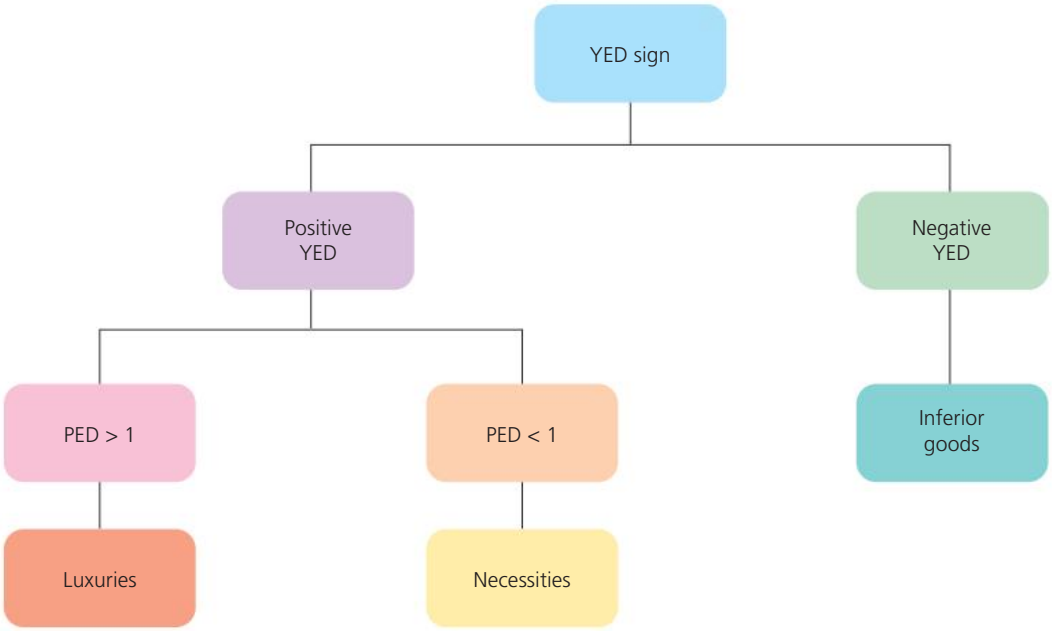


Figure 8.8 The significance of the YED sign

Top tip!

Remember the economic distinction between normal and inferior goods and services – inferior products are associated with a negative YED, whereas normal goods (necessities and luxuries) are associated with a positive YED (see Figure 8.8).

Top tip!

While it is not necessary to show the negative coefficient for PED (owing to the law of demand), it is vital that you use the negative coefficient for YED when referring to an inferior good (less is bought as real income levels rise) and the positive YED value when referring to normal goods.

Top tip!

It can be difficult to classify certain products as inferior or normal goods. For example, economy class air travel may be considered as a luxury good as it tends to account for a large percentage of the average person's income. However, this is relatively cheap (or inferior) compared to business class or first-class air travel (luxury products). The key skills to demonstrate here are reasoning and critical thinking.

Common mistake

'Inferior goods are those that are of poor quality and/or low price.'

It is incorrect to assume that all inferior goods are those that are of poor quality because this is not the meaning of inferior goods. For example, a Honda car is unlikely to be of poor quality but may be considered inferior to more luxurious or superior brands such as Audi, BMW or Mercedes Benz cars.

PAPER 1 EXAM PRACTICE QUESTION 8.1

Using examples, distinguish between *normal goods* and *inferior goods*. *[10 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 8.2

Assume the income elasticity of demand for sausages is –0.25 and that consumer incomes have increased by 3 per cent. Calculate the percentage change in the demand for sausages, assuming all other factors remain constant. *[2 marks]*

Importance of YED (AO3) (HL only)

Knowledge of income elasticity of demand is important for both firms and the government. It also helps to explain changes in the sectoral structure of the economy. Firms and governments can estimate the impact on different markets following changes in real national income (GDP). For example, during a recession, with falling real incomes, inferior goods are likely to face higher demand. Hence, supermarkets and other retailers can sell more inferior products during the recession because there will be higher demand from consumers for these products. Fast food restaurants can also thrive during an economic downturn as people switch demand in favour of inferior food products.

By contrast, firms that supply luxury goods and services are the most affected during an economic downturn as falling real incomes lead to a greater percentage drop in demand. However, the opposite is true during an economic boom, when rising real incomes are likely to lead to falling demand for inferior goods. Hence, firms will increase the supply of the alternative superior products that are income elastic in order to increase sales revenue.

Firms selling income inelastic products tend not to be so affected by fluctuations in the business cycle (see Chapter 16). For example, primary sector products such as oil and agricultural output have a relatively low YED value, as these are necessities for production and consumption respectively. This is likely to be beneficial for producers operating in the primary sector as demand is relatively stable, despite fluctuations in the business cycle (and, hence, variations in real incomes). Other 'recession proof' (income inelastic) goods and services include education, healthcare and law enforcement.

Knowledge of YED is also important for understanding sectoral change as an economy grows. Economic growth (see Chapter 18) is the long-term increase in the economy's national output, which is best measured by the increase in real GDP over a period of time.

Typically, as economic growth occurs and real incomes rise, the primary sector tends to become less important in terms of its contribution towards national output and employment in the economy. As a country's living standards rise, the secondary and tertiary sectors tend to become relatively more important, due to the importance of YED.

Common mistake

Some students confuse YED with PED, stating that *'demand is income elastic if customers buy more when the price falls'*. Although there is an income effect if price falls, YED measures how demand changes following a change in income, rather than changes in price.

Figure 8.9 Fast food retailers can prosper during recessions

PAPER 1 EXAM PRACTICE QUESTION 8.3

Explain **two** reasons why the demand for primary commodities might be relatively price inelastic. *[10 marks]*

For firms operating in the primary sector of the economy, demand is relatively income inelastic. In other words, demand is relatively stable as people do not demand much more following an increase in real income or significantly less when real income falls. For example, customers do not eat considerably more fresh fruits and vegetables following an increase in their real incomes. As a result, the agricultural sector of the economy is likely to grow only slowly as real incomes in the economy increase over time. This is also one reason why countries that rely on the exports

of primary commodities lack the economic growth and development of countries that rely on the export of manufactured goods and services

However, it is important to note that certain primary commodities can be highly income elastic, such as diamonds, oil and gas. Countries that have a natural abundance of oil, such as the USA, Saudi Arabia and Kuwait have prospered from the higher demand for oil as other nations develop their manufacturing and service industries. This shows how reliance on primary commodities can play a vital part in a country's real GDP.

Figure 8.10 The demand for tourism is highly income elastic

In general, secondary sector output (manufactured goods) has a higher YED value. This means demand is more sensitive to changes in income levels than for primary sector output. As standards of living rise, there tends to be a more than proportionate increase in the demand for manufactured goods such as motor vehicles, consumer electronics and furniture. Hence, the secondary sector grows much more rapidly than the primary sector as real incomes increase.

In more economically developed countries (or high-income countries), the tertiary sector grows the most rapidly with rising levels of real income. This is because tertiary sector output (services) has a relatively high YED value. As standards of living rise, the demand for banking, education, and leisure and tourism increase by a larger percentage than the increase in real income. However, this can mean that firms and the economy suffer during an economic recession as customers are highly responsive to the fall in their real income. Hence, YED has an important effect on the allocation of resources within an economy as it determines the nature of sectoral change as the country develops.

Theory of Knowledge

How reliable are statistics and empirical evidence in generating knowledge in economics?

WISE ChoICES

To what extent does higher real disposable income improve **equity** and **economic well-being**?

Top tip!

The concept of elasticity of demand (PED and YED) is vital to the understanding of many aspects of microeconomics, so this is a commonly examined topic (SL and HL). Make sure you learn the definitions (elasticity of demand, price elasticity of demand and income elasticity of demand), the formula for calculating both PED and YED, the factors that affect these, and also why knowledge elasticity is important for households, firms and governments.

Chapter summary

- When the demand for a product is *income elastic*, a change in consumer income levels leads to a greater percentage change in quantity demanded, and vice versa.
- *Normal goods* (necessities and luxury products) have a positive YED value because consumers will buy more of these items as their real incomes increase, ceteris paribus.
- *Necessities* are normal goods with a positive YED value of between 0 and 1, that is, demand is income inelastic as customers continue to buy such essential needs even if prices rise.
- *Luxuries* are normal goods with a positive YED value that is greater than 1, that is, demand is income elastic because customers are relatively responsive to changes in income when buying superior products.
- *Inferior goods* are products with a negative YED value. In other words, as real income levels rise, quantity demanded falls as customers buy alternative superior quality products.
- Knowledge of YED is also important for understanding sectoral change as an economy grows. Typically, economic growth results in a country moving from reliance on primary sector output to secondary and then tertiary sector output.

9 Elasticity of supply

Real-world issue

How do consumers and producers make choices in trying to meet their economic objectives?

Conceptual understandings (WISE ChoICES)

- Consumer and producer **choices** are the outcome of complex decision-making.
- Welfare is maximized if allocative **efficiency** is achieved.
- Constant **change** produces dynamic markets.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:
- price elasticity of supply (PES), including the formula (AO2)
- the degrees of PES, i.e. the theoretical range of values for PES (AO2)
- diagrams that show relatively price elastic and relatively price inelastic supply (AO4)
- diagrams that show constant PES: perfectly price elastic, perfectly price inelastic and unitary PES along a supply curve (AO4)
- the determinants of PES (AO2)
- PES, changes in price or quantity supplied from data provided (AO4)
- the reasons why the PES for primary commodities is generally lower than the PES for manufactured products (AO3) (HL only).

Price elasticity of supply (AO2, AO4)

Price elasticity of supply (PES) measures the degree of responsiveness of quantity supplied of a product (along a given supply curve) due to a change in its price. PES is calculated using the formula:

◆ **Price elasticity of supply** (PES) measures the degree of responsiveness of quantity supplied of a product due to a change in its price.

$$\frac{\text{Percentage change in quantity supplied}}{\text{Percentage change in price}}$$

Or in annotation:

$$\text{PES} = \frac{\%\Delta Qs}{\%\Delta P}$$

Supply is price elastic if firms can quite easily increase the quantity supplied without any time delay if there is an increase in the price of the product. For example, it is fairly easy for Coca-Cola to mass produce more of its soft drinks due to an increase in the price. In such a circumstance, firms that can increase their output at short notice can gain a competitive advantage as they are able to respond to the higher market price.

By contrast, supply is price inelastic if firms find it difficult to change the quantity supplied in a short period of time following a change in the market price. For example, housing has a low PES value, meaning that firms are not highly responsive to changes in price in the

short-term. If house prices fall, for example, firms cannot simply reduce the quantity supplied as it takes time for property developers to put their construction plans on hold. Similarly, if property prices increase, construction companies and property developers cannot easily increase the supply of residential and office buildings, especially for large projects such as commercial skyscrapers and large housing complexes.

Figure 9.1 The supply of housing is price inelastic

The value of PES is positive owing to the law of supply. This means there is a positive correlation between a change in price and quantity supplied. The theoretical range of values for PES (or the degree of PES) is outlined in Table 9.1.

Table 9.1 Degrees of PES values

PES value	Degree of PES	Diagram (PES along a supply curve)
PES > 1	Price elastic supply	Relatively elastic supply
PES < 1	Price inelastic supply	Relatively inelastic supply
PES = 0	Perfectly price inelastic	Constant PES – perfectly inelastic
PES = ∞	Perfectly price elastic	Constant PES – perfectly elastic
PES = 1	**Unitary price elastic supply**	Constant PES – unitary PES

Price elastic supply

If PES > 1 then supply is *price elastic*. In other words, firms are responsive to changes in price (see Figure 9.2). This is because the percentage change in price causes a larger proportionate change in the quantity supplied. An example is mass produced goods, such as soft drinks or bottled water.

Common mistake

'Price elastic supply means that as the price of a good or service goes up, supply rises.'

As the price of a product goes up, supply will rise, ceteris paribus – irrespective of whether supply is price elastic or price inelastic (this is simply the 'law' of supply). With price elastic supply, the *percentage* increase in quantity supplied is greater than the percentage increase in price.

◆ **Unitary price elastic supply** refers to a theoretical situation when the value of PES = 1, that is, the percentage change in quantity supplied is equal to the percentage change in the price of the product.

In Figure 9.2, if price increases by 15 per cent from P_1 to P_2, there is plenty of spare capacity for the firm to raise output in a short period of time. Hence, the quantity supplied increases by a greater proportion of 150 per cent from Q_1 to Q_2. In this case, the PES = 150% ÷ 15% = 10. Firms that have a high PES (elastic supply) are highly competitive as they are very responsive to changes in price so can adjust supply according to changing market conditions.

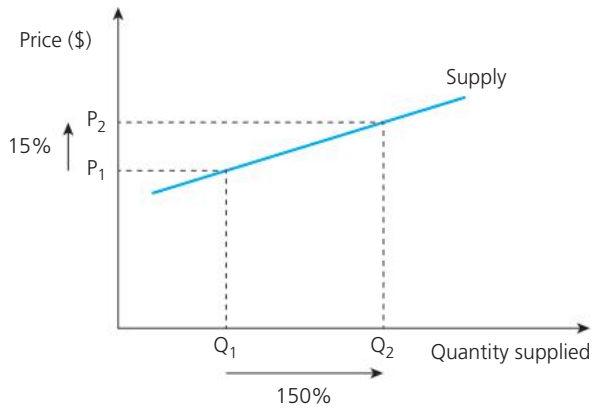


Figure 9.2 The price elastic supply curve

Top tip!

Diagrammatically, at a low enough price, a relatively price elastic supply curve will intersect the *y*-axis.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 9.1

A product has a price elasticity of supply of +3.0. Calculate how much quantity supplied changes if the price falls from $13 to $11.70, while output was originally 5,000 units. *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 9.2

Calculate the value of price elasticity of supply from the diagram below, if the firm chooses to increase its price from $25 to $30. *[2 marks]*

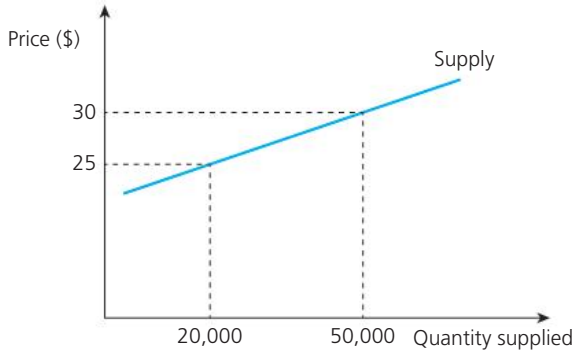


Figure 9.3 The price elastic supply curve

Price inelastic supply

If PES < 1 then supply is *price inelastic*. In other words, firms are relatively unresponsive to changes in price (see Figure 9.4). This occurs when firms find it difficult to react swiftly to changes in price. For example, organic fruits and vegetables take a long time to grow and harvest, so quantity supplied is relatively unresponsive to changes in price.

In Figure 9.4, supply is price inelastic because the percentage change in quantity supplied (10%) is less than the percentage change in the price (40%). When price increases from P_1 to P_2, there is minimal spare capacity so the quantity supplied can only increase by a smaller proportion from Q_1 to Q_2. In this case, the PES = 10% ÷ 40% = 0.25.

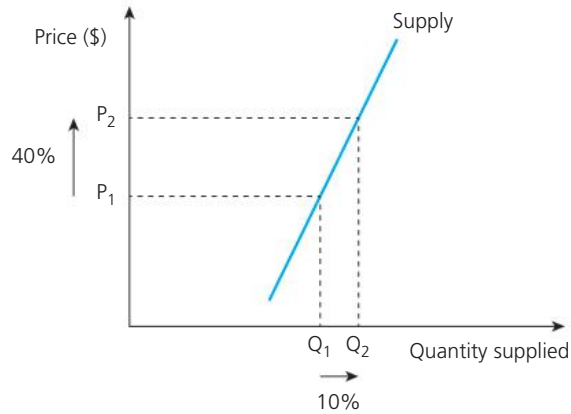


Figure 9.4 The price inelastic supply curve

PAPER 2 AND 3 EXAM PRACTICE QUESTION 9.3

Calculate the value of the price elasticity of supply if the market price of beans increases from $2 per kilogram to $2.20 per kilogram, and causes quantity supplied to rise from 10,000 kilograms to 10,500 kilograms. *[2 marks]*

Top tip!

Diagrammatically, at a low enough price, a relatively price inelastic supply curve will intersect the *x*-axis.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 9.4

The price of luxury watches at Murtagh Corp is reduced from $300 to $250, resulting in quantity supplied falling from 95 units to 85 units per month. Calculate the value of PES for the firm (to 2 decimal places) and comment on your result. *[4 marks]*

Perfectly price inelastic supply

If PES = 0, then supply is *perfectly price inelastic*. In other words, a change in price has no impact on the quantity supplied as there is no spare capacity to raise output (see Figure 9.5). Numerically, the percentage change in quantity supplied is 0, because the supply of a product will remain the same (it is fixed), regardless of any change in price.

In Figure 9.5, supply is perfectly price inelastic at Q_e. Irrespective of any change in the price, firms can only supply a maximum of Q_e, so the change in price from P_1 to P_2 has no impact on the quantity supplied. In other words, PES = 0. An example is a football stadium or a concert hall that cannot accommodate more than the maximum seating capacity.

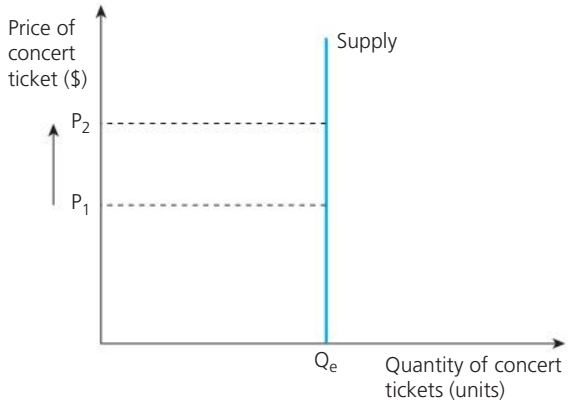


Figure 9.5 The perfectly price inelastic supply curve

CASE STUDY

The value of art

Some real-world examples of perfectly price inelastic supply come from the world of high-end artwork and masterpieces (all figures in US dollars):

- Claude Monet's 1905 *Water Lilies* painting sold for $43.7m at a New York auction in November 2012.
- *L'Homme qui Marche (Walking Man)* by Alberto Giacometti (1961) sold for $104m in January 2010.
- *Garçon à la pipe (Boy with a Pipe)* by Pablo Picasso (1905) was sold for $104m in May 2004.
- *Nude, Green Leaves and Bust* (1932) by Pablo Picasso sold for $106m in April 2010.
- Edvard Munch's *The Scream* (1893) sold for $120m in May 2012.
- In December 2019, Italian artist Maurizio Cattelan's *Comedian*, consisting of only a banana duct-taped to a wall, sold for a staggering $120,000! Cattelan stated he purchased the banana at a local market, for around $0.30.

ATL ACTIVITY 9.1

Discuss with a partner other examples of products that might have (close to) perfectly price inelastic supply. Be prepared to share your answers with the rest of the class.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 9.5

Suggest why an original masterpiece artwork by Picasso or Van Gogh is likely to have perfectly price inelastic supply. *[2 marks]*

Perfectly price elastic supply

If PES = ∞ (infinity) then supply is *perfectly price elastic*. In other words, the quantity supplied can change without any corresponding change in price owing to the spare capacity that exists at the current price level (see Figure 9.6). This theoretical situation means that firms can supply any amount of output at the given market price owing to constant unit costs of production and no limits on its productive capacity.

In Figure 9.6, supply is perfectly price elastic at a price of P_e. For example, Berkshire Hathaway might have a huge stock of Duracell batteries, so any increase in demand will simply result in the company mass producing more Duracell batteries, without the price having to be raised. Numerically, the percentage change in price is zero. Hence, as quantity supplied can be increased from Q_1 to Q_2 without the need for higher prices, the PES = infinity.

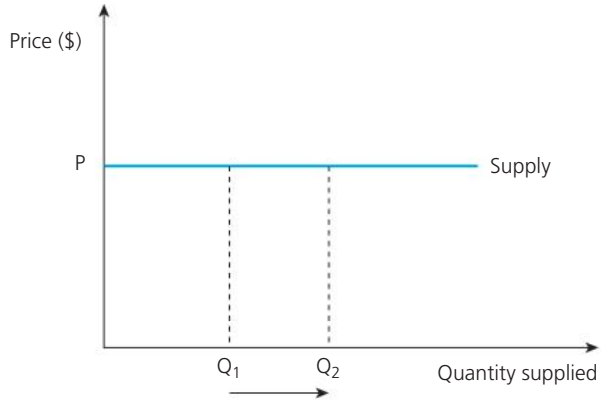


Figure 9.6 The perfectly price elastic supply curve

Unitary price elastic supply

If PES = 1 then supply has *unitary price elasticity*. In other words, the percentage change in the quantity supplied matches the proportional change in price (see Figure 9.7). Numerically, the percentage change in quantity supplied is equal to the percentage change in the price of the product.

In Figure 9.7, any supply curve that starts at the origin (such as S_1, S_2 or S_3) has a PES value equal to 1. This theoretical possibility means that a change in price causes the same proportional change in quantity supplied.

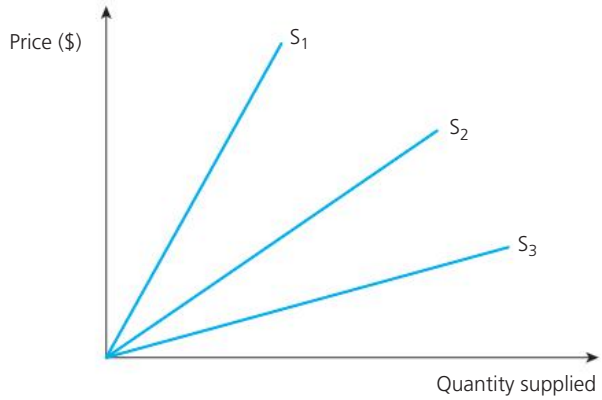


Figure 9.7 The unitary price elasticity supply curve

Top tip!

Mathematically, supply has a different PES value at different points along the linear supply curve. Supply is more elastic at lower prices and more inelastic at higher prices (owing to capacity limits). However, when referring to a supply curve, it is acceptable to describe the supply as relatively elastic (see Figure 9.2) or relatively inelastic (see Figure 9.4).

Determinants of PES (AO2)

There are five main determinants of price elasticity of supply: time, mobility of factors of production (factor substitution), unused capacity, the ability to store (inventory), and the rate at which costs increase (marginal costs).

Top tip!

The determinants of price elasticity of supply can be remembered by the acronym TICCS: **t**ime, **i**nventory (stocks), (marginal) **c**osts, **c**apacity and **s**ubstitution (of factors of production).

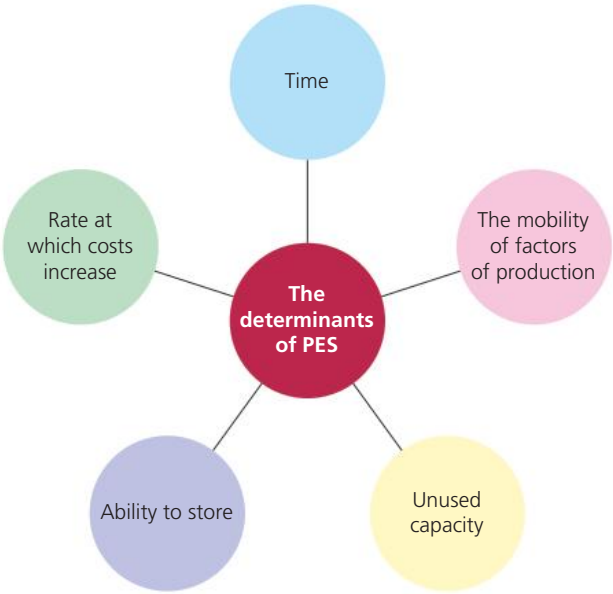


Figure 9.8 The determinants of price elasticity of supply for a firm

The time period

The production process time affects the price elasticity of supply for a product. Supply tends to be price inelastic in the short run as firms do not necessarily have the resources to suddenly increase their output owing to the length of the production process. For example, the supply of organic fruits and vegetables is dependent on the time it takes to grow and harvest these products. Therefore, supply is price inelastic when it takes a long time to produce a good or service, as the quantity supplied cannot be easily increased following an increase in price.

In the same way, the shorter the production time, the more price elastic the supply. For example, the PES of milk is much higher than the PES of cows as it takes farmers significantly less time to milk cows than to breed and rear cattle for meat. In the same way, mass produced products, such as computer chips or canned drinks, are relatively price elastic in supply.

In the long run, firms can adjust their levels of production according to price changes in the market. There is sufficient time for all resources (factors of production) to be increased and for new firms to enter the market. Hence, in the long run, supply tends to be more responsive to changes in price, so the value of PES will be higher.

The mobility of factors of production (factor substitution)

The **mobility of factors of production** refers to the level of ease and cost of factor substitution in the production process. The easier it is to substitute factors of production (such as labour for capital), the more price elastic supply tends to be as producers can adjust supply in response to changes in market prices. In other words, the more mobile factors of production are, the greater the value of PES will be, ceteris paribus. For example, workers on a farm can be easily retrained to grow and harvest carrots instead of broccoli if the price of carrots rises faster relative to the price of broccoli. Similarly, publishers can use machinery to switch between printing textbooks, magazines, comics and novels.

◆ The **mobility of factors of production** refers to the level of ease and cost of factor substitution in the production process. The easier it is to substitute factors of production, the more price elastic supply tends to be.

By contrast, if resources cannot be easily moved to the production of another good or service, then the PES will be inelastic. Factor immobility occurs when it is difficult to switch between factors of production in the production process. This can occur for two reasons:

- Geographical immobility – when it is difficult to move factors of production from one geographical location to another, for example, owing to costs of relocation, the high cost of rent or the lack of available land.
- Occupational immobility – when it is difficult to move factors of production from one type of job or unit of output to another, for example, owing to the lack of skills or experience.

Essentially, if factors of production are relatively easy to substitute in the production process, the price elasticity of supply for the product is higher. This means that if the production of a good or service can be varied, supply is more price elastic.

WISE ChoICES

Examine how factor mobility is important in facilitating economic **change** and **efficiency**.

Unused capacity

Unused capacity refers to the degree of spare productive capacity. A firm with plenty of spare (unused) capacity and access to raw materials and components can increase the quantity supplied with relative ease, without experiencing an increase in its average cost of production. Hence, the quantity supplied can change without any corresponding change in price owing to the unused capacity that exists at the current price level (see Figure 9.6). Therefore, supply is relatively price elastic. For example, the quantity supplied of goods and services is highly price elastic during a recession, because there is plenty of spare capacity and unused labour and capital resources.

In the theoretical case of perfectly price elastic supply (see Figure 9.6), firms can supply any amount of output at the given market price owing to constant unit costs of production and no limits on its productive capacity.

The opposite applies to firms with no spare capacity as they are unable to increase output in response to higher prices. If there are no unused resources (no spare capacity), then a firm is unable to increase output without a significant rise in costs. In other words, the quantity supplied is inelastic in response to a change in price. Essentially, having spare (unused) capacity means that it is easier to increase production if there is an increase in price (as shown by a rightwards shift of the demand curve) of a product.

Top tip!

The supply of most goods and services tends to be relatively price elastic during a recession owing to the spare capacity that exists.

Ability to store

The ability to store refers to the level of stocks or inventory held by a firm. Firms with high levels of inventory (stocks of unused raw materials, work-in-progress and finished goods) tend to have relatively price elastic supply as they are able to respond quickly to a change in market prices. For example, it is relatively easy for firms to store canned meats rather than fresh meat. Hence, the ability to store canned (processed) meats means it has a higher PES value than fresh meats.

By contrast, firms with low levels of stock or the inability to store inventory are unable to respond to meet rising demand or higher prices, so have a relatively low PES. As stocks are low, dwindling supplies force the market price higher owing to the concept of scarcity. A fishmonger is unable to easily store large volumes of fresh fish, so the quantity supplied cannot respond easily to changes in price. Hence, as the fishmonger is unable to store large stocks of fresh fish, supply is price inelastic.

Figure 9.9 Supply of fresh fish is relatively price inelastic

Essentially, firms with high inventories (stocks of raw materials, work-in-progress and finished goods) tend to have relatively price elastic supply.

The rate at which costs increase (marginal costs)

The rate at which costs increase with additional units of output is called **marginal cost** (MC). It is the additional cost associated with producing one extra unit of output. It is calculated using the formula: MC = Change in total costs ÷ Change in output. Note that the MC of production is derived from the variable cost of production, given that fixed costs of production do not change as output changes. In other words, there are no additional fixed costs incurred in producing an extra unit of output once production has already begun.

Owing to the *law of diminishing marginal returns* (see Chapter 4), marginal costs of production will rise after a certain point of output due to the marginal factors of production becoming more expensive to employ. As a result, a firm will not be willing or able to supply at a price below its MC as this would incur a loss for the firm. To do so, the firm needs to receive a price equal to or above its MC of production in order to at least cover the opportunity cost of production.

If a firm can increase output without the marginal costs of production rising at a faster rate, supply is price elastic. In other words, if the rate at which costs rise is low, it will stimulate an increase in the quantity supplied in order to maximize profit. However, if marginal costs rise rapidly, then a much higher market price is needed to stimulate an increase in production, see Figure 14.14 (page 193). The significantly higher cost of production also represents greater risks for businesses, so can stifle production.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 9.6

Angry Birds is a highly popular video game created by Finnish company Rovio, with customers paying $0.99 each to download the game from online app stores. According to the company's website, the game has been downloaded more than 2 billion times across all platforms. With the use of an appropriate diagram, explain why the high level of demand for *Angry Birds* games has no effect on the selling price. *[4 marks]*

Knowledge of PES can enable firms to be more responsive and competitive to changes in the market. There are several ways that firms can achieve this, although these will incur some additional costs:

- create spare capacity
- keep larger volumes of stocks (inventories)
- invest in improved storage systems to prolong the shelf-life of products
- use or upgrade to the latest technologies to increase production capacity and improve capital mobility
- improve distribution systems (how and when the products get to the customers)
- develop and train employees in order to improve labour mobility to perform a range of jobs.

Top tip!

The above analysis outlines the determinants of PES for individual firms in a market. For the industry as a whole, there is another important factor: the number of firms in the industry. In general, the greater the number of firms in the industry, the higher the value of PES as the market is able to respond to higher prices by increasing the quantity supplied.

Theory of Knowledge

To what extent should economic decision-making be based on cultural perspectives of ethical considerations of knowledge claims rather than empirical evidence?

Top tip!

A key determinant of PES not explicitly stated in the IB Economics syllabus is the nature of **barriers to entry** in an industry. The existence of intellectual property rights (such as copyrights, trademarks and patents) reduces the number of potential firms in the industry and therefore lowers the value of PES, ceteris paribus.

Figure 9.10 Barriers to entry are a determinant of PES

PES for primary commodities and manufactured products (AO2) (HL only)

Higher Level students are required to understand the reasons why the PES for primary commodities is generally lower than the PES for manufactured products.

The price elasticity of supply for primary products is relatively low owing to the comparatively long time it takes to increase primary sector output, such as oil, iron, coal and agricultural harvests. By contrast, the PES for manufactured products is generally higher because many of these can be mass produced in shorter time periods; for example, toothpicks, nails and LEGO toys. We can apply the five determinants (TICCS) of price elasticity of supply to explain this:

- Time – It takes a lot longer to grow and harvest wheat or to extract gold than it does to manufacture cereal or bread or jewellery. The output of primary sector products tends to take a relatively long time, so supply is fixed in the short run. Hence, the supply of primary sector output is relatively price inelastic, whereas it is relatively price elastic for manufactured output.
- Inventory (stocks) – Although it is possible to store primary output (such as wheat or coffee beans) these are perishable and can easily become damaged. Hence, it can be difficult to hold stocks of primary products for any significant period of time. Manufactured products (such as laptops or furniture) can generally be stored for longer periods of time. Although these manufactured products may become dated over time (due to newer and improved products being available), they do not perish or become obsolete in the same way as primary products. Hence, the supply of manufactured products tends to be more price elastic in supply.
- (Marginal) costs – The additional costs of supplying an extra unit of mass produced products, such as Krispy Kreme doughnuts, is very low. With mass production, output is continuous with large numbers of identical goods being produced. Therefore, it is relatively easy to produce more without having to increase price as supply is highly elastic. However, in the primary sector, it is not so simple to produce an extra unit of oranges. Hence, this would contribute to the higher price needed to ensure firms have an incentive to produce additional units of primary products.
- Capacity – Capital intensive output in the secondary (manufacturing) sector means it is relatively easy to increase output, whereas production tends to be land-intensive or labour-intensive in the primary sector. Land is fixed in supply – once fields have been planted for the season, the capacity is effectively fixed. Only in the long run can agricultural land be changed significantly. Also, labour-intensive production in the primary sector can be expensive, so this also limits the ability of firms to increase quantity supplied despite any increase in price.
- Substitution (of factors of production) – Capital equipment used in manufacturing is more flexible and adaptable than resources used in the primary sector. Firms can easily adapt capital equipment to increase supply of

Figure 9.11 The marginal cost of producing an extra doughnut is minimal

manufactured products, especially those that are mass produced. For example, machinery can be easily and quickly adjusted to mass produce nails, screws, nuts, bolts and drill bits of various shapes and sizes. In the primary sector, capital tends to be more specialized, such as combine harvesters and tractors. Hence, quantity supplied is more price inelastic as it is more difficult for primary sector firms to increase output especially in the short run.

Chapter summary

- Price elasticity of supply (PES) is a measure of the degree of responsiveness of supply to a change in the price of a product.
- PES is measured using the formula: PES = Percentage change in quantity supplied ÷ Percentage change in price.
- Supply is *price elastic* when PES > 1, that is, the percentage change in the quantity supplied of a product is greater than the percentage change in its price.
- Supply is *price inelastic* when PES < 1, that is, the percentage change in quantity supplied of a product is less than the percentage change in its price.
- Supply is *price unit elastic* (or has *unitary elasticity*) when PES = 1, that is, the percentage change in quantity supplied of a product is equal to the percentage change in its price.
- Supply is *perfectly elastic* when any percentage change in price leads to quantity supplied falling to zero, that is, supply only exists at one price.
- Supply is *perfectly inelastic* when the percentage change in quantity supplied is 0, because the quantity supplied will remain the same (it is fixed) regardless of the change in the price of a product.
- PES is lower if the production time is relatively long, because suppliers cannot easily respond to higher market prices. Supply tends to be more responsive to changes in price in the long run, when all factors of production are variable.
- If a firm has spare (unused) capacity, with access to raw materials and components, production can be easily increased in response to a rise in price, that is, PES is elastic.
- The PES for primary products is generally lower than the PES for manufactured products. This is because the output of primary products tends to take a relatively long time compared to manufactured products, so supply is fixed in the short run. (HL only)

10 Role of government in microeconomics

Real-world issue

When are markets unable to satisfy important economic objectives – and does government intervention help?

Conceptual understandings (WISE ChoICES)

- The market mechanism may result in socially undesirable outcomes that do not achieve **efficiency**, environmental **sustainability** and/or **equity**.
- Resource overuse, resulting in challenges to environmental **sustainability**.
- Government **intervention** is effective, to varying degrees, in different real-world markets.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- the reasons for government intervention in markets (AO2)
- the main forms of government intervention in markets: price controls, indirect taxes and subsidies, direct provision of services, and command and control (regulation and legislation) (AO2)
- diagrams that show the possible effects on markets and stakeholders of the following forms of government intervention in markets: price ceiling, price floor, indirect taxes and subsidies (AO2, AO4)
- the effects on markets and stakeholders of price ceilings, price floors, indirect taxes and subsidies (AO4) (HL only)
- consumer nudges (AO2) (HL only)
- government intervention in markets in terms of consequences for markets and stakeholders (AO3).

Reasons for government intervention in markets (AO2)

This section of the IB Economics syllabus considers situations when the free market (price mechanism) is unable to satisfy important economic objectives, and discusses whether government intervention can help. The government may choose to intervene in economic activity because the free market (price mechanism) may result in socially undesirable outcomes or fail to achieve economic efficiency, environmental sustainability and/or equity.

The main reasons for government intervention in markets (to influence market outcomes) are to: earn government revenue, support firms, support households on low incomes, influence the level of production, influence the level of consumption, correct market failure and promote equity.

To earn government revenue

Governments intervene in markets for several reasons, such as to correct market failures (to provide essential goods and services that are under-provided or not provided by the private sector), to redistribute income and wealth, to influence the level of economic activity in the

country, and to improve overall standards of living. However, in order to do this, the government needs to raise revenue to fund its expenditure. Government revenue can come from several sources, including:

- Tax revenue – This is the main source of government revenue. Taxes can be classified as either *direct taxes* (levies or charges on incomes and wealth) or *indirect taxes* (levies imposed on spending).
- The sale of goods and services – This source of revenue comes from intervention via state-owned enterprises, such as national providers of postal services, the national broadcasting corporation, the airport authority and nationwide railway operators.
- Privatization proceeds – Privatization refers to the sale of state-owned enterprises and government-owned assets, which can be a major source of government revenue. However, this is a short-term policy as state-owned assets can only be sold once to the private sector.
- Sovereign wealth funds (SWF) – These are state-owned investment funds such as stocks and shares, the bonds of other governments, investment in property and gold reserves. These sources of revenue help to generate income to further fund government expenditure.
- Public sector borrowing – Government borrowing is required when its other sources of revenue (from taxes, sale of goods and services from state-owned enterprises, privatization proceeds and sovereign wealth funds) do not meet its spending needs.

Figure 10.1 Reasons for government intervention in markets

To support firms

Government intervention also takes place to provide support for firms. This can take many forms, such as:

- Subsidizing firms, such as agricultural farmers or providers of apprenticeships and training programmes, to lower their production costs and encourage more output.
- Tax concessions – offering tax allowances or discounts reduces the burden on firms, especially new business start-ups and smaller firms struggling to survive or grow.
- Protecting domestic infant industries (newly established firms in markets possibly dominated or threatened by foreign competitors). Such support can be in the form of:
  - tariffs – taxes imposed on imported products from foreign suppliers
  - quotas – quantitative limits on the value or volume of imports
  - subsidies to domestic producers – to lower their costs of production
  - administrative barriers – imposed on foreign products entering the country.
- Business development loans – finance offered to businesses at relatively low interest rates to help fund business operations and growth.
- Research and development funding – finance provided to help firms with their innovation, product development and research and development (R&D) expenditure.
- Financial bailouts – in some cases, the government may choose to intervene by saving failing businesses deemed to be strategically important to the economy. For example, the global financial crisis of 2008 led to a huge number of bankruptcies, causing the US government to bail out Bank of America and Citigroup with $45 billion of funds each.

Top tip!

A useful way to remember the reasons for government intervention is the acronym **FETCH PM** (as in 'fetch the Prime Minister') – intervention is used to influence: **f**irms, **e**quity, **t**ax revenues, **c**onsumption, **h**ouseholds, **p**roduction and **m**arket failure.

To support households on low incomes

Governments are concerned about the huge inequalities and unequal income distribution that can be caused by market forces. While the market economy rewards those who are willing and naturally able to work, have the desired skills and qualifications to succeed, or are gifted with wealth, not everyone in society is so fortunate. Those in poverty lack basic income so are deprived of being able to meet their basic human needs necessary for survival. They may suffer from malnutrition, hunger, a lack of clean water, poor healthcare and inadequate shelter. Hence, government intervention is needed to support low-income households so they can enjoy a better standard of living. For example, government funding for the provision of essential infrastructure (such as sanitation and clean water supplies) helps to promote equity to those on low incomes and to help eradicate absolute poverty.

Figure 10.2 Government intervention is needed to support low-income households

Income redistribution through government tax and expenditure policies can reduce such inequalities, such as using *progressive taxes* and *transfer payments* (see Chapter 19) to low-income individuals and households. Transfer payments are the use of tax revenues to redistribute income and wealth in order to promote income equity; for example, unemployment benefits to those without a job so they can meet their basic human needs, and state pensions to support elderly people in retirement.

To influence the level of production

Government intervention in markets takes place to influence the level of production and consumption in the economy. For example, regulation and taxation policies are used to limit production and consumption of goods that create negative externalities (see Chapter 11). Negative externalities are external costs imposed on stakeholders not involved in an economic transaction. Examples include the production and consumption of tobacco products, alcohol and gambling. By contrast, government subsidies enable firms to provide socially desirable goods and services thereby by making these merit goods (see Chapter 11) available to everyone. Merit goods are those which have positive external benefits to stakeholders not directly involved in an economic transaction. Examples include education and healthcare services.

Governments also intervene in markets to promote general economic fairness, including the fairness of competition in different markets. For example, markets dominated by oligopolistic or monopolistic firms (see Chapter 14) can be anti-competitive, by limiting supply in order to charge customers higher prices and establishing artificial barriers to entry to prevent competition. Oligopolistic and monopolistic markets are those dominated by a few large firms, so competition is limited. Thus, competition policy is used to stimulate competition in the market, thereby influencing the level of production in the industry.

To influence the level of consumption

Government intervention in markets also occurs in order to influence the level of consumption. For example, regulation and indirect taxes are used to reduce the consumption of demerit goods such as alcoholic beverages, sugary drinks, cigarettes, recreational drugs, single-use plastics and fast food (junk food). Another example is regulation and legislation to ensure consumers are provided with adequate information and data to make informed decisions, such as rules of labels used on food products. Information asymmetries (imperfect information) in certain industries, such as insurance, banking and private healthcare, can also cause inefficient or suboptimal decision-making.

Figure 10.3 Governments use 'sugar taxes' to limit the consumption of soft drinks

Government expenditure also enables more socially desirable goods and services to be consumed, such as healthcare services, education, museums, public libraries, recreation centres and apprenticeship programmes. They might do this by directly providing the services or by subsidizing the output or consumption of these services.

To correct market failure

In a socially optimal market economy, all resources are allocated efficiently. However, in reality, there are many instances of market failure (see Chapters 11–14). Governments intervene in markets to address imperfections, inequities and inefficiencies. For example, in an unregulated inefficient market, firms with monopoly power (see Chapter 14) can create barriers to entry and act against the public interest by limiting supply and/or raising prices.

At other times, government intervention takes place as the market fails to provide certain goods and services, as it is impossible to exclude those who are not able or willing to pay. See Chapter 12 to read more about the provision of **public goods**, such as national defence, and law and order. These are goods and services that are **non-excludable** (non-payers can benefit directly) and **non-rivalrous** (one person's consumption of the public good does not reduce the amount available to other parties).

Similarly, there is under-provision of **merit goods** in free market economies, such as inadequate provision of healthcare and education services.

Economic activity often creates external costs (called **negative externalities**) to third parties, such as pollution. These are costs imposed on third parties who are not directly involved in an economic transaction. One of the consequences of economic growth in major cities, such as Bangkok, Dhaka, Paris and Shanghai, is chronic air pollution from traffic-congested roads. Jakarta, Indonesia, is expected to overtake Tokyo as the world's most populous city by 2030 with 35 million people. Without direct government intervention, Jakarta will continue to suffer some of the world's worst traffic jams (drivers expect to spend a couple of hours in traffic to move just three miles or 4.83 kilometers). The traffic problems account for 75 per cent of Jakarta's air pollution. One solution was the government's introduction of Jakarta's MTR (mass transportation system) metro (subway) system in 2019.

Essentially, government intervention is required to prevent or correct such market failures.

◆ **Public goods** are collective consumption goods provided by the government that are non-rivalrous and non-excludable.

◆ **Merit goods** are products that create positive externalities when they are produced or consumed, that is, MSB > MPB.

◆ **Negative externalities** (or **external costs**) are expenses incurred by third parties in an economic transaction for which no compensation is paid.

WISE ChoICES

Based on your learning and understanding of Economics so far, discuss the extent to which there is a trade-off between the concepts of **equity** and **efficiency**.

To promote equity

The free market economy does not always promote equity. Equity is a normative concept concerned with 'justness' or 'fairness' in the allocation of scare resource, although this can be difficult to define so is challenging to measure accurately. For most economists, equity relates to how fairly income is distributed between different groups of people in society. Market economies

face inequitable distribution of income owing to the natural unequal ownership of factors of production. For example, consider the wage differentials between those who earn the national minimum wage and the remuneration for professional footballers, brain surgeons, commercial pilots, Hollywood actors and elite supermodels.

However, equity does not mean equality in income. Free market economists argue that inequities are necessary to create incentives to work, that is, equity is about being fair not equal. So, people with higher levels of qualifications, skills and experience should be paid more. In other words, justified inequalities exist. Nevertheless, governments intervene in markets to promote greater degrees of equity in the distribution of income.

Figure 10.4 How equitable is the pay for elite professional sports athletes?

Nevertheless, the extent to which governments should intervene in markets to promote greater equity remains an area of much debate for economists.

Main forms of government intervention in markets (AO2, AO4)

The government tries to combat market imperfections and inequities for the reasons outlined above, and to improve social welfare. The main forms of government intervention in markets are:

- price controls: price ceilings (maximum prices) and price floors (minimum prices)
- indirect taxes and subsidies
- direct provision of services
- command and control regulation and legislation
- consumer nudges (HL only).

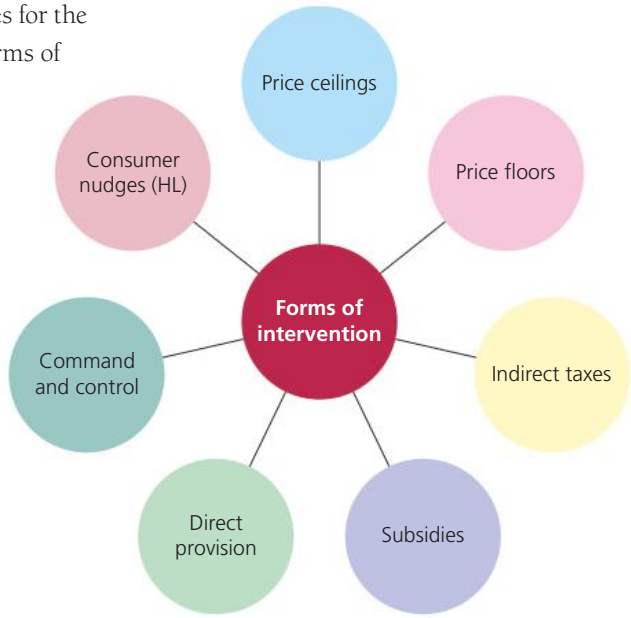


Figure 10.5 The main forms of government intervention in markets

Price controls (AO2, AO4)

Price controls are government regulations establishing a maximum or minimum price to be charged for certain goods and services. They consist of price ceilings (maximum price) and price floors (minimum price).

A **price ceiling** is a type of price control that limits the maximum price in order to encourage output and consumption. Hence, it is also known as a **maximum price**. It occurs when the government sets a price *below* the market equilibrium price. The purpose of a price ceiling is to protect consumers of a certain good or service from soaring prices, such as escalating food prices or rents in the property market. The government wants to ensure the product is affordable for as many consumers as possible by establishing a price that cannot be exceeded.

◆ **Price controls** are government regulations establishing a maximum or minimum price to be charged for certain goods and services.

◆ A **price ceiling** (or **maximum price**) occurs when the government sets a price below the market equilibrium price to encourage output and consumption.

Following the imposition of a binding price ceiling, due to the price being forced below the free market equilibrium, consumer surplus is increased (because many more consumers can obtain a product for a lower price than they would be willing to pay) but producer surplus is reduced. There is also excess demand (see Chapter 5 and Figure 10.6) when the price is below the equilibrium, so can cause inefficiencies in the market.

The binding price ceiling or maximum price, imposed to control residential rents, results in more demand (Q_d) than is supplied (Q_S) at the lower price (P_c). This results in excess demand at all prices below the equilibrium price (P_e), as shown by the shaded area. One option for dealing with this shortage is made up by the government supplying housing at a price of P_c to stabilize residential rents in the economy. Rent controls are regulations that allow only a gradual increase in rent, if at all, compared with what the free market may demand in order to protect tenants. Without rent controls, the high demand for housing in cities could cause rent prices to increase substantially in an unsustainable manner thereby displacing tenants and their families.

Consumers who can access the good or service at the lower price can benefit from the imposition of the price ceiling. However, price ceilings distort market forces and therefore can result in an inefficient allocation of scarce resources, such as a lack of housing. This creates a deadweight loss (see Figure 10.7).

In Figure 10.7, consumer surplus was originally area A + B, and producer surplus was area C + D + E. The imposition of the maximum price of P_2 reduces the quantity supplied to S_1, thereby creating a shortage. This changes consumer surplus to area A + C and reduces producer surplus to area E, so there is a deadweight (welfare) loss as shown by the shaded area. In other words, a drawback of binding price ceilings is a decrease in total community surplus (the sum of consumer and producer surplus).

Owing to the shortage created, maximum prices can also cause underground parallel markets to appear. In Figure 10.7, traders in parallel markets will charge will charge a higher price than P_2 thereby eliminating the excess demand. Such markets are generally illegal.

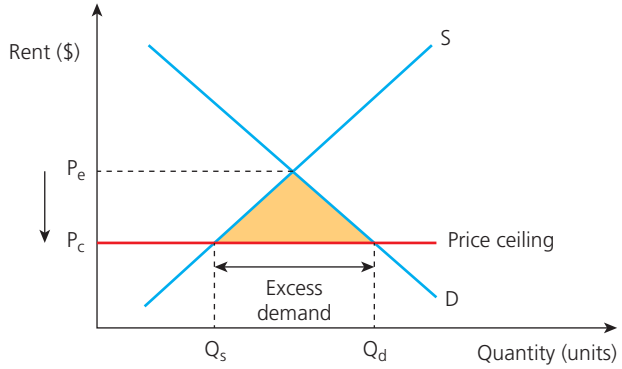


Figure 10.6 Impacts of price ceilings on the residential rental property market

Top tip!

Note that excess demand is shown by the *horizontal distance* between the amount of demand and supply at each price level. The triangular shaded area in Figure 10.6 represents excess demand at all price levels between P_e and P_c.

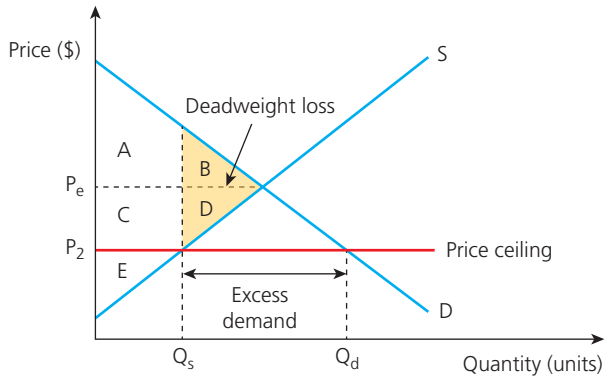


Figure 10.7 Welfare loss of price ceilings

Top tip!

The purpose of using a price ceiling is to protect consumers by making the good or service affordable for as many consumers as possible. Hence, for a price ceiling to be effective, it must be set below the free-market equilibrium price.

ATL ACTIVITY 10.1

During the coronavirus outbreak, the world saw a huge increase in the demand for surgical masks. Read this article about how the government in Laos intervened to make surgical masks, deemed to be merit goods, more affordable for low-income households.

https://www.thestar.com.my/news/regional/2020/02/10/commerce-ministry-of-laos-sets-price-ceiling-for-face-masks

Discuss the extent to which intervention through price ceilings is the most effective way to correct the under-provision of surgical masks.

PAPER 3 EXAM PRACTICE QUESTION 10.1 (HL ONLY)

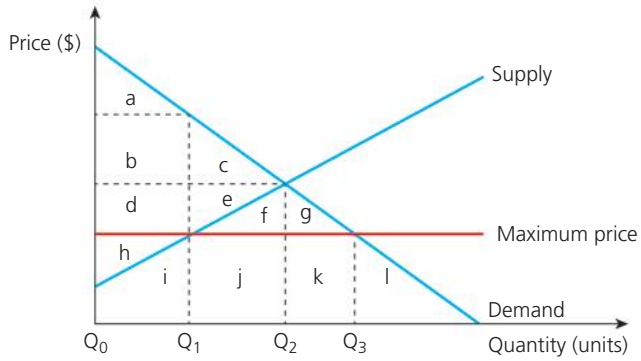


Figure 10.8 The imposition of a maximum price

With reference to Figure 10.8, identify the following:

a The change in consumer surplus following the imposition of the price ceiling. *[1 mark]*

b The change in producer surplus after imposition of the price ceiling. *[1 mark]*

c The change in the firm's sales revenue. *[1 mark]*

d The shortage in the market. *[1 mark]*

PAPER 3 EXAM PRACTICE QUESTION 10.2 (HL ONLY)

The table below shows the demand and supply schedules for product Y.

Demand	Price ($)	Supply
30,000	10	12,000
25,000	15	16,000
20,000	20	20,000
15,000	25	24,000
10,000	30	28,000
5,000	35	32,000

a Define the term *price ceiling*. *[2 marks]*

b Identify the equilibrium price of product Y. *[1 mark]*

c Briefly explain the impact of the government imposing a price floor of $25 for product Y. *[2 marks]*

WISE ChoICES

In highly popular places such as San Francisco, Singapore and Hong Kong, property prices are literally through the roof, with house prices soaring faster than average pay rises and the inflation rate. Growing populations in these areas, along with scarce land supply, have resulted in smaller-sized dwellings, but higher property prices.

Discuss how housing crises are linked to the concepts of **economic well-being**, **intervention**, **equity** and **sustainability**.

Figure 10.9 Apartment buildings in Quarry Bay, Hong Kong

By contrast, a **price floor** is a type of price control that sets a binding minimum price in order to encourage production and supply. Hence, it is also known as a **minimum price**. It occurs when the government sets a price *above* the market equilibrium price. The purpose of a price floor is to protect producers of a certain good or service by limiting how low a price can be charged for a good or service. In the labour market, this is done by imposing a minimum wage that employers must pay their workers.

In Figure 10.10, the guaranteed minimum price offered to agricultural farmers gives an incentive for producers to supply more (S_1) than is demanded (D_1) at the higher price (P_2). This results in excess supply at all prices above P_1, as shown by the shaded area. This surplus is either stored as excess inventory by producers, who bear the costs of storing it, or is more likely to be bought by the government at a price of P_2 to support the farmers. In this case, the surplus is released onto the market during times of bad harvests thereby stabilizing food prices.

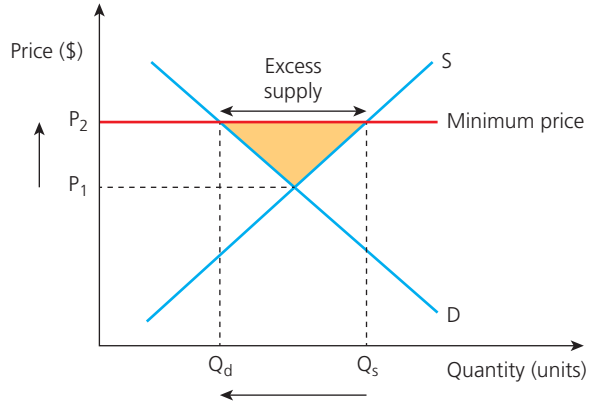


Figure 10.10 Consequences of a price floor

◆ A **price floor**, also known as a **minimum price**, is a form of price control that imposes a price guarantee set above the market equilibrium price to encourage supply of a certain good or service.

Figure 10.11 Minimum prices are used to stimulate increased supply of agricultural output

In the labour market, a national minimum wage (NMW) set above the market equilibrium wage rate, ensures all workers have enough income to provide for their basic needs. Employers can pay more than the minimum amount, but they cannot pay less. In this case, the suppliers are employees (who supply their labour services) and employers are the consumers (who demand labour for their services in the production process). In Figure 10.12, the NMW being higher than the market equilibrium wage rate (where $D_L = S_L$) causes the supply of labour (S_L) to be greater than the demand for labour (D_L). In theory, the NMW results in excess supply of labour. In other words, the higher costs of labour causes unemployment, at least in the short run, ceteris paribus.

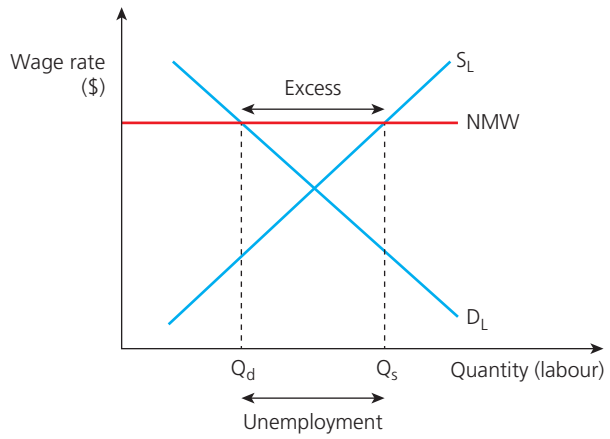


Figure 10.12 Consequences of a national minimum wage

Like price ceilings, minimum prices also distort market forces so can result in an inefficient allocation of scarce resources. In Figure 10.13, the price floor reduces the quantity demanded to D_1 (as consumers need to pay higher prices), creating a surplus (excess supply). This reduces both consumer and producer surplus as shown by the shaded area, so there is a decrease in social surplus. In other words, there is a deadweight loss to society.

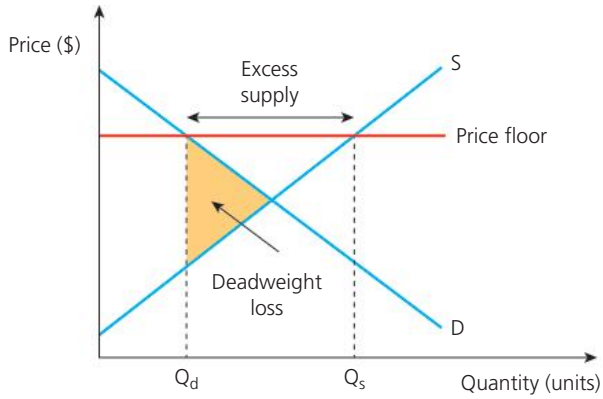


Figure 10.13 Welfare loss of price floors

Top tip!

For a price floor to affect the market, it must be greater than the free-market equilibrium price in order to stimulate greater supply.

PAPER 3 EXAM PRACTICE QUESTION 10.3 (HL ONLY)

Refer to Figure 10.14 and answer the questions that follow.

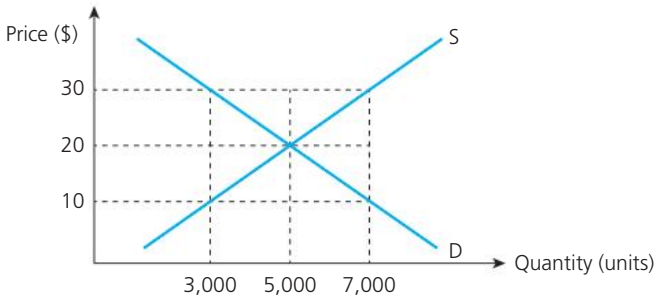


Figure 10.14 The imposition of a minimum price

a Explain what situation arises if the government imposes a price floor of $30 for the product. *[2 marks]*

b Calculate the change in consumer spending following the imposition of the price floor. *[2 marks]*

c Calculate the change in producer revenue following the imposition of the price floor. *[2 marks]*

d Suppose the government exports all the excess supply at $20 per unit. Calculate the amount of taxpayers' money needed to support this price control scheme. *[2 marks]*

Top tip!

For this section of the IB Economics syllabus, both SL and HL students need to be able to construct four essential diagrams to demonstrate the impact of government intervention: price ceilings (maximum prices), price floors (minimum prices), indirect taxes and subsidies. You also need to be able to use these diagrams to explain the possible effects on markets and stakeholders.

Indirect taxes and subsidies (AO2, AO4)

Indirect taxes

An **indirect tax** is a government levy or charge on the sale of goods and services, rather than on incomes or wealth. Indirect taxes are imposed on producers (suppliers), although they often pass on at least some of this to consumers in the form of higher prices. Indirect taxes can be categorized as either:

- **Specific** – These taxes charge a fixed amount of tax per unit sold. Examples include taxes on cigarettes, air passenger tax, landfill taxes, electronic road pricing and highway (motorway) tolls. Specific taxes are indiscriminate as the tax per unit is fixed, so are easier to calculate and administer.

◆ An **indirect tax**, is a payment taken indirectly from the consumer's income, through their expenditure on goods and services.

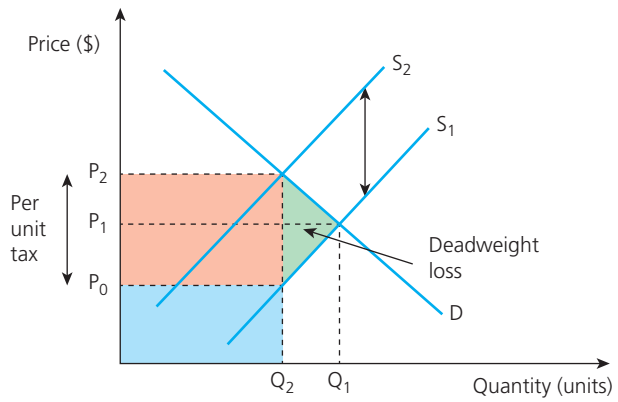
◆ A **specific tax** (also known as a **per unit tax**) imposes a fixed amount of tax on each product that is sold.

- **Ad valorem** – These taxes impose a percentage tax on the value of a good or service. Examples include stamp duty (on the purchase or sale of commercial and residential property), tariffs (taxes on imports) and sales taxes (such as Value Added Tax charged in most European countries, and the Goods & Services Tax used in the USA and other parts of the world). As a percentage tax, ad valorem taxes charge a larger amount of tax based on the value of the purchase. For example, the higher the property value the greater the indirect tax will be.

◆ An **ad valorem** tax imposes a percentage tax on the value of a good or service that is sold.

Governments usually impose indirect taxes on demerit goods (see Chapter 11). For example, many countries use a 'sugar tax' to discourage the excess consumption of products with high sugar content, such as soft drinks. As a result, consumers are more likely to choose sugar-free or low-sugar versions of these drinks (such as switching from Coca-Cola to Diet Coke or Coke Zero). Other demerit goods that are taxed in this way include petrol, fast foods, cigarettes and alcohol.

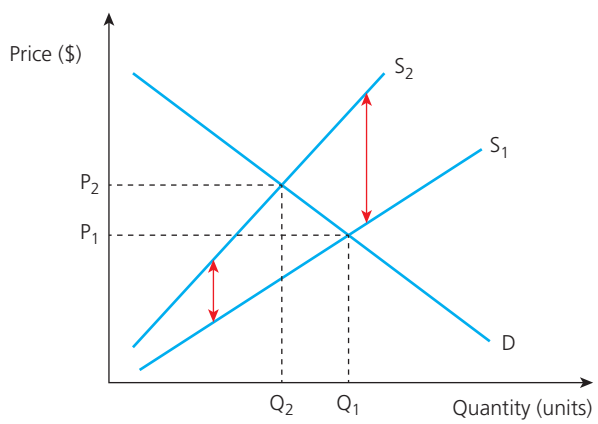
Nevertheless, indirect taxes are a major source of government revenue. The government will often intervene if the price mechanism fails to establish a more socially desirable equilibrium, such as taxing demerit goods to reduce their consumption; for example, imposing fuel taxes on motorists. On a macroeconomic scale, indirect taxes can be used to influence the level of economic activity, by changing the rate of sales taxes. They are also used to protect domestic firms from overseas rivals by imposing tariffs on imports.



■ **Figure 10.15** The imposition of a specific tax (per unit tax)

Indirect taxes raise the costs of production. Specific taxes cause a parallel shift in the supply curve to the left (see Figure 10.15), while ad valorem taxes pivot the supply curve (see Figure 10.16). In either case, the impact is a higher market price and lower quantity demanded.

In Figure 10.15, the specific tax causes a parallel shift in the supply curve from S_1 to S_2, indicating an increase in costs of production. This raises the market equilibrium price from P_1 to P_2 and reduces the quantity demanded from Q_1 to Q_2. The vertical distance between S_1 and S_2 (or P_0 and P_2) is the value of the specific tax. Customers pay P_2 and purchase Q_2 units of output. The shaded red area represents the amount of tax revenue collected by the government, determined by the per unit tax ($P_2 - P_0$) multiplied by the quantity sold (Q_2). The shaded blue area is the sales revenue collected by the firm.



■ **Figure 10.16** The imposition of an ad valorem tax (percentage tax)

The ad valorem (percentage) tax causes a pivotal shift in the supply curve from S_1 to S_2 in Figure 10.16 because the absolute amount of the tax is greater as the price increases (as indicated by the two red arrows). The ad valorem tax raises the price from P_1 to P_2, thereby causing a contraction in the quantity demanded from Q_1 to Q_2.

The imposition of indirect taxes means that consumers tend to lose out as they must pay higher prices, especially if the price elasticity of demand (PED) is relatively low because this suggests there are few, if any, substitute products. Producers also lose out as their costs of production increase. However, if demand is price inelastic then they can pass most of the tax onto customers without largely affecting the level of demand, although the corresponding contraction in quantity demanded will reduce their profitability. Therefore, the imposition of an indirect tax creates a deadweight loss (the combined loss of consumer and producer surplus).

Common mistake

Students often comment that indirect taxes cause a shift in the demand curve because the tax reduces consumers' ability to pay. This is true for direct taxes, whereas indirect taxation is imposed on the producer, so will cause an inwards shift of the supply curve. This, in turn, causes the price to increase – and owing to the price change, there is a contraction in quantity demanded (rather than a shift of the whole demand curve).

The effectiveness of indirect taxes and subsidies depend on the concepts of PED and PES (see Chapters 7 and 9 respectively). A greater proportion of the incidence of tax falls on consumers if the PED is low, that is, there is a lack of substitutes, such as in the case of cigarettes or petrol (gas). This means that firms can pass on most of the higher costs to consumers (see Figure 10.17). The opposite is true when demand is price elastic, that is, producers are unable to pass on most of the indirect tax to the consumer because customers are highly responsive to an increase in price (see Figure 10.18).

In Figure 10.17, the imposition of a specific indirect tax shifts the supply curve from S_1 to S_{+tax}, thus raising the equilibrium price from P_1 to P_2. With the relatively price inelastic demand curve, most of the incidence of the tax is paid by consumers (shown by the area in orange). This is due to the lack of close substitutes available, so consumers are less responsive to the change in price. The rest of the tax is paid by the producer (shown by the area in red).

In Figure 10.18, the relatively price elastic demand curve suggests there are plenty of other close substitutes, so the firm cannot simply pass on the indirect tax to consumers in the form of higher prices. The high PED means consumers will be very responsive to an increase in price. Thus, the producer absorbs the larger proportion of the tax incidence.

If the price elasticity of supply (PES) is high, consumers pay a greater proportion of the incidence of tax. This is because firms are less able and willing to supply if production costs increase, so buyers end up paying a greater burden of the tax (see Figure 10.19). The opposite is true if PES is low (price inelastic).

In Figure 10.19, the relatively price elastic supply curve means firms are highly responsive to changes in price, so must pass on the tax to consumers in order for them to be able to supply. Thus, the consumer pays the larger proportion of the tax incidence.

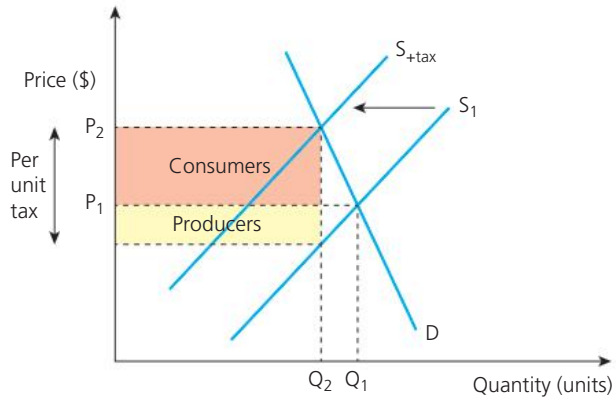


Figure 10.17 Tax incidence and price inelastic demand

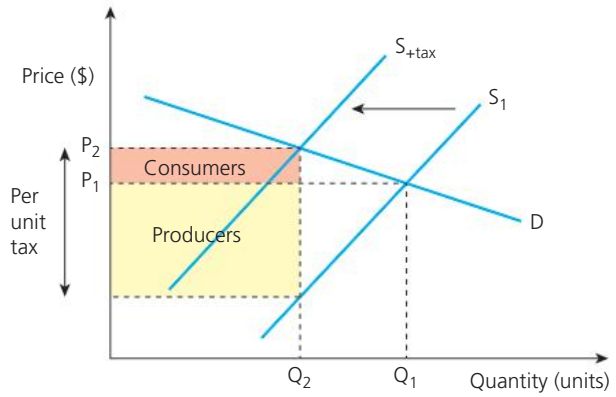


Figure 10.18 Tax incidence and price elastic demand

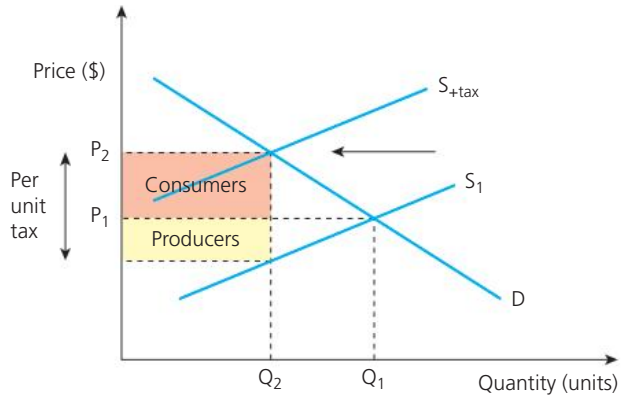
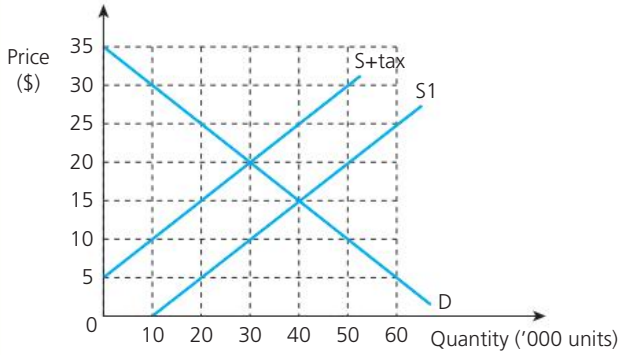


Figure 10.19 Tax incidence and price elastic supply

PAPER 3 EXAM PRACTICE QUESTION 10.4 (HL ONLY)

Refer to the graph (Figure 10.20) and answer the questions that follow.



■ **Figure 10.20** The imposition of a tax

a Calculate the total tax revenue collected by the government from the imposition of the tax. *[2 marks]*

b Calculate the incidence of tax paid by the consumer. *[2 marks]*

c Calculate the change in consumer spending following the imposition of the tax. *[2 marks]*

d Calculate the deadweight loss resulting from the imposition of the tax. *[2 marks]*

e Calculate the value of the producer surplus after the imposition of the tax. *[2 marks]*

f Calculate the change in the value of consumer surplus after the tax has been imposed. *[2 marks]*

Subsidies

A subsidy is a form of financial assistance from the government to encourage output (such as the sale of goods for exporting), to reduce the costs of producing or supplying certain merit goods (such as education, training and healthcare), or to keep down the cost of living for its citizens (such as food prices or the cost of rented accommodation). It is generally given to reduce the financial burden of important stakeholders in targeted markets or industries considered to be in the overall interest of the general public. They are typically used to offset market failures in order to achieve greater economic efficiency.

Examples of subsidies used to encourage greater production and consumption include:

- Offering free vaccinations to children, which is funded by the government. It is therefore more likely that parents will choose to do this for their children given that they do not have to pay.
- Subsidized interest rates on student loans to encourage more people to go on to higher and further education, such as attending university.
- Healthcare subsidies based on household income and size, designed to support lower income households and those with a large number of children by lowering the costs of health insurance premiums.
- Subsidizing the output of merit goods (see Chapter 11), such as education, training and apprenticeship programmes, public transport, family counselling services, public libraries, museums, art galleries and public parks.
- To limit negative externalities such as pollution by subsidizing 'green' technologies, such as biofuels and renewable energy sources such as wind and solar power.

- Strategically protecting certain industries, such as agricultural subsidies to protect farmers and to incentivize production.
- Export subsidies help to protect domestic businesses struggling to compete with international competition, enabling them to lower export prices and/or to compete in a more profitable way. This helps to prevent a subsequent decline in employment.

Subsidies reduce the costs of production for firms, thus shifting the supply curve outwards to the right (see Figure 10.21). The subsidy enables firms to produce more output at any given price, thus shifts the supply curve from S_1 to S_2. The per unit subsidy is shown by the vertical distance between S_1 and S_2, as illustrated by the red arrow. The increase in supply reduces the market price for consumers from P_1 to P_2, resulting in the quantity demanded expanding from Q_1 to Q_2, ceteris paribus. Hence, producers pass on the green section of the subsidy to consumers, in the form of lower prices, and retain the red section for lowering their production costs or higher profit margins.

Consumers generally benefit from subsidies as the market price is lowered, thus more people are able and willing to buy the good or service. In Figure 10.22, consumer surplus increases by the area C + D. Producers also benefit from the subsidy as their production costs are reduced. This helps to improve their competitiveness and profitability. Producer surplus increases by the area A + B. Nevertheless, subsidies distort market forces and may subsequently protect inefficient firms. This results in a deadweight loss, as shown by area E.

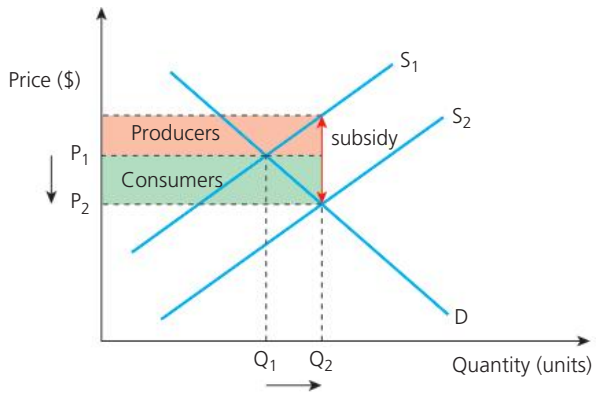


Figure 10.21 Subsidies and market outcomes

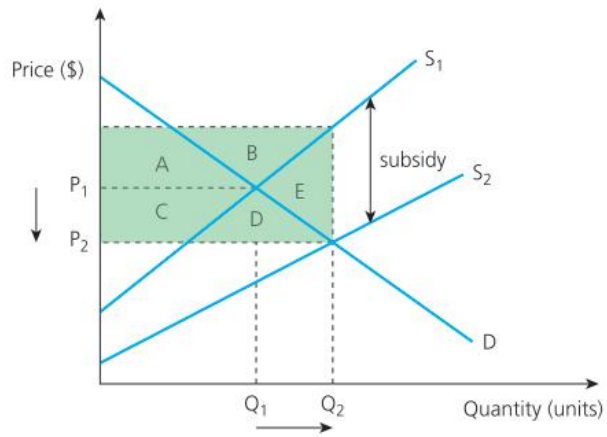


Figure 10.22 The effects of subsidies

Disadvantages of government subsidies

- It is expensive for the government as it would have to raise a significant amount of tax revenue to finance its expenditure on subsidizing the businesses and industries.
- Subsidies can reduce incentives for firms to cut costs or to be more competitive.
- There are opportunity costs in using government funds in this way, such as using the tax revenues for direct provision of public goods (see Chapter 12) or a smaller tax burden on taxpayers.

Nevertheless, while the government spends money on financing subsidies, and there is an opportunity cost in doing so, the net benefits to society in the long-term may outweigh the costs (in theory, at least). For example, if the free market fails to provide a socially optimal output of merit goods, there are sound economic arguments for the use of subsidies. For example, firms that develop environmentally friendly technologies may be able to create huge positive externalities for society in a sustainable way, thus could justify the use of government subsidies.

Top tip!

The incidence of tax falls mainly upon the stakeholder group (consumers or producers) that is less responsive to changes in price, that is, the group that has the most price inelastic function.

ATL ACTIVITY 10.2

There has been much debate about the economic merits of agricultural subsidies in the European Union, such as the EU's Common Agriculture Policy. Investigate the arguments for and against such a policy and evaluate the extent to which you believe agricultural subsidies are in the best interest of the general public.

Theory of Knowledge

How can we know which industries are worth subsidizing and who should make such decisions?

PAPER 1 EXAM PRACTICE QUESTION 10.5 (HL ONLY)

Explain two ways in which a government can intervene in markets to influence the price of a product. *[10 marks]*

PAPER 3 EXAM PRACTICE QUESTION 10.6 (HL ONLY)

The table below shows the demand and supply schedules for Product Z, where Q_d = quantity demanded and Q_s = quantity supplied.

Q_d	Price ($)	Q_s
30,000	7.5	90,000
40,000	7.0	80,000
50,000	6.5	70,000
60,000	6.0	60,000
70,000	5.5	50,000
80,000	5.0	40,000
90,000	4.5	30,000

a Identify the equilibrium price and quantity. *[2 marks]*

b Suppose the government grants a subsidy of $1 per unit to the producers of Product Z. Calculate the new equilibrium price and quantity. *[3 marks]*

c Calculate the total cost to the government of passing the subsidy to producers of Product Z. *[2 marks]*

d Determine the value of the incidence of the subsidy that is passed on to consumers of Product Z. *[2 marks]*

PAPER 3 EXAM PRACTICE QUESTION 10.7 (HL ONLY)

The diagram below (Figure 10.23) shows the situation following a government subsidy being provided for the producers of baby milk powder.

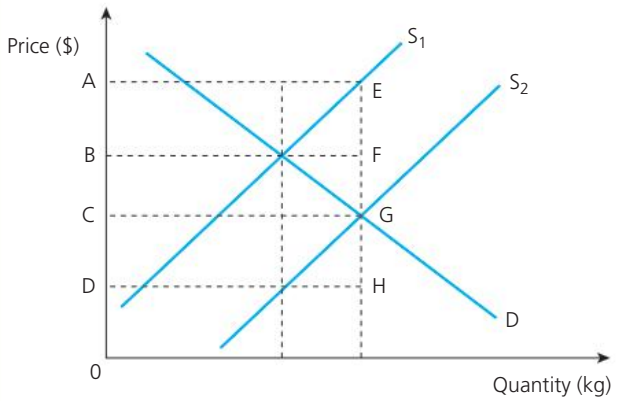


Figure 10.23 The imposition of a subsidy

a Define the term *subsidy*. *[2 marks]*

b Use the diagram above to determine the area that represents the total amount of money spent by the government on subsidizing baby milk powder. *[2 marks]*

c Use the diagram to determine the area that represents the incidence of the subsidy retained by the producers. *[2 marks]*

Direct provision of services (AO2)

The free market is often criticized for under-providing and under-consuming merit goods (see Chapter 11), and public goods might not even be provided at all owing to the free rider problem (see Chapter 12). The free rider problem exists when someone is able to access a good or service without having to directly pay for it. To solve this problem of market failure, the government may decide to directly provide these goods and services. This is known as **direct provision**. Note, direct provision does not mean that the good or service is necessarily provided for free; the government may choose to charge a price for it, although this is likely to be significantly lower than if private profit-maximizing firms were to be involved.

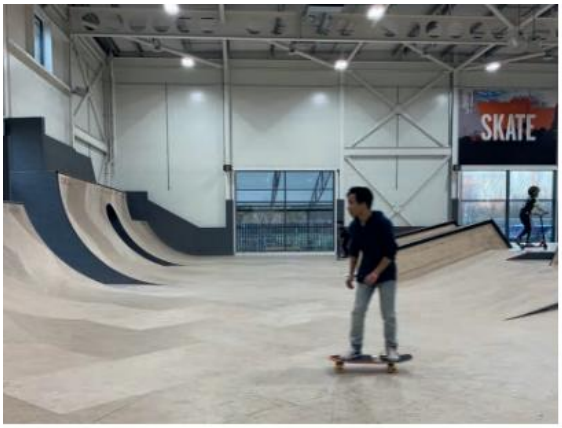


Figure 10.24 Many skateparks and leisure centres are principally funded by the government

Examples of direct provision of services by the government include:

- mental health and intellectual disability services
- public healthcare services, such as government-funded hospitals
- state-funded primary and secondary schools
- postal services to cater for all areas in the country, including remote areas
- sewage systems and sanitation infrastructure
- law and order services, including crime control, police services and the legal system
- national defence and security
- public broadcasting services
- leisure centres, skateparks and other recreational services.

◆ **Direct provision** occurs when the government directly provides or supplies goods and services deemed to be in the best interest of the public.

Public goods are generally provided free of charge by the government, as the market economy fails to create any incentive for private firms to supply these products, such as street lighting and sidewalks (pavements). This is because private firms have no way of preventing free riders (non-payers) from using the public goods. Therefore, the government provides these goods and services, funded through tax revenues.

◆ **Command and control (CAC) regulation and legislation** refers to the direct rules or laws governing an activity or industry, stating what is permitted and what is illegal.

Command and control regulation and legislation (AO2)

Command and control (CAC) regulation and legislation refers to the direct rules or laws governing an activity or industry, stating what is permitted and what is illegal. For example, the Environmental Protection Agency (founded in 1970) oversees all environmental laws in the USA. This includes establishing environmental regulation and legislation that limits pollution to a certain maximum level. In 2020, the UK government banned the use of credit cards to place bets, in an attempt to curb the problem of gambling. A total of 24 million adults in the UK gamble, with 10.5 million of those doing so online. In 2019, the UK opened its first National Health Service (NHS) Gambling Clinic for children owing to growing concern that gambling is being fuelled by online gaming sites and targeted adverts aimed at young people.

Figure 10.25 Smoking – the choice is yours but be warned (nudged)

Other examples of CAC regulation and legislation include:

- Minimum age laws for the purchase of demerit goods such as alcohol, tobacco products, high-sugar and high-caffeine energy drinks, and betting (gambling).

- Environmental protection laws that specify allowable quantities of pollution (legally enforceable numerical limits to pollution) or the pollution-control technologies that must be used in certain industries.
- Outlawing tobacco advertising in the mass media, such as television, radio and newspaper advertisements.
- Banning smoking in public areas such as parks and beaches.
- Laws requiring packaging to carry healthcare warnings about the dangers of smoking.

WORKED EXAMPLE

The table below shows the two firms in the industry that create 700 and 600 units of pollution respectively. If the government wishes to reduce pollution in the industry by 60 per cent, what are the additional costs of the CAC regulation to the industry?

Firm	Current pollution level	Unit cost of cutting pollution
A	700 units	$1,000
B	600 units	$600

Solution

- Firm A needs to reduce pollution by 70% to 210 units at a cost of $1,000 per unit = $210,000
- Firm B also needs to reduce pollution by 70% to 180 units at a cost of $600 per unit = $108,000
- Hence, the total cost to the industry of the CAC regulation = $210,000 + $108,000 = $318,000.

ATL ACTIVITY 10.3

Investigate the use of command and control regulation as a means to combat the global issue of climate change. Evaluate the extent to which climate control measures have proved to be effective in a country of your choice.

However, there are several limitations of CAC regulation and legislation, such as:

- It can be expensive and time consuming for firms to change their processes to meet these standards. As production costs increase, this can reduce their competitiveness and profitability, thereby possibly leading to job losses in some industries.
- They do not offer any incentive for firms to improve the quality of their production beyond the standards set by the law. For example, as soon as car manufacturers have met the stipulated environmental standards there is no incentive to do any better.
- Legislation is inflexible and applies to all firms in the regulated industry. CAC policies do not draw any distinction between large multinational firms that would find it relatively easy and inexpensive to meet these standards and smaller firms that are likely to struggle with the compliance costs of meeting the standards and could even go out of business as a result.

Theory of Knowledge

To what extent might command and control measures contravene human rights and civil liberties? Does the world of politics try to resolve problems that other areas can't resolve?

Consumer nudges (AO2) (HL only)

Nudge theory is the practice of influencing the choices that people make (see Chapter 6). Nudges are created by choice architects, such as the government, using small prompts or tweaks to alter social and economic behaviour, but without taking away the power for people to choose. Nudges are subtle ways of making it easier for consumers to make a certain decision. American economist Richard Thaler, who coined the term 'nudge', wrote *'By knowing how people think, we can make it easier for them to choose what is best for them, their families and society'* (2008). Some examples of consumer nudges as a form of government intervention include:

- Providing information about social norms (what others are doing) in order to influence choice and behaviour. For example, people can be motivated in a certain way by the timely use of personalized messages about:

◆ **Nudge theory** is the practice of influencing the choices that people make. Nudges are created by choice architects using small prompts or tweaks to alter social and economic behaviour, but without taking away the power for people to choose.

- the percentage of people in the local area who pay their taxes on time
- their energy consumption at home compared to that of their neighbours
- socially acceptable behaviour on public transport
- the proportion of the population signed up to donate their organs in the event of an accident leading to death.

- Making use of online technologies to encourage recycling, such as smartphone apps that motivate users to recycle more. This can be providing useful information on the nearest local e-waste and recycling centres and by using a reward system for app users who recycle regularly.
- Using automatic enrolment schemes, which require individuals to make an effort to opt out of these schemes. For example, some governments use an opt-out scheme for pensions, which automatically requires employers to make deductions and contributions to their employees' pension schemes unless workers formally request to be exempted. Similarly, automatic enrolment in organ donation schemes means that if people die in an accident their organs could be used to save someone else's life. Using an opt-in scheme would take more effort to register for these schemes.
- Using highly visible speed cameras and other speed warning systems to prompt motorists to drive slower and to observe speed limits on residential roads.

Figure 10.26 Radar speed systems are used to nudge (influence) consumer behaviour

Just as is the case for IB Economics students, consumers are more likely to change their behaviour if they get immediate and relevant feedback on the benefits of doing what is deemed to be socially desirable.

Theory of Knowledge

Critics of nudges argue that it is akin to manipulation and deception. Can manipulation and deception ever be good for individuals and societies?

Theory of Knowledge

Humans are social beings whose behaviour is often influenced by the behaviour and expectations of others. How can we know whether humans make better decisions based on social norms and expectations or in isolation in order to serve their own self-interest?

Government intervention in markets – consequences for markets and stakeholders (AO3)

Governments intervene in markets when they are inefficient in allocating the economy's scarce resources and in order to deal with market inequities, such as through the use of price controls, regulation and legislation, taxation and subsidies. Government interventions in markets have both advantages and disadvantages on different stakeholders (see Table 10.1). Positive consequences of government intervention in markets include:

- promoting general economic well-being and fairness
- maximizing social welfare
- creating greater economic opportunities.

◆ **Government failure** arises when the cost of attempting to prevent or correct free market imperfections or distortions turn out to be greater than the social costs of the original market failure itself.

Table 10.1 – Government intervention in markets: consequences for markets and stakeholders

Government intervention	Impact on markets		Impact on stakeholders		
	Price	**Output**	**Consumers**	**Firms**	**Government**
Price ceilings (maximum price)	Lower than free market equilibrium	Excess demand (shortage)	Increase in consumer surplus	Decrease in producer surplus	Supply goods and services to match the shortage
Price floors (minimum price)	Higher than free market equilibrium	Excess supply (surplus)	Decrease in consumer surplus	Increase in producer surplus	Buy G&S to eliminate the surplus
Indirect taxation	Higher than free market equilibrium	Decrease in quantity demanded	Decrease in consumer surplus; tax incidence depends on PED	Decrease in producer surplus; lower profitability	Increase in tax revenues, depending on value of PED
Subsidies	Lower than free market equilibrium	Increase in quantity demanded	Increase in consumer surplus; subsidy incidence depends on PED	Increase in producer surplus; higher profitability	Increase in government spending to finance the subsidies

Government intervention is not always successful. **Government failure** arises when the cost of attempting to prevent or correct free market imperfections or distortions turns out to be greater than the social costs of the original market failure itself (see Chapters 11–14 for further details on market failures).

For example, in the labour market, is it not always clear whether government intervention by introducing a National Minimum Wage (NMW) (see Figure 10.12) is beneficial to the economy. Whether the NMW causes unemployment is debatable and the empirical evidence is not conclusive. For example, the higher average wage rates stimulate consumption expenditure in the economy, thereby creating employment opportunities in the long run. The government might also have other priorities in their choice of using a NMW, such as correcting market failures in the labour market including the exploitation of low-skilled labour, to combat poverty in the economy or simply to create incentives to work.

The effectiveness of government intervention is dependent on several factors, including:

- Value judgements – As a social science, economics is subjected to different perspectives. Many economists and policymakers want a particular form of intervention (such as agricultural subsidies) as they have their own vested interest.
- The ceteris paribus assumption – The analysis in this section is subject to the changing nature of social behaviour of individuals and societies. Hence, economic models and theories can only go so far in predicting the behaviour of economic agents. Furthermore, policies used in isolation (such as only taxing cigarettes) are rarely successful.

WISE ChoICES

Economists and policymakers often debate the extent to which governments should intervene in markets, such as the use of subsidies. To what extent does **intervention** compromise **efficiency** in the allocation of scarce resources in the economy?

- The value of price elasticity of demand (PED) and price elasticity of supply (PES) for different products, such as demerit goods.
- The law of unintended consequences – Government intervention does not always work in the way planned or the way the economic theory predicts. Quite often, intervention creates problems of its own, adversely affecting some stakeholder groups for the benefit of others.

Hence, when tackling microeconomic problems, such as the negative externalities caused by smoking or gambling, there is a need to consider alternative ways of dealing with the issue other than through government intervention.

Chapter summary

- Government intervention occurs because the free market (price mechanism) may result in socially desirable outcomes or fail to achieve economic efficiency, environmental sustainability and/or equity.
- The reasons for government intervention are to influence: **f**irms, **e**quity, **t**ax revenues, **c**onsumption, **h**ouseholds, **p**roduction and **m**arket failure (the FETCH PM acronym).
- Price controls are a form of government intervention involving regulation to establish a price ceiling (maximum price) or price floor (minimum price) for certain products.
- Price ceilings cause excess demand (shortages) whereas price floors create excess supply (surpluses). In both cases, society experiences a welfare loss (or deadweight loss).
- Indirect taxes are government levies on the sale of certain products, with the tax being specific (a fixed amount) or ad valorem (percentage).
- The effectiveness of indirect taxes and subsidies depends on the concepts of PED and PES. A greater proportion of the incidence of tax falls on consumers if the PED is low. If the PES is high, consumers pay a greater proportion of the incidence of tax.
- Subsidies are a form of financial assistance from the government, used mainly to encourage the production and consumption of targeted products.
- Direct provision is a form of government intervention that occurs when the government directly provides or supplies products that are deemed to be in the best interest of the general public.
- As a form of government intervention, nudges are created by choice architects (such as the government) using small prompts or tweaks to alter social and economic behaviour, but without taking away the power for people to choose. (HL only)
- Government failure arises when the costs of trying to correct market failures (imperfections or distortions) are greater than the social costs of the original market failure itself.

11 Market failure – externalities and common pool (common access) resources

Real-world issue

When are markets unable to satisfy important economic objectives – and does government intervention help?

Conceptual understandings (WISE ChoICES)

- The market mechanism may result in socially undesirable outcomes that do not achieve **efficiency**, environmental **sustainability** and/or **equity**.
- Resource overuse, resulting in challenges to environmental **sustainability**.
- Government **intervention** is effective, to varying degrees, in different real-world markets.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- social optimum output occurs where marginal social benefit equals marginal social cost (AO2)
- that allocative efficiency and social (community) surplus are maximized when MSB = MSC (AO2)
- a diagram that shows allocative efficiency (AO4)
- positive externalities of production and consumption and welfare loss in the context of merit goods (AO2)
- negative externalities of production and consumption and welfare loss in the context of demerit goods (AO2)
- diagrams that show market failure due to: negative externalities of production and consumption, and positive externalities of production and consumption (AO4)
- welfare loss from a market failure diagram (AO4) (HL only)
- common pool resources, including the characteristics (with reference to tragedy of commons, and CPRs being rivalrous but non-excludable) (AO2)
- unsustainable production of common pool resources, thus creating negative externalities (AO2)
- various methods of government intervention in response to externalities and common pool resources (AO2)
- diagrams that show government responses to externalities: indirect (Pigouvian) taxes, carbon taxes, subsidies, legislation and regulation, and education (AO4)
- the strengths and limitations of government policies to correct externalities and approaches to managing common pool resources (AO3)
- the importance of international co-operation to correct externalities and managing common pool resources (AO3).

Socially optimum output (AO2, AO4)

The price mechanism (see Chapter 5) is not always efficient in allocating scarce resources in the economy. **Market failure** occurs when the signalling, incentive and rationing functions of the price mechanism fail to operate optimally, which leads to a loss in economic welfare. From society's point of view, there is a misallocation of resources in the economy. For example, the market may fail to consider the external costs (such as congestion and pollution) and external benefits (from the provision of public and merit goods) arising from production and consumption. Essentially, market failure or inefficient resource allocation results in one of any combination of four outcomes:

◆ **Market failure** refers to any situation when the price mechanism (the free market forces of demand and supply) allocates scarce resources in an inefficient way.

- under-provision of certain goods and services
- over-provision of certain goods and services
- under-consumption of certain goods and services
- over-consumption of certain goods and services.

Examples of market failure include:

- The under-provision of merit goods, such as education and healthcare, as these services would be provided in a market economy only to those who were willing and able to pay.
- The over-provision of demerit goods such as tobacco, alcohol and gambling due to the absence of government intervention in such markets.
- The under-provision of public goods (see Chapter 12), such as street lighting and public roads, because producers cannot exclude those who do not pay benefiting from the provision of these services.
- The abuse of monopoly power (see Chapter 14) by charging customers prices above market equilibrium, and, hence, the inefficiencies associated with a suboptimal allocation of scare resources.

Market failure occurs where there is a divergence between private costs and benefits and social costs and benefits of production and/or consumption.

- External benefits (or positive externalities) are the advantages or gains of production or consumption to a third party, i.e. individual and/or firms not directly involved in an economic transaction.
- External costs (or negative externalities) are the disadvantages or losses incurred from production or consumption by a third party not directly involved in an economic transaction.
- **Private benefits** are the advantages or gains of production and consumption enjoyed by an individual firm or person.
- **Private costs** of production and consumption are the actual expenses incurred by an individual firm or person.
- **Social benefits** are the true (or full) benefits of consumption or production, that is, the sum of private benefits and external benefits (also referred to as positive externalities). Hence, social benefits = private benefits + positive externalities.
- **Social costs** are the true (or full) costs of consumption or production, that is, the sum of private costs and external costs. Hence, social costs = private costs + negative externalities.

◆ **Private benefits** are the gains of production and consumption enjoyed by an individual firm or person.

◆ **Private costs** of production and consumption are the expenses incurred by an individual firm or person.

◆ **Social benefits** are the total benefits of consumption or production to society, that is, the sum of private benefits and external benefits.

◆ **Social costs** are the total costs of consumption or production to society, that is, the sum of private costs and external costs.

For example, the private costs of driving include the purchase of a vehicle, petrol (fuel), insurance, parking fees and maintenance costs. However, there are also external costs (negative externalities) imposed on others in society, including pollution, traffic congestion and possible road accidents. Social costs of driving include the total of all private and external costs.

Figure 11.1 Driving incurs both private costs and external costs

PAPER 2 AND 3 EXAM PRACTICE QUESTION 11.1

Outline **two** social costs of building a new airport. *[4 marks]*

The socially optimum output occurs when the marginal social benefit (MSB) equals the marginal social cost (MSC). In other words, at the output level where MSB = MSC.

- **Marginal private benefit** (MPB) is the additional value enjoyed by households and firms from the consumption or production (output) of an extra unit of a particular good or service.
- **Marginal private cost** (MPC) is the additional expense of production for firms or the extra charge paid by customers for the output or consumption of an extra unit of a good or service.
- **Marginal social benefit** (MSB) refers to the total gains to society from an extra unit of production or consumption of a particular good or service. It is the sum of the benefits for private individuals and the positive externalities to others in society.
- **Marginal social cost** (MSC) is the total expenses to society from an extra unit of production or consumption of a particular product. The total costs of producing or consuming an additional unit of output includes both private or direct costs of producers and consumers plus the costs to others in society.

Top tip!

MSB > MPB if there are positive externalities of production or consumption (such as education and healthcare). MSB < MPB if there are negative externalities of production or consumption (such as smoking and gambling).

Allocative efficiency is achieved when social or community surplus is maximized (see Chapter 5). This occurs at the socially optimum level of output, that is, where MSB equals MSC. This is because at this level of output, it is not possible to reallocate resources to make one party better off without making others worse off.

Social surplus (also known as **community surplus**) is the sum of consumer and producer surplus at a given market price. It is the net benefit available to society from an economic transaction or activity. In Figure 11.2, social surplus is shown by the area A + B, which is made up of both consumer surplus (area A) and producer surplus (area B). Community surplus is maximized when the price mechanism clears the market of any excess demand or supply. At the market equilibrium price (P_e) and equilibrium quantity (Q_e), there are no shortages or surpluses. In other words, resources are allocated efficiently (at point X) thereby maximizing social welfare.

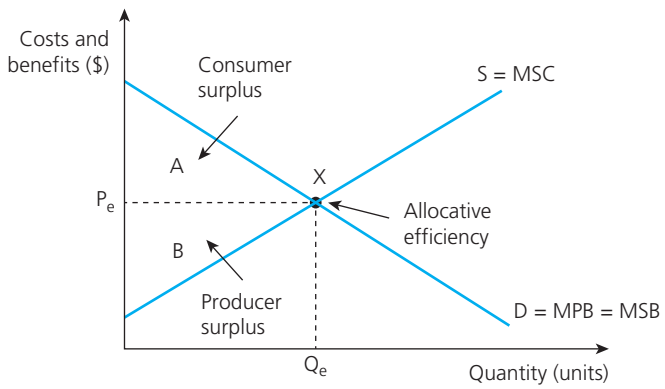


Figure 11.2 Allocative efficiency: maximizing social surplus

If the economy was operating at a level below the socially optimal (Q_e), then both consumer and producer surplus could be increased by producing and consuming a greater amount, thereby improving social surplus and economic welfare. Allocative efficiency occurs at point X in Figure 11.2 because it represents the best possible use of an economy's scarce resources to produce what society deems to be the optimal amount. This means all factors of production are used efficiently and there are no unemployed resources.

Top tip!

Community (social) surplus and allocative efficiency are important economic ideas to consider when discussing market failures and the role of government intervention in markets.

Top tip!

When price is set equal to marginal cost (P = MC), economic welfare is maximized, that is, consumer and producer surplus are maximized. This is because the price that consumers are willing and able to pay matches that of the producers.

Theory of Knowledge

How can we know when an economic problem truly justifies the need for government intervention?

Top tip!

Higher Level students need to be able to calculate consumer surplus and producer surplus from a diagram. Have a go at some of the practice questions in Chapter 5 to test your understanding of consumer and producer surplus.

Positive externalities of production and consumption and welfare loss (AO2, AO4)

This section of the IB Economics syllabus covers positive externalities (also referred to as external benefits) of production and consumption, welfare loss and merit goods. **Externalities** (also known as **spillover effects**) are the external costs or benefits of an economic transaction, causing the market to fail to achieve the social optimum level of production or consumption, where marginal social benefits equal marginal social costs (MSB = MSC) as shown in Figure 11.2 above.

◆ **Externalities** (or **spillover effects**) are the external costs or benefits of an economic transaction, causing the market to fail to achieve the social optimum level of production or consumption.

Positive externalities (also known as **external benefits**) of production and consumption refer to the benefits enjoyed by a third party from an economic transaction. The production and consumption of public goods (see Chapter 12) and merit goods exert positive externalities. Examples of products with external benefits for which no money is paid directly by the beneficiary include:

- flood defence systems
- law and order system, including police and emergency services
- lighthouses
- national defence and security
- public firework displays
- public parks, libraries and museums
- sewage and waste disposal systems
- street lighting.

Figure 11.3 Sewage systems create positive externalities

Common mistake

When students are asked to explain why externalities are a source of market failure, they often focus only on negative externalities, such as pollution and damage to the environment. However, remember that externalities are spillover effects of economic activity, so can be positive (external benefits) as well as negative (external costs).

◆ **Positive externalities** (or **external benefits**) are the benefits enjoyed by a third party not directly involved in an economic transaction.

Merit goods

Merit goods are products that create positive externalities when they are produced or consumed. Hence, the marginal social benefits from the production and consumption of merit goods is greater than the marginal private benefits. In other words, MSB > MPB owing to the external benefits of production and consumption. They are deemed to be of value by society owing to the spillover effects to third parties, so can have a significant impact on social welfare. In other words, the provision and consumption of merit goods improves the general standard of living in the economy.

◆ **Merit goods** are products that create positive externalities when they are produced or consumed. Hence, the MSB of production and consumption of merit goods exceeds the MPB.

Examples of merit goods include:

- community services, such as family counselling and recycling programmes
- education (pre-school, primary, secondary and tertiary education)
- flu vaccinations

- healthcare services
- museums, art galleries and public libraries
- public housing
- public transportation
- research and development
- sports and recreational services (such as public swimming pools and sports grounds)
- training and development.

However, merit goods are rivalrous and excludable. **Rivalrous** means that consumption of a merit good (such as taking up vacant hospital bed or a place at a school) reduces the amount available to others at that point in time. **Excludable** means that it can be possible for suppliers to prevent non-payers benefiting from the merit good (such as private healthcare clinics and fee-paying schools). Since they are both rivalrous and excludable, merit goods tend to be under-consumed and under-produced in free market conditions (see Figures 11.5 and 11.6 respectively). In addition, individuals may not act in their own best interest owing to imperfect information. In other words, people do not necessarily know what is good for them! This means that government intervention is often needed to ensure a more optimal allocation of scarce resources.

Figure 11.4 The provision of public housing is an example of merit goods

In Figure 11.5, positive consumption externalities exist because the marginal social benefit of consumption (MSB curve) is greater than the marginal private benefit (MPB curve) at all levels of output up to the socially optimal level of output (Q_{opt}). This is due to the existence of positive externalities of consumption. Hence, there is market failure at the free market equilibrium (Q_e and P_e), that is, there is under-consumption of the merit good. The socially desirable level of output is where MSB = MSC. In other words, at output level Q_{opt}. The green shaded area represents the welfare loss due to the under-consumption.

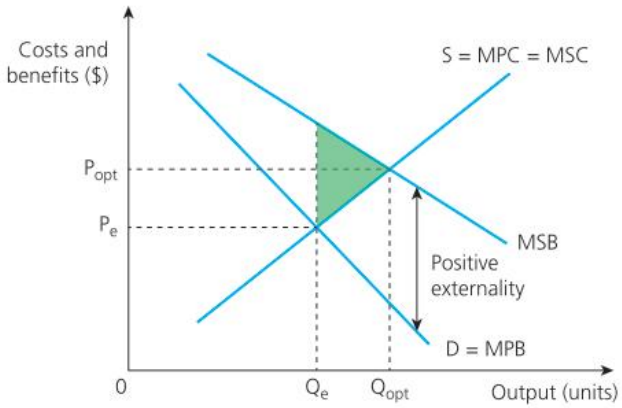


Figure 11.5 Positive externalities of consumption (of merit goods)

Top tip!

Note that the welfare loss in market failure diagrams also represents the possible welfare gain if production and consumption are increased in the case of merit goods, and if the production and consumption are reduced in the case of demerit goods.

In the case of a positive externality of production (see Figure 11.6), the marginal private cost (MPC curve) is greater than the marginal social cost (MSC curve). This is true for all levels of production up to the socially optimal level of output (Q_{opt}), when the MSC = MPB, when the existence of positive externalities (as shown by the vertical distance between MPC and MSC) is fully exploited.

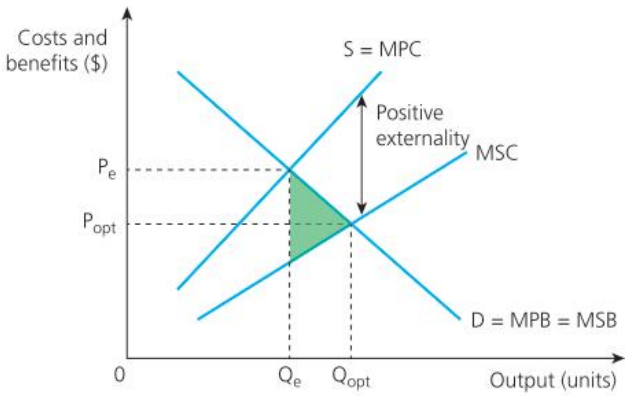


Figure 11.6 Positive externalities of production (of merit goods)

Hence, there is market failure at the free market equilibrium (Q_e and P_e), that is, there is under-consumption of the merit good. The green shaded area represents the welfare loss as a result of the under-consumption. Government intervention is needed to lower the price from P_e to P_{opt} thereby encouraging an increase in the consumption of the merit good up to Q_{opt} units of output.

Common mistake

Do not confuse the definition of public goods with that for merit goods. Both the public and private sectors of the economy provide merit goods, such as education and health services. Unlike public goods, merit goods can be *rivalrous* and *excludable*, for example, selective schools, single-gender schools, religious-based schools or private healthcare clinics.

Top tip!

Higher Level students need to be able to calculate the value of welfare loss from a market failure diagram. This applies to both positive and negative externalities (see Practice question 11.2 later in this chapter).

Theory of Knowledge

Some people argue that contraception is a merit good as it can prevent the personal and social costs of unwanted pregnancies. However, others argue that contraception is sinful as it encourages sexual behaviour and undermines family values. How can economists know whether contraception is a merit or demerit good?

ATL ACTIVITY 11.1

Completing the IB Diploma Programme is short-term pain for long-term gain! With a partner, discuss why education (as a merit good) creates positive externalities of consumption and decide whether you agree with the above statement.

WISE ChoICES

Discuss whether the provision of merit goods always improves **economic well-being** and **equity**.

Negative externalities of production and consumption and welfare loss (AO2, AO4)

◆ **Negative externalities** (or **external costs**) are expenses incurred by third parties in an economic transaction for which no compensation is paid.

This section of the IB Economics syllabus covers negative externalities of production and consumption, welfare loss and demerit goods. **Negative externalities** (also known as **external costs**) are expenses incurred by third parties in an economic transaction for which no compensation is paid. For example, the use of fossil fuels (those formed by natural processes, such as coal, crude oil and natural gas) in the production process can pose a threat to economic, social and environmental sustainability. Fossil fuels can take millions of years to generate, so economic activity depletes these scarce and non-renewable resources. The production and consumption of demerit goods also exert negative externalities such as pollution. Examples of external costs include:

- air pollution caused by fumes from factories or traffic congestion
- child obesity from consumption of junk food and carbonated soft drinks
- climate change fuelled by continual demand for higher levels of consumption and production

Figure 11.7 Advertising clutter and light pollution in Manhattan, New York City

- excessive advertising (known as *advertising clutter*) that causes visual blight
- light pollution
- litter on public beaches
- noise pollution from a night club or from a nearby airport
- second-hand (passive) smoking.

Demerit goods

Demerit goods are products that create negative spillover effects (or negative externalities of production or consumption) to third parties not directly involved in an economic transaction. The production and/or consumption of demerit goods can be damaging to third parties in terms of physical, mental or moral harm. Hence, their production and consumption results in marginal social costs being greater than marginal private costs of production and consumption, that is, MSC > MPC. In a free market, without government intervention, demerit goods are over-produced and over-consumed from society's point of view. Examples of demerit goods include the over-provision and over-consumption of:

- alcoholic beverages
- gambling and betting services
- junk food
- pornography and prostitution
- recreational drugs and hard drugs
- single-use plastics
- tobacco products, such as cigarettes and cigars.

◆ **Demerit goods** are products that create negative spillover effects (or negative externalities) to others in society. Hence, their production and consumption results in MSC > MPC, causing a welfare loss to society.

Top tip!

Note that what might be deemed as a demerit good is sometimes dependent on context, culture and values. For example, there are different perceptions on the use of pornography, prostitution or recreational drugs. After all, economics is not entirely free from value judgements. What is important is that you are able to explain and justify your arguments in the examinations.

Figure 11.8 Junk foods are demerit goods

Although it is assumed that people are rational economic agents, it is often the case that they are not the best judges of their own welfare. Hence, the government intervenes to discourage the production and consumption of demerit goods (see Figure 11.9 and Figure 11.10 respectively), such as placing age limits for drinking, smoking, gambling and driving.

In Figure 11.9, the negative externalities of production occur when producing a good or service causes a harmful effect to a third party, such as burning fossil fuels, which creates environmental pollution. Hence, the marginal social cost (MSC curve) is greater than the marginal private cost (MPC curve). In a free market, without government intervention, the level

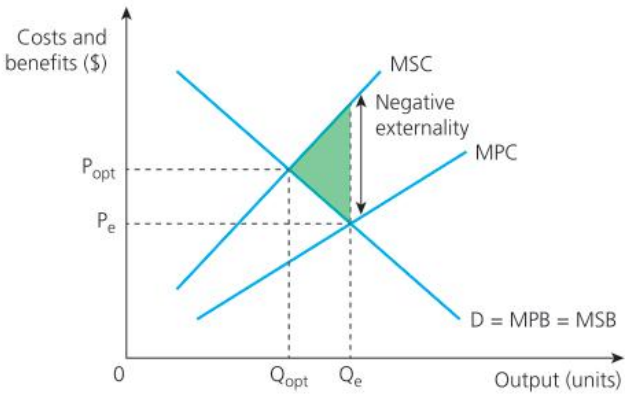


Figure 11.9 Negative externalities of production (of demerit goods)

of output will be at Q_e where the MPC = MPB of production. However, the socially optimal level of output is at Q_{opt} where the MSC = MSB, with a higher price of P_{opt} being charged. Hence, from society's point of view, there is over-production of demerit goods and a corresponding welfare loss as shown by the green shaded area. Essentially, society would benefit from reducing output of the demerit good from Q_e to Q_{opt}, as this level of output eliminates the negative externalities (as shown by the vertical distance between MSC and MSB at all output levels beyond Q_{opt}).

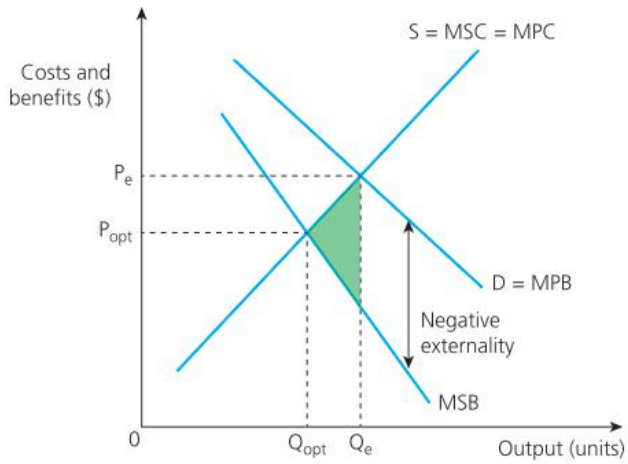


Figure 11.10 Negative externalities of consumption (of demerit goods)

In Figure 11.10, the negative externalities of consumption occur when consuming a good or service causes a harmful effect to a third party, such as smoking, which causes passive (second-hand) smoking to others in the surrounding area. In this case, the marginal social benefit (MSB curve) is less than the marginal private benefit (MPB curve). In a free market, without government intervention, equilibrium output is at Q_e, which exceeds the socially optimal level at Q_{opt}, where the MSC = MSB of consumption. Hence, from society's point of view, there is over-consumption of demerit goods (such as tobacco or alcohol) and a corresponding welfare loss as shown by the green shaded area. Essentially, society would benefit from reducing consumption of the demerit good from Q_e to Q_{opt}, as this level of output eliminates the negative externalities (as shown by the vertical distance between MSC and MSB at all consumption levels beyond Q_{opt}).

PAPER 3 EXAM PRACTICE QUESTION 11.2

a Describe an example of a product with negative externalities. *[2 marks]*

b The diagram below (Figure 11.11) shows the situation for the production of a demerit good.

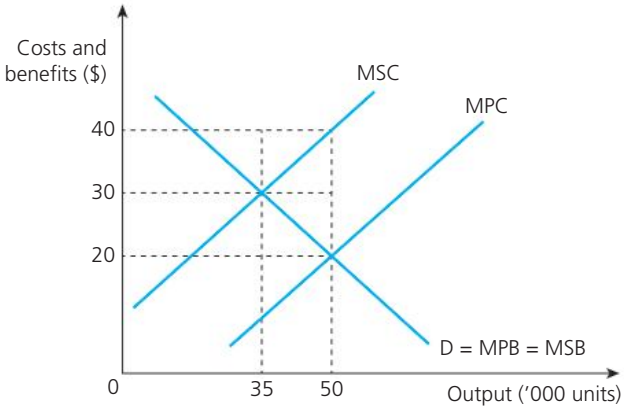


Figure 11.11 The impact of demerit goods

i Identify the level of output in a market without government intervention. *[1 mark]*

ii Identify the socially optimal level of output and the price charged. *[2 marks]*

iii Calculate the value of the welfare loss due to the overproduction of the demerit good. *[2 marks]*

Common pool resources (AO2, AO4)

A third category of market failure (in addition to positive externalities of production and consumption, and negative externalities of production and consumption) is **common access resources** (CPRs). They are not owned by any private individual or firm, so do not have a price, but are available for anyone to use without direct payment. Hence, without government intervention, these resources are vulnerable to overuse and abuse. Examples include fisheries (fishing grounds), irrigation and drainage systems, pastures, rivers, forests and the atmosphere.

◆ **Common access resources** (or **common pool resources**) are those that are non-excludable but rivalrous in consumption and create a situation of tragedy of the commons. Hence, they result in negative externalities and unsustainable production.

Even for renewable resources, such as fish stocks, the overuse of these resources can lead to a long-term decline beyond the sustainable level in the stock of natural capital. Common pool resources, sometimes referred to as **common pool resources** (CARs), have three key characteristics:

1 **Non-excludable**

Non-excludability means that it is not possible to prevent those who are unwilling or unable to pay from benefiting once the resource is provided. These users are known as free riders. As CPRs are freely available, there tends to be over-consumption of such resources. For example, in the absence of property rights or in areas where fishing grounds are poorly protected, commercial fishing businesses will try to maximize their fish catches, thereby leading to overfishing. Hence, common access resources are difficult to sustain, which is why governments intervene to prevent the depletion or destruction of such goods.

◆ **Non-excludability** is one of the two characteristics of pure public goods, where it is not possible to exclude someone from using a good or service even if they have not paid for it, because it is not practical to charge a price.

2 **Rivalrous**

This characteristic means that the usage of a resource (such as fisheries or forestries) reduces the amount that is available for others to use or consume. Common pool resources are rivalrous because consumption of these resources diminishes the quantity and possibly quality of remaining resources for the current and future generations to use. In the case of over-consumption of CPRs, sustainability becomes a real issue so generations to come may be deprived of any access to these resources.

3 **Tragedy of the commons**

Common pool resources create a situation known as the **tragedy of the commons**. This means, degradation, depletion or destruction of a common access resource caused by the problems of overuse (or abuse) and over-consumption. This economic and social problem was first articulated by American ecologist Garrett James Hardin (1915–2003) who warned of the threats of over-population and over-consumption. His article titled 'The Tragedy of the Commons' in *Science* journal (1968) used the example of grazing of the commons (fields) to explain that it was in each herder's best interest to allow any additional cows onto the grazing land. This rational economic decision means that each additional cow added to the individual herder's marginal utility, while the damage to the common land was shared by all herders who used the land. However, the long-term consequence of each herder's rational economic decision was market failure as their actions caused degradation, depletion and destruction of the common pool resource to all users and, hence, society as a whole.

◆ The **tragedy of the commons** refers to the degradation, depletion or destruction of a common pool resource caused by the problems of rivalry and overuse.

Another contemporary example is commercial fishing. Firms face the temptation to catch as many fish as possible – or even overfish – because if they do not, rival firms will. Collectively, this leads to fish stocks being depleted beyond their sustainable level, even though no single firm may have intended to do so. A tragedy of the commons occurs because there is over-grazing or over-harvesting of a particular common pool resource. Again, the absence of a price mechanism means the market fails to allocate scarce resources in an efficient manner.

The same applies to pollution of the natural environment and the subsequent damage to the ecosystem. Examples include air pollution caused by over-congested roads in cities such as Bangkok, Mumbai, New Delhi, Washington DC and Toronto.

CASE STUDY

The world's most congested cities

The table below shows the world's most congested cities, according to *Business Insider*, based on the average number of hours lost in congestion per year, adjusted for the city's population:

Rank	City	Country	Hours lost p/a
1	Bogota	Colombia	272
2	Rome	Italy	254
3	London	United Kingdom	227
4	Mexico City	Mexico	218
5	Moscow	Russia	210
6	Saint Petersburg	Russia	200
7	Rio de Janeiro	Brazil	199
8	Boston	USA	164
9	Istanbul	Turkey	157
10	Sao Paulo	Brazil	154

Source: https://www.businessinsider.com/cities-with-worst-traffic-in-the-world-2019-2?r=US&IR=T#1-moscow-russia-20

Unsustainable production creating negative externalities

Owing to the nature of non-excludability and the resulting tragedy of the commons, production is unsustainable. Essentially, this means the exploitation of common pool resources results in unsustainable production and negative externalities of production and consumption. Examples of such external costs include the damaging impact to societies and economies from deforestation, overfishing, congestion, pollution, soil erosion, destruction of natural habitats and climate change.

CASE STUDY

Overfishing

Overfishing is the commercial practice of removing too many fish from their natural habitat such that the species cannot replenish themselves in time. This can result in the fish stock becoming severely depleted or under-populated, thus endangering the species. There is a global trend in overfishing, with some fishing businesses illegally overfishing. Overfishing creates a threat to sustainability and causes negative externalities of production (such as pollution caused by too many commercial trawlers operating in the oceans) and negative consumption externalities (such as the immense amount of food waste from over-consumption of seafood buffets).

Figure 11.12 Overfishing is not a sustainable economic activity

CASE STUDY

Coca-Cola

Coca-Cola is the world's most successful consumer drinks company, with over $40 billion in sales revenue (which equates to $109,589,041 every day of the year!). However, this also means the company is the planet's largest plastic polluter. Although the multinational giant is often associated with fuelling child obesity, the company is increasingly being linked with plastic waste and pollution, with more than 100 billion plastic bottles produced each year – that is the equivalent of 273,972,602 plastic bottles every day of the year! In response to negative media exposure, the company's global chief executive, James Quincey, stated that the Coca-Cola Company aims to recover every plastic bottle that the company sells, and to use 50 per cent of this for new bottles, by 2030. The company also announced that it would replace the plastic shrinkwraps used in multipacks with 100 per cent recyclable cardboard.

Top tip!

Common access resources differ from public goods (see Chapter 12) because they are *subtractable* (overused), rather than *non-rivalrous* (the use by an individual does not diminish the amount available for others).

ATL ACTIVITY 11.2

Is there really plenty more fish in the sea, as the American proverb says? Watch this short interview from Sky News with Mark Carney (former Governor of the Bank of England and Chair of the Monetary Policy Committee), who talks about the causes and possible consequences of overfishing and pollution in the oceans caused by plastic waste:

https://youtu.be/1aLyYI8-5TM

Discuss the possible and most effective solutions to these market failures.

◆ An **indirect tax** (or **Pigouvian tax**) is a form of intervention in response to negative externalities and common pool resources by taxing the producer or consumer in order to make them pay for the negative externalities of production and consumption.

Government intervention in response to externalities and common pool resources (AO2, AO4)

In the IB Economics syllabus, there are nine types of government intervention in response to externalities and common pool resources (CPRs). These can be remembered by the acronym **PIGLETS CC**:

Indirect (Pigouvian) taxes

An **indirect tax** is used to 'internalize' negative externalities. In other words, the buyer and/or seller pays for the true costs of their actions without any burden being placed on third parties. For example, indirect taxes are imposed on fuel (petrol or gas), alcohol, tobacco, sugary drinks, congestion charge and the use of plastic carrier bags. Indirect taxes are corrective measures used in order to reduce or eliminate the external costs (negative externalities) associated with the production of certain products. They are also known as **Pigouvian taxes**, named after British economist Arthur Cecil Pigou (1877–1959),

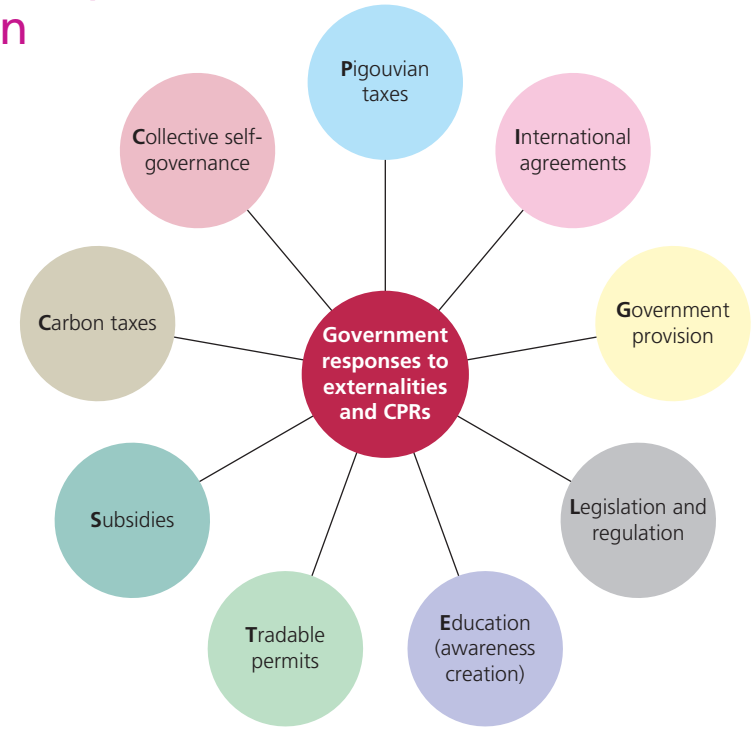


Figure 11.13 Government intervention in response to externalities and common pool resources (CPRs)

who developed this idea in his book *The Economics of Welfare* (1920). Pigou noted that private firms, in pursuit of their own marginal private benefits, were not concerned with any negative externalities (external costs) to others in society. This is because without government intervention, they do not have any incentive to do so. As Pigouvian taxes are imposed on demerit goods, they are often also referred to as 'sin taxes'.

In Figure 11.14, the Pigouvian tax is imposed on a producer creating negative externalities. For example, a logging company causes a divergence between the marginal private costs (MPB) and marginal social costs (MSC) of production. The main purpose of using an indirect tax is to make the price of the product equal to D = MPB = MSB, thereby creating a more socially efficient allocation of resources. Point A shows the free market equilibrium, whereas point B shows the socially optimal level of output (where MSC = MSB). The indirect tax raises costs for the producer, thereby causing the S = MPC curve (or the supply curve of the private firm) to shift leftwards towards the S_2 = MSC curve. The lower supply causes the price to rise from P_e (the market equilibrium price) to P_{opt} (the socially optimal price level) and the quantity demanded to fall from Q_e (the market equilibrium quantity) to Q_{opt} (the socially optimal quantity).

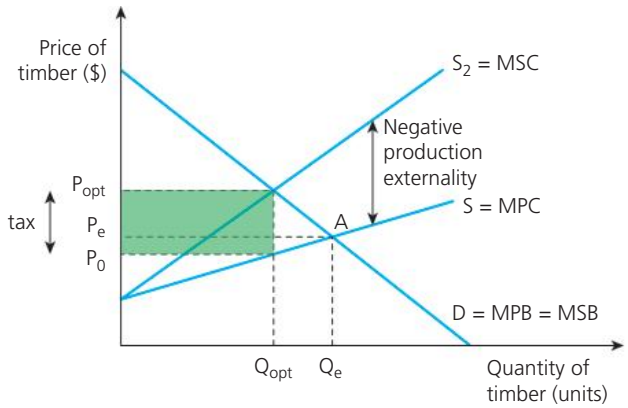


Figure 11.14 Indirect tax on a logging company

The advantages of imposing an indirect tax on the production or consumption of products with negative externalities are that:

- It increases the price of the product and therefore should decrease the quantity demanded. For example, Ireland's tax on plastic carrier bags aimed at reducing plastic waste and litter problems reduced the quantity demanded by 90 per cent in the first three months of implementation.
- It creates tax revenue for the government (shown by the shaded green area in Figure 11.14) which can be used to raise public funds to deal with the external costs of negative externalities or to fund other interventionist measures, such as providing subsidies for firms to adopt green technologies or to fund government provision of merit and public goods and services.

Figure 11.15 Scotland introduced a 5 pence ($0.07) plastic bag tax in October 2013, following success in Ireland, Wales, Germany and Hong Kong

Top tip!

Chapter 4 showed that an indirect tax can be shown by either a parallel shift in the supply curve (caused by the imposition of a *specific tax*) or a non-parallel shift (caused by the imposition of an *ad valorem tax*).

Top tip!

A Pigouvian tax is shown by the vertical distance between the two supply curves in the indirect tax diagram (Figure 11.14). The consumer pays a higher price ($P_{opt} - P_e$) and the producer pays the remainder ($P_e - P_0$). The more price inelastic the demand for the product, the greater the proportion of the tax is paid by the consumer.

The disadvantages of imposing a tax on the production or consumption of products with negative externalities are that:

- the demand for many of these products (such as cigarettes, alcohol, gambling and petrol) tend to be price inelastic, so the tax may have little impact on the level of consumption and, hence, production
- the tax on many of these products is regressive (see Chapter 19) so has a greater impact on low-income earners than high-income earners
- it can encourage smuggling and unofficial market activities, creating a parallel market for demerit goods.

ATL ACTIVITY 11.3

Investigate the economic costs of sugary drinks in Mexico (or another country of your choice). Evaluate the extent to which a Pigouvian tax has reduced the problems associated with the production and consumption of sugary drinks.

ATL ACTIVITY 11.4

Investigate the extent to which indirect taxes on unhealthy choices (such as tobacco, alcohol and sugary drinks) are effective in making people choose healthier options. Be prepared to share your research findings with the rest of your class.

Carbon taxes

A **carbon tax** is similar to a Pigouvian tax but imposes the tax on carbon emissions. Essentially, the carbon tax identifies the price of greenhouse gas emissions or the carbon content of fossil fuels (such as coal, diesel fuel, propane, gasoline and natural gas) and imposes this external cost on the polluting industry. Scientists and economists associate excess carbon emissions with global warming and climate change, which has created negative externalities such as heatwaves, flooding, landslides, storms and droughts. Without carbon taxes, market failures exist as firms and households are not accurately charged for using fossil fuels. The purpose of the tax is to create incentives for firms (and households depending on the scope of the carbon tax) to reduce pollution, thereby reducing their tax liability. According to the World Bank, more than 40 countries use carbon taxes to combat the problems of negative externalities caused by over-production.

◆ A **carbon tax** is a tax on greenhouse gas emissions or the carbon content of fossil fuels in order to reduce pollution from particular industries by internalizing negative externalities of production.

For example, a factory that pollutes the air causes a divergence between the marginal private costs (MPC) and marginal social costs (MSC) of production, as shown in Figure 11.17. The effect of the per unit carbon tax ($P_{opt} - P_0$) is a fall in production, from Q_e to Q_{opt}, and a higher price for the product, from P_e to P_{opt}. The government earns tax revenue as shown by the shaded green area.

Figure 11.16 Carbon taxes are imposed on particular polluting industries

Carbon taxes are set by assessing the pollution costs and the associated administrative costs of controlling the pollution, thereby internalizing negative production externalities. The main advantage of using carbon taxes is that they impose additional costs to users of carbon fuels, thus users pay for the environmental and climate damage caused by their economic activities. In addition, an effective carbon tax can make it more economically rewarding for firms to switch to alternative (non-carbon) fuels, green technologies and energy efficient production. Like a Pigouvian tax, establishing the correct tax rate is key to the effectiveness of using carbon taxes in order to correct market failures. If the tax is too low, firms will simply continue to pollute, whereas if the tax is too high the escalated costs negatively affect profits, employment and consumers. The Organization for Economic Co-operation and Development (OECD) reported that the average carbon tax in 2018 was $8 per tonne. However, the United Nations has recommended a per unit carbon tax of at least $135 to prevent temperatures from rising above 1.5 degrees Celsius by 2030.

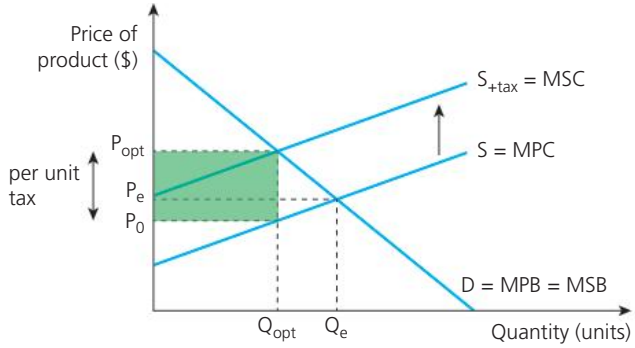


Figure 11.17 Carbon tax on polluting firms

The effectiveness of carbon taxes also depends on the extent to which firms are able to pass on higher energy costs to their customers. If the demand for their products is price inelastic (see Chapter 7), customers are less sensitive to price changes, so firms are able to charge higher prices. In other words, consumers bear a greater proportion of the carbon tax burden. However, carbon taxes have been criticized for being regressive, as they account for a larger proportion of the income of poorer households. In addition, higher income households are in a better financial position to switch to using more energy-efficient technologies. Nevertheless, the simplicity and impartiality of carbon taxes mean that this method of combating pollution can be highly effective for many countries.

ATL ACTIVITY 11.5

To find out more about carbon taxes and the problems associated with carbon emissions, visit the Carbon Tax Center website: www.carbontax.org. It provides a wealth of resources and case studies to further develop your understanding of this part of the IB Economics syllabus. For example, find out which countries are the world's largest polluters of carbon emissions, which fuels are the worst pollutants and what alternative solutions there are.

WISE ChoICES

To what extent does **intervention** in the form of indirect and carbon taxes help to achieve **equity** and **sustainability**?

Legislation and regulation

Another government response to deal with the externalities of production and consumption and common pool resources is the use of legislation and regulation. In the case of merit goods, laws can be used to encourage greater consumption of goods and services with positive externalities. For example, legislation and regulation are used to ensure there is compulsory education for children and making it a legal requirement for school-aged children to be vaccinated against certain diseases. Governments also use legislation and regulation to control the volume of traffic, such as in London, Milan, Oslo, Paris, Stockholm and Singapore.

Legislation refers to the laws on the use of scarce resources, such as laws to protect fish and marine life or legislation to curb carbon emissions from private and commercial motor vehicles. **Regulation** refers to the management of complex rules, laws and policies that firms need to comply with. For example, fishing businesses needs to act in accordance with government regulations, such as fishing licences, quotas and rules about the types and weight of fish that can be caught at sea.

◆ **Legislation** (in the context of asymmetric information) refers to laws stipulated by the government as a response to tackling imperfect information in markets, such as requiring health warnings on cigarette packets.

◆ **Regulation** (in the context of asymmetric information) refers to the act of monitoring and controlling the activity of firms, such as advertising information required to be legal, decent, honest and truthful.

In the case of cigarettes and tobacco products, legislation and regulation regarding where people can smoke will lower the marginal private benefit as consumers are inconvenienced. Smoking bans in restaurants, the workplace and other public areas mean that smokers will tend to derive less pleasure from smoking as a result of government intervention.

In Figure 11.18, legislation and regulation shift the demand curve for cigarettes (or any other demerit good) from MPB to MSB owing to reduced benefits of consumption. Thus, consumption falls from Q_1 to Q_2. Therefore, this also reduces the negative externalities associated with smoking, such as passive (second-hand) smoking and litter, as shown by the area of welfare loss.

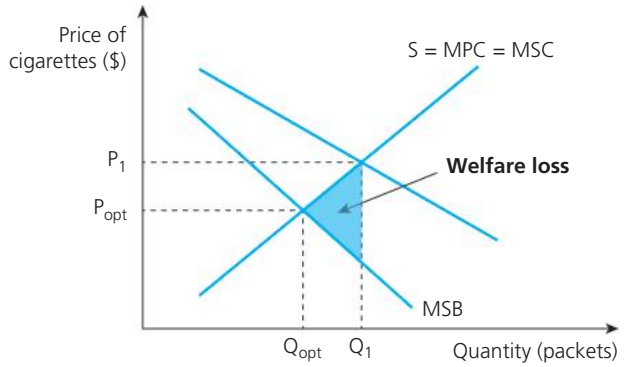


Figure 11.18 Regulation and negative consumption externalities

However, such legislation and regulations can cause underground (illegal) markets to develop where the demerit good or service can be purchased, often at a very high price. In addition, unless the penalties are extremely high and consistency enforced, people may choose to break the rules and regulations, such as under-age smoking, drinking and gambling.

CASE STUDY

Legislation and regulation in action

Government legislation and regulation can be very effective in correcting market failures. Parents in the UK are fined if they do not send their children to school – since 2015, full-time education in the UK is compulsory for children up to the age of 18. Since 2003, the Hong Kong government has imposed a fine of HK$1,500 ($193) for each item of litter. Smoking in public places has been banned in Hong Kong since 2007. There is also a HK$5,000 penalty ($645) for those who eat on public transport such as on Hong Kong buses or trains or smoke in public areas, such as shopping malls, bars, restaurants, airports, railways stations and the beach.

PAPER 1 EXAM PRACTICE QUESTION 11.3

With the aid of a suitable market failure diagram, explain why a government might use congestion zone charging to reduce the use of motor vehicles in city centres such as London and Paris. *[10 marks]*

Education – awareness creation

Educating the general public about the costs of consuming demerit goods and the benefits of consuming merit goods is another approach to correcting market failures. Education, or awareness creation perhaps through advertising, is about influencing consumer behaviour. Examples include:

- the '5 A Day' programme run in countries such as Germany, UK and USA in line with the World Health Organization's recommendation that people should eat at least 400g (or 5 portions) of fruits and vegetables each day (see Figure 11.19)
- schools educating students about safe sex and family planning

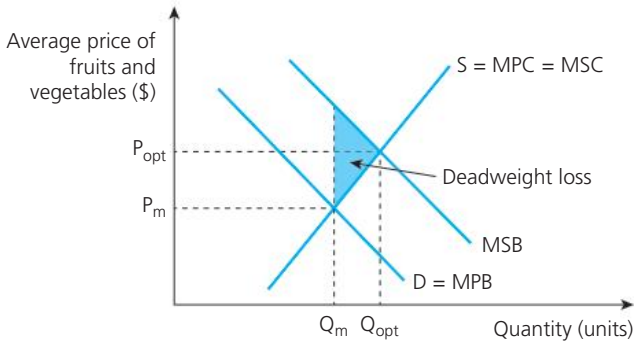


Figure 11.19 Education (awareness creation) and positive consumption externalities

- awareness creation programmes about health screening, such as encouraging more people to measure their levels of cholesterol, blood pressure and risks of chronic illnesses such as diabetes, stroke and cancer
- advertising about the importance of environmental protection, including the social benefits of conserving energy, waste minimization and recycling.

The impact of education (awareness creation) on the demand curve for fruits and vegetables is an outwards shift from MPB to MSB. This increases consumption from Q_m to Q_{opt}. Healthier people should mean fewer absences from work and school, with external benefits to society and the economy as a whole. However, as there are costs for funding education and awareness creation, which does not translate to consumer or producer surplus (see Chapter 5), and also administrative costs associated with the provision of education, there is a deadweight loss.

Advantages of education and awareness creation to combat the problems of market failure include:

- Behaviour and consumption patterns of individuals and firms change – there is a rise in the consumption of merit goods and a fall in the demand for demerit goods. For example, people learn about the dangers of smoking or gambling, so fewer people smoke or go to casinos.
- Successful advertising may lead to a cultural change in behaviour, such as healthier diets, increased use of electric cars, recycling, waste reduction and the use of renewable energies.

Disadvantages of education and awareness creation to tackle the problems of market failure include:

- Not all awareness creation campaigns are effective. Education and marketing tactics do not necessarily work in changing people's social behaviours.
- It can take a long time to educate people and for the advertised message to be accepted and acted upon. For example, although people are aware of the dangers of diesel pollution and plastic wastes, it can take many years for society to change its routine habits.
- There is an opportunity cost of government expenditure on education and awareness creation. This means, the money might have been spent on something else deemed to be more beneficial to society. Some people argue that alternative measures such as taxation and regulation are more effective than spending time and resources to educate people to reduce their consumption of demerit goods.

CASE STUDY

Plastic waste in the world's longest river

Plastic waste is illegally dumped into the River Nile, the world's longest river. Over 300 million people in eleven countries rely on the River Nile for their livelihoods. However, the unlawful dumping of plastics has threatened ecosystems and people's lives as massive amounts of micro plastic waste get back into the food chain in countries such as Ethiopia, South Sudan, and Uganda.

Tradable permits

Tradable permits are government-regulated emissions trading schemes using a market-based approach to reduce production to a more socially efficient level. The regulator sets a limit (the cap) on the total amount of emissions or pollution allowed in an industry and firms are issued emissions permits, such as the right to pollute up to 100 units of carbon dioxide per month. The scheme allows pollution permits to be freely traded (based on the price mechanism) thus enabling more efficient firms to sell their excess permits. Hence, tradable permits are often referred to as **cap and trade schemes** (CATS).

◆ **Tradable permits** (or **cap and trade schemes**) are government-regulated emissions trading schemes that limit pollution in an industry to a more socially efficient level. Efficient firms can sell any excess permits that they do not need to use.

While CATS help to reduce carbon emissions, they can also raise a significant amount of revenue for the government. The European Union's Emissions Trading Scheme (ETS), established in 2005, is one of the world's largest cap and trade schemes, covering more than 11,000 factories and power stations across all 27 EU member countries. The scheme aims to regulate economic activity to prevent dangerous climate change.

With economic growth and higher demand for goods and services, the demand for pollution permits can also increase but the cap on pollution creates higher prices for tradable permits,

as shown in Figure 11.20. The increased demand for tradable permits from D_1 to D_2 increases the market price from P_1 to P_2. There is no change in the quantity supplied as the government has fixed (capped) the number of pollution permits for the industry. The aim of the pollution permits is to make the polluters pay for the external costs of their production. Government revenue also increases from $0,P_1,A,Q_e$ to $0,P_2,B,Q_e$.

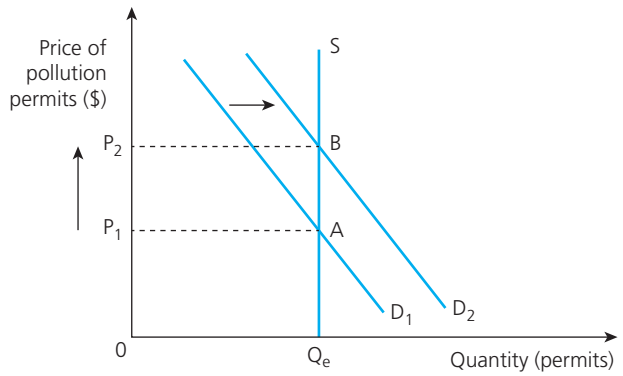


Figure 11.20 Impact of higher demand for tradable permits

Tradable permit schemes mean that the largest polluters have to purchase more permits, which increases their costs of production, so makes them less competitive and less profitable. By contrast, the more environmentally friendly producers can gain extra revenue from selling their surplus permits, thereby making them more competitive and more profitable. Hence, tradable permit schemes create an incentive for firms to develop greener or cleaner technologies in the long run, thereby leading to a fall in the demand for and supply of pollution permits.

Over time, the government may choose to reduce the number of pollution permits, thereby shifting the supply of carbon emissions permits from S_1 to S_2 in Figure 11.21. This raises the market price of pollution permits from P_1 to P_2, creating a greater incentive for firms to be more efficient and invest in clean technologies.

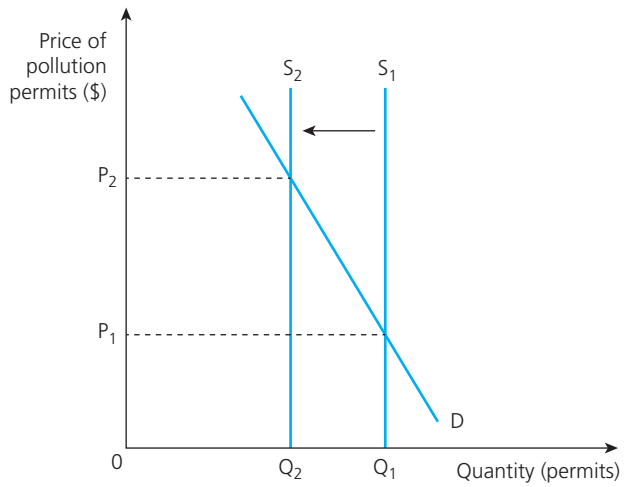


Figure 11.21 Impact of reducing tradable permits

Critics of cap and trade schemes argue that CATS are anti-competitive (against smaller firms) and they can cause job losses (due to higher costs of production). Furthermore, in a more globalized world, multinational companies can shift production to other countries to avoid the constraints of pollution permits and CATS. Environmentalists have also argued that higher prices of tradable permits have proved ineffective in reducing carbon dioxide emissions to levels necessary to stop global warming and climate change, especially as the demand for pollution permits is highly price inelastic. Notice that government revenue from the sale of permits also increases from P_1Q_1 to P_2Q_2 in Figure 11.21 due to the relatively price inelastic demand for pollution permits. Another criticism is that the cap is often deemed to be too high – China's national cap and trade programme, for example, sets the limit at up to 5 billion tons of carbon emissions. This does little to combat the underlying problems of pollution and climate change.

International agreements

Governments can also create and develop international agreements to respond to the growing problems of negative externalities of production and consumption. Some of these may be bilateral agreements (between two countries) while most are multilateral agreements but are nevertheless essential for environment protection. Most international agreements are legally binding for the member countries that have formally ratified the treaties (contracts). According to the International Environmental Agency (IEA), more than 3,000 international environmental agreements have been registered, covering areas of environmental governance such as freshwater resources, hazardous substances, global marine environments, nature conservation and nuclear safety.

For example, the Kyoto Protocol is an international agreement based on the United Nations Framework Convention on Climate Change (UNFCCC) of 1992, which commits member

states to reducing greenhouse gas emissions, including carbon dioxide, methane, nitrous oxide, hydrofluorocarbons, perfluorocarbons and sulphur hexafluoride. The Kyoto Protocol (effective 2005) involves 192 party states and is based on the principle of common but differentiated responsibilities. In other words, the agreement recognizes that individual countries have different resources and abilities to deal with global warming and climate change. Essentially, this means high-income countries are more obligated to do their part in reducing global emissions of greenhouse gases, especially as they are historically responsible for contributing to most of the greenhouse gases in the atmosphere today.

While international agreements hold countries to account for their actions and potential harm to the environment, they require full commitment from all countries, not just those that sign treaties at different points in time. For example, Canada has withdrawn from the Kyoto Protocol while the USA has not ratified the agreement. In addition, high-income countries and large multinational companies have been known to deliberately exploit low-income countries despite national and international environmental laws and agreements. An example is the immense environmental damage caused by deforestation in Amazon countries. This means, the destruction of rainforests for commercial reasons such as logging, cattle ranching, colonization and commercial agriculture. Another example is the number of large oil companies exploiting Nigeria's oil supplies, contributing to loss of ecosystems and damage caused by air and water pollution.

Essentially, the failure (unwillingness) of high-income countries to be part of international treaties or the tendency for them to ignore them seriously undermines international agreements to protect the natural environment.

ATL ACTIVITY 11.6

Research the economic reasons for deforestation in the Amazon. Discuss the impact of commercial deforestation on the nine countries that share the Amazon. A good starting point is **https://rainforests.mongabay.com/amazon/amazon_destruction.html** where you can also find out about the importance of rainforests for our planet.

Collective self-governance

Collective self-governance refers to voluntary communal actions to combat the problems of negative externalities and the problems associated with the exploitation of common pool resources. Elinor Ostrom, the first female winner of the Nobel Memorial Prize in Economic Sciences (2009), made huge contributions regarding how to overcome the problem of the tragedy of the commons by using collective self-governance and voluntary mutually agreed rules, including community-agreed sanctions if these rules are broken.

◆ **Collective self-governance** refers to voluntary communal actions to tackle the problems of negative externalities and the problems associated with the exploitation of common pool resources.

Ostrom (1993–2012) argued that environmental sustainability is possible if people work together to create democratically agreed and adaptable rules for the exploitation of common pool resources. This is because local communities make more informed decisions about the use of community resources rather than governments or other regulating authorities, because they have knowledge about the local culture and context and because they are directly affected. Hence, local collective self-governance of resources and the collective self-determination to protect their best interest can be effective in ensuring resource sustainability. This includes formulating social norms of proper conduct, such as social pressures to conform to the 3Rs of environmental protection (reducing, recycling and reusing the scarce resources of society).

Collective self-governance often relies on other methods, such as taxation and education, but can have significant and long-lasting impacts. For example, it is becoming increasingly unacceptable in many societies to consume so many plastics (especially single-use plastics such as plastic straws, food packaging, drinks containers and balloons). For example, the Plastic Free Challenge (**http://www.plasticfreechallenge.org**) uses social media to challenge people to voluntarily engage in reducing their use of plastics; some 300 million tons of plastic waste each year.

Collective self-governance has gained attention in the sustainable tourism industry. This includes the collective efforts to conserve natural resources, protect natural environments, safeguard wildlife, and preserve the local area's heritage and culture. Self-governance has become an

important aspect when developing and managing sustainable tourism activities while providing authentic experiences for tourists. Collective self-governance therefore also creates socio-economic benefits for local communities such as employment opportunities and higher income earnings.

CASE STUDY

Car litter fines in the UK

Motorists are liable for a £150 ($195) fine for littering from their cars – even if a passenger does it. The on-the-spot fine acts as a huge deterrent to the antisocial behaviour of some people. The government estimates the cost to taxpayers to clean litter from UK streets caused by car littering is about £700 million ($910m) per year. The scheme also encourages people with dash cams in their cars or those with smartphones to take video evidence of offenders in a communal attempt to keep local roads clean. The huge opportunity cost of littering also encourages people to take their litter home or to take responsibility for disposing of their own trash.

Figure 11.22 Nudge theory is used to deter littering

Source: Adapted from Express https://tinyurl.com/vm6k5tt

◆ **Subsidies** are a form of financial assistance from the government to domestic firms by lowering their costs of production in order to help the firms compete against foreign imports. Subsidies encourage output, reduce the price of certain products, or keep down the cost of living for the domestic country's citizens.

Subsidies

A **subsidy** is a sum of money given to a producer to reduce the costs of production (thereby encouraging higher levels of output) or to consumers to reduce the price of consumption. For example, some governments subsidize public transport to discourage people from using private cars (see Figure 11.23).

In Figure 11.23, the specific subsidy given to the public transport provider is shown by the vertical distance between the MPC and MSC curves. This represents the value of the positive externalities of production and consumption. The subsidy helps to reduce production costs, thus shifts the supply

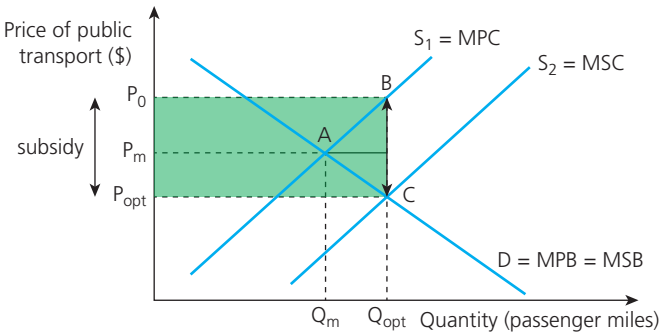


Figure 11.23 The effects of a producer subsidy on public transportation

curve outwards from MPC to MSC. The producer passes some of this per unit subsidy to consumers in the form of lower prices (from P_m to P_{opt}) and keeps the remainder (P_0 to P_m) in the form of lower production costs. Hence, the subsidy provides an incentive for more people to consume the good, thereby increasing the quantity from Q_m to Q_{opt}, and reducing congestion on roads. The total amount spent by the government on the subsidy is shown by the green shaded area.

However, as the cost of funding the subsidy does not equate to additional producer or consumer surplus (see Chapter 5), and there are also administrative costs associated with the provision of subsidies, the subsidy creates market inefficiencies known as a **deadweight loss**. This is shown by the triangular area ABC in Figure 11.23 and clarified in Figure 11.24 below. The following can be seen from this diagram:

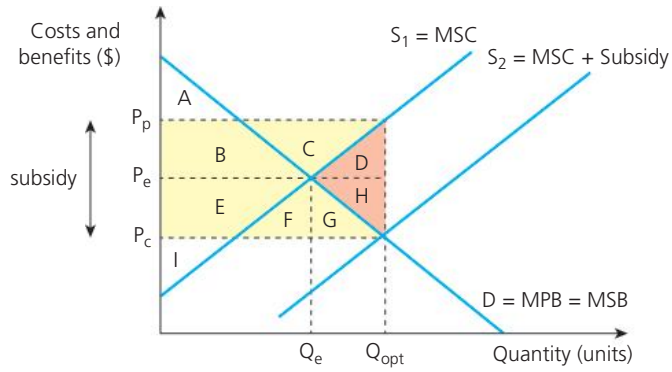


Figure 11.24 Subsidies and deadweight loss

- Before the subsidy, the price was P_e, so consumer surplus was shown by the area A + B.
- Producer surplus was originally E + I.
- Following the subsidy, the price for the consumer falls to P_c. Hence, consumer surplus increases to area A + B + E + F + G.
- Producers now receive a price of P_p owing to the subsidy. Hence, producer surplus is now shown by the area B + C + E + I.
- The total amount spent by the government on the subsidy is shown by the shaded green area (B + C + D + E + F + G + H).
- However, as area D + H does not get transformed into either consumer or producer surplus, this represents inefficiency, that is, a deadweight loss (or welfare loss).

Limitations of using subsidies to correct market failures include:

- Just as with indirection taxation imposed on negative production and consumption externalities, it is difficult to set a precise subsidy that ensures an optimal allocation of resources (merit goods, for example).
- The social return on the production and consumption of merit goods (such as education or the provision of public libraries, sports facilities and museums) is difficult, if not subjective, to measure.
- If the price elasticity of demand for the good or service is inelastic, the lower price (due to the subsidy) has little impact on the quantity demanded.
- There is always an opportunity cost to the provision of subsidies as the money could have been used for other government projects that may reap greater social gains.

Other similar government schemes include:

- Subsidizing consumers to switch to less damaging or polluting activities. For example, the government in China has given subsidies of up to 60,000 yuan ($9,800) to buyers of electric and hydrogen vehicles in a bid to combat rising air pollution in the country.
- Providing funding for clean technologies to create incentives for firms to adopt green technologies, such as subsidies for the adoption of renewable energy sources, including wind, sun and biomass.
- Forest investment funding programmes help to reduce pollutants from deforestation and the negative impacts of soil degradation.
- Subsidies to provide free drinking water in shopping malls and other public places to discourage the overuse of single-use plastic bottles.

Government provision

Direct provision of goods and services is a final government response to market failure, such as provision of public transportation services, public libraries, museums, communal parks, education and healthcare services. As the market fails to provide an adequate supply of merit and public goods, government provision helps to ensure that the socially optimal level of output is achieved.

Figure 11.25 Government provision of education services creates positive externalities

The key advantage of direct government provision is that merit goods and public goods become accessible to everyone, regardless of their income or social status. This helps to improve the economic well-being of society as a whole.

However, government provision also has its limitations, including:

- Economic inefficiency – Direct provision of healthcare services, for example (such as the UK's National Health Service), might encourage consumption beyond the socially optimum level because people do not have to pay for these. Hence, people are still likely to overuse the services.
- Opportunity cost – Direct government provision of a good or service entails an opportunity cost as the money could have been spent on something else, such as paying off government debt or possibly lowering tax rates in the economy.
- Issues of inequality – In the case of a shortage of supply due to excess demand for the good or service provided by the government, it can be difficult to decide who should be able to take advantage of the free or subsidized product. For example, the waiting list for a hip replacement operation in a government hospital may be very long.

Strengths and limitations of government policies (AO3)

The previous section referred to the strengths and limitations of the various government policies to correct externalities and the approaches to managing common pool resources (the PIGLETS CC acronym). These reasons can be examined in three ways:

Challenges involved in measurement of externalities

There are difficulties and challenges involved in measuring the value of externalities in an accurate way. For example, what is the value of negative externalities such as noise pollution or carbon emissions, and can these even be measured precisely? What is the value of positive

externalities from the provision of public goods such as street lighting and flood control systems or from merit goods such as the provision of universal primary education? Nevertheless, the challenges in measuring the impact of noise pollution does not mean the government has an excuse to ignore the problem. For example, airport authorities regulate the number of night flights to limit noise pollution to nearby residents.

Figure 11.26 Is it possible to accurately measure the true value of education?

Degree of effectiveness

It is not always clear whether government intervention corrects market failures in an effective way. Economists use the *ceteris paribus* assumption (see Chapter 1) to explain how changes in policy can influence market outcomes. However, in reality, there are other factors that may change the amount of demerit and merit goods consumed in the economy. For example, the global decrease in the number of smokers is partly due to several interrelated factors (rather than any single government interventionist policy), such as:

- laws banning smoking in public areas
- health initiatives established by the government
- major retailers opting to stop selling cigarettes (as part of their own corporate social responsibility)
- anti-smoking advertisements in the mass media, including television, radio and newspaper adverts
- pressure group activities from organizations such as Action on Smoking and Health (ASH), publicizing the risks associated with tobacco smoking and campaigning for greater restrictions on cigarette and tobacco sales
- growing education and awareness creation about the dangers and financial costs of smoking.

In addition, government policies can be limited by *time lags* (such as education and awareness creation) and the *costs* of enforcement (such as fining people who litter) or cost of government provision (such as pubic healthcare services). For merit goods, there also needs to be consideration of the costs of financing government provision. For example, according to *Healthline Media*, around 900,000 Americans caught the common flu in 2018, which caused 80,000 deaths.

CASE STUDY

The economic costs of the common flu in the USA

- The USA spends $10.4bn per year in direct medical expenses.
- The country loses $15.3bn a year from the common flu due to loss of earnings (that's the equivalent of 17 million workdays a year!).
- Productivity losses amount to more than $21 billion a year.
- Around 80 per cent of children who die from flu had not received a flu vaccination.

Source: **Healthline:** ***The Flu: Facts, Statistics, and You***
https://www.healthline.com/health/influenza/facts-and-statistics#7

Consequences for stakeholders

Different stakeholders are affected by government intervention in different ways. As not all stakeholder groups are necessarily equal, there is scope for exploitation. For example:

- Large multinational companies can exploit their market power by relocating to low-income countries where legal limits on pollution are not imposed or controlled and monitored to the same extent.
- Government legislation and regulations such as environmental standards tend to favour large multinational companies that have the financial resources to comply, whereas smaller firms that face higher costs of compliance may struggle to survive. For example, larger firms with greener and cleaner technologies are more likely to be able to operate within their pollution permits and sell excess rights to smaller and less-efficient rival firms.
- Pigouvian taxes are regressive, so have a larger impact on low-income earners or smaller firms while high-income earners and large firms with strong market share can afford to continue burdening other stakeholders (third parties). Their low income and price elasticities of demand (see Chapter 7) mean government interventionist policies have minimal impact on their spending behaviour.

Hence, value judgements can impact the use of certain policies to deal with market failures and common pool resources. For example, there are growing concerns about drive-thru restaurants and coffee shops and the pollution generated by motorists purchasing their food or drinks in this way. Should drive-thru restaurants and customers be charged for the additional pollution created from their production and consumption? Clearly, some stakeholder groups (such as environmentalists and the local community) will be more in favour of regulating this industry than other stakeholders (such as restaurant owners and their customers).

Figure 11.27 Should drive-thru operators and customers pay for the pollution they create?

Importance of international co-operation (AO3)

To deal with the limitations of government policies in dealing with market failures and common pool resources, there is a need for international co-operation. For example, policy responses to the threat to sustainability are limited by the global existence of the lack of ownership of common pool resources, such as fishing in international waters. International co-operation is also vital

owing to the escalating global demand for scarce resources, including non-renewable fossil fuels, due to higher levels of economic activity and globalization. This requires consideration of the global nature of sustainability issues, the challenges faced in international co-operation, and the need for monitoring and enforcement.

Global nature of sustainability issues

Global warming, climate change and increased concerns about inequalities have prompted governments around the world to consider ecological and social sustainability rather than just economic sustainability. Essentially, this means governments are placing equal emphasis on the planet and people, not just profits. The most widely used framework in dealing with sustainability issues is the Sustainable Development Goals (SDG) of the United Nation's Development Programme (UNDP).

There are 17 interconnected targets as part of the UNDP's SDG, which Ban Ki-moon (Secretary-General of the United Nations at the time) described as the 'most transparent and inclusive process in UN history'. This international agreement between 193 countries is a major scheme to achieve a better and more sustainable future for everyone on the planet. The SDGs address global challenges to sustainability, which are largely caused by market failures and negative externalities, including issues related to poverty, healthcare, education, infrastructure, inequalities, climate change and environmental degradation, and human rights. The seventeenth goal itself is about global partnerships for sustainable development, including lobbying governments to boost development financing. However, there is much debate and concern about how the goals are to be funded in order to achieve all 17 goals by the end of 2030, and whether all countries have the resources and willpower to meet these targets. The SDGs are covered in detail in Chapter 32.

WISE ChoICES

With reference to common access resources, discuss the role of **intervention** due to the threat to economic **sustainability**.

Challenges faced in international co-operation

Cultural differences and local contexts can also create challenges for international co-operation. Laws differ regarding the minimum age that a person must be before being legally allowed to purchase cigarettes or alcohol. Smoking also generates huge tax revenues in countries such as Andorra (where the average person smokes more than 6,400 cigarettes a year!), Luxembourg, Belarus, Albania, Belgium, Russia and China. Hence, they may be more reluctant to ban the sale of tobacco products. While it is illegal to sell alcohol in Iran, Bangladesh, Brunei and Saudi Arabia, other countries collect huge tax revenues from the sale of alcoholic beverages (including Belarus, the Czech Republic and Germany). Some countries including Belgium, Spain, France, Netherlands, Mexico, Japan and China also rely on the export earnings from the global sales of alcoholic drinks.

There will also be debate about how much each country can and should contribute to the problems of global warming and climate change. For example, Canada's withdrawal from the Kyoto Protocol and the USA's refusal to ratify this international agreement poses huge threats to environmental protection and sustainability. Similarly, Japan (being the world's largest consumer of tuna and whale meat products) refuses to sign international protocols regarding overfishing. The nation consumes more than 600 million tons of seafood each year, but decades of overfishing have led to fish stocks reaching unsustainable levels. China, the USA, Indonesia and South Korea have also been accused of overfishing beyond ecologically sustainable levels.

Monitoring and enforcement

There are both financial and opportunity costs of monitoring and enforcing the policies used to correct market failures. For example, the following corrective measures need to be monitored and enforced, with sufficient funding and resources:

- laws to regulate where people can drive, cycle and gamble
- laws to make it illegal for people to smoke, eat or to talk on a mobile phone while driving

- motorcyclists being made to wear a helmet and car passengers having to wear seat belts
- achieving the UNDP's SDGs requires careful monitoring and control, such as achieving gender equality in society, inclusive and equitable quality education, and the reduction of inequalities within and between countries.

Furthermore, the government has no direct control over the quality of the goods or services produced in underground (parallel) markets, which in some cases can be dangerous for human consumption. For example, there have been cases of illegally distilled vodka in Europe and tainted baby milk powder cases in China. People may still also choose to break the rules, such as underage smokers and drinkers who try to bypass the law by obtaining false identity cards. This is more likely to occur if the fines or penalties are not set high enough to discourage consumption of the demerit goods or if the laws are not consistently and thoroughly enforced.

WISE ChoICES

To what extent does the lack of a pricing mechanism for common pool resources pose a threat to **sustainability** and limit **economic well-being**?

Chapter summary

- Market failure occurs when the price mechanism cannot allocate resources efficiently but causes external costs or external benefits of production or consumption.
- Marginal social benefit (MSB) is the additional gain to society from the production or consumption of an extra unit of output, that is, the sum of MPB and marginal external benefits.
- Marginal social cost (MSC) is the extra cost of an economic transaction to society, that is, the sum of MPC and marginal external costs.
- Positive externalities (external benefits) arise from the production or consumption of goods or services that create positive spillover effects to third parties. These products are known as merit goods.
- Negative externalities (external costs) arise from the production or consumption of goods or services that create negative spillover effects to third parties. These products are known as demerit goods.
- Common pool resources (CPRs) are goods or services that have the characteristics of non-excludability and rivalry in consumption, such as pastures (grazing land) or fish stocks. The over-exploitation of CPRs (or common access resources – CARs) can lead to the tragedy of the commons, such as overfishing and deforestation.
- As firms and individuals do not have to pay for the use of a CPR, it is inevitable that individuals and firms driven by self-interest will overuse the resource as they do not have to pay to use it. The tragedy is that market failure occurs as a result of there not being a market or price to preserve the resource.
- Non-excludability means it is not possible (or extremely difficult) to exclude people from using a resource even if they do not pay. Rivalrous means that one person's use of the resource diminishes the availability and benefits of the resource that is available to others in society.
- The lack of an effective pricing mechanism for CPRs results in overuse (or abuse), depletion and degradation of the resource because consumers do not pay for the use of the resources. Clearly, this is inefficient and poses a threat to the sustainability of CPRs as it becomes difficult to maintain or preserve the resources for future generations.
- There are numerous methods of government intervention in response to negative externalities and common pool resources, which can be remembered by the acronym PIGLETS CC: Pigouvian (indirect taxes), International agreements, Government provision, Legislation and regulation, Education (awareness creation), Tradable permits, Subsidies, Carbon taxes, and Collective self-governance.
- The government can internalize the costs of negative externalities by imposing an indirect tax (Pigouvian tax) equal to the external marginal cost of production or consumption, thus forcing producers and consumers to pay the full marginal social cost.

12 Market failure – public goods

Real-world issue

When are markets unable to satisfy important economic objectives – and does government intervention help?

Conceptual understandings (WISE ChoICES)

- The market mechanism may result in socially undesirable outcomes that do not achieve **efficiency**, environmental **sustainability** and/or **equity**.
- Resource overuse, resulting in challenges to environmental **sustainability**.
- Government **intervention** is effective, to varying degrees, in different real-world markets.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- public goods with reference to the characteristics (non-rivalrous and non-excludable) and the free rider problem (AO2)
- government intervention in response to public goods in terms of direct provision and contracting out to the private sector (AO3).

Public goods (AO2)

Public goods are collective consumption goods that have the two key characteristics of being **non-rivalrous** and non-excludable.

- **Non-rivalrous** – A person's consumption of a public good does not limit the benefits available to other people. The marginal cost of supplying a public good to an extra person is zero. If the public good is supplied to one individual, it is also available to everyone. By contrast, a person's consumption of a *private good* reduces the amount available for other individuals.
- **Non-excludable** – Firms cannot exclude people from the benefits of consumption, even if they do not pay. This means that non-payers can enjoy the benefits of consumption at no direct financial cost to them. By contrast, the consumption of private goods ultimately depends on an individual's ability and willingness to pay.

Examples of public goods include:

- emergency services (ambulance, fire and police)
- flood control systems, including dams
- free downloadable 'apps' for smartphones and tablet computers
- free streaming apps and websites
- knowledge and information systems, such as official government statistics
- law and order systems
- lighthouses
- national defence (national security)
- open access free Wi-Fi (such as in shopping malls and airports)
- open source software, such as Wikipedia, Khan Academy, YouTube, and creative commons websites
- public fireworks displays
- public radio broadcasts

◆ **Public goods** are collective consumption goods that have the two key characteristics of being non-rivalrous and non-excludable.

◆ **Non-rivalrous** is one of the two characteristics of pure public goods, where the consumption of the good or service by one person does not reduce the amount available for others

- public roads and motorways (highways)
- public water supplies
- road signs and traffic control systems (such as traffic lights)
- street lighting
- waste disposable systems, including drainage and sewage systems
- weather warning systems.

Figure 12.1 Lighthouses are an example of public goods

A good or service that has both characteristics (non-rivalrous and non-excludability) is known as a **pure public good**. Those with one of these characteristics only are known as **impure** or **quasi-public goods**. Some goods and services are non-rivalrous and non-excludable to some extent, such as public roads. There is rivalry to a degree (which is why many large cities suffer from heavily congested roads) and there is some degree of excludability (as drivers need a licence and money to own and maintain a car). Other examples include public beaches and parks. They are open to everyone, but they have a finite capacity – once they reach a certain capacity, no one else can enjoy the beach or park at that time.

By contrast, a private good is both excludable and rivalrous. For example, a smartphone is a private good as a supplier can exclude (prevent) those who are unwilling and/or unable to pay from using it. Also, once the smartphone has been purchased or consumed, it is not available for others. The same condition applies to all private goods, such as clothes, food, toys, books, nail polish, theatre tickets, restaurants, hotels and private car park spaces.

Common mistake

Students often claim that public goods are those provided by the public sector. This is an incorrect definition. Public goods have two key characteristics resulting in no incentives for any private sector firm to provide such goods and services. Hence, the government must do so if these goods are deemed to be desirable for society. However, this is a *consequence*, rather than the *definition*, of public goods.

Top tip!

While public goods exert positive externalities of consumption, not all goods that have positive externalities are public goods. For example, private healthcare and private education (both private goods) exert positive externalities of consumption.

ATL ACTIVITY 12.1

Discuss with a partner whether all public goods are examples of economic goods. Be prepared to share your answers with the rest of the class.

Free rider problem

Non-excludability means that it is not possible to prevent someone from using the public good even if they do not pay for it. Therefore, profit-maximizing firms in the private sector do not have any incentive to provide public goods. This results in public goods being under-provided, if at all, in a free market system. Paul A Samuelson (1954), the economist who coined the term public goods, defined a collective consumption good as products '*which (we) all enjoy in common in the sense that each individual's consumption of such a good leads to no subtractions from any other individual's consumption of that good*'. Hence, once public goods are provided, they can be consumed at zero marginal cost. Since there are no incentives for the private sector to provide public goods, such as street lighting or national defense, this represents a type or cause of market failure. Instead, public goods tend to be provided by the government.

This problem arises owing to the **free rider problem**, which occurs when people have access to (use or benefit from) a good or service without having to directly pay for it. For example, if a lighthouse is built, all sailors and fishing vessels in the area can benefit, without having to contribute any payment towards it. This happens because public goods are non-excludable. This means that, those who are unwilling and/or unable to pay cannot be excluded from using the good or service if it is provided.

Hence, the existence of the free rider problem prevents private sector firms from producing such goods and services as they are unable to charge users directly. In other words, the market demand for a public good does not actually exist. In addition, the supply of a public good will be significantly below the social optimum level – if it exists at all. Free riders take advantage of the public goods or services provided by the government without directly contributing to such provision.

ATL ACTIVITY 12.2

Using your knowledge of economics, convince a classmate why the following are *not* examples of public goods:

a private roads and tunnels with toll charges

b primary and secondary education

c public transportation services.

◆ The **free rider problem** occurs when people have access to (use or benefit from) a good or service without having to pay for it. As a result, the good or service will be underprovided or not provided at all in a free market.

Figure 12.2 As a private service, there is no such thing as a free taxi ride

Free-riding is regarded as an economic problem for several reasons:

- Free-riding leads to the under-production and/or over-consumption of a public good or service, so is highly inefficient. For example, most Wikipedia users do not pay to use the online information website yet are able to access the information freely.
- Over-consumption and over-exploitation of public goods can often lead to the *tragedy of the commons* (see below), resulting in their destruction or even extinction, such as public beaches being destroyed by the growing number of tourists.
- Free riders do not consider the external costs (negative externalities) of their activities, such as litter and plastic wastes left in public parks and on public beaches, possibly causing damage to the natural environment and ecosystems.

PAPER 1 EXAM PRACTICE QUESTION 12.1

Explain why merit goods and public goods are examples of market failure. *[10 marks]*

Common mistake

Students often confuse merit goods (see Chapter 11) with public goods. While it is true that merit goods and public goods both have positive externalities, public goods (such as national defence and street lighting) are highly unlikely to be provided by private sector firms owing to the free rider problem. By contrast, merit goods are excludable (such as education, healthcare and staff training) and are often provided by private sector firms.

Government intervention in response to public goods (AO3)

If the free rider problem cannot be addressed, valuable (public) goods and services will remain unproduced by the market economy. The non-rivalrous and non-excludability characteristics of public goods provide a strong case for the government to provide and pay for their provision, rather than relying on the market mechanism to do so. Government provision can help to prevent the under-provision and under-consumption of public goods, thereby improving social welfare. The government has two methods of intervention in response to public goods – it can directly provide these or pay a third-party contractor to do so. Whether direct provision or contracting out is used, the funding of public goods comes from the government.

Figure 12.3 Government intervention is needed to fund the emergency services

Common mistake

'Public goods are those provided by the public sector.'

This mistake is made too often by students. This is the *outcome* (or consequence) of the free rider problem facing public goods. Being non-rivalrous means one person's consumption of a product does not reduce the consumption of that product for others. Besides, there are other goods provided by the public sector that are *not* public goods, such as education, healthcare and public housing.

Direct provision occurs when the government directly provides the public goods, without having to use a third-party provider. For example, this could apply for the provision of law and order systems, national defence (national security), emergency services (ambulance, fire and police), and knowledge and information systems, such as official government statistics.

◆ **Direct provision** occurs when the government directly provides or supplies certain goods and services deemed to be in the best interest of the public.

Arguments in favour of direct provision of public goods include:

- public goods would simply not be available if it were not for government provision
- direct provision enables large-scale operations, so the government benefits from economies of scale and efficiency gains
- direct provision is generally seen to be favourable so would have the support of the general public as public goods improve the well-being of the individuals and societies.

However, the main arguments against direct provision are:

- the cost to finance such expenditure, which ultimately is passed on to taxpayers
- the opportunity cost of such provision, that is, could the money have been better spent on other priorities of the economy?
- potential government failure from intervening in markets because governments do not necessarily know what is best for individuals and the societies.

■ **Figure 12.4** Should the government spend more money to tackle the problems of potholes?

Alternatively, the government might choose to use **contracting out** to the private sector, where it pays a specialized third-party firm with the expertise to provide the public good. For example, this could apply for the provision of public fireworks displays (to celebrate major events such as New Year or National Day celebrations), highway maintenance services, waste disposable systems, including drainage and sewage systems, and refuse collection.

◆ **Contracting out** is an approach to government intervention in response to public goods by paying a specialized third-party firm with the expertise to provide a public good, such as refuse collection services.

The imperfections of the market economy in the provision, or lack of public goods need to be weighed against the imperfections of government intervention. Government intervention in markets is not always efficient, especially when there are bureaucratic processes involved and conflicting value judgements of policymakers (see Chapter 2). Direct provision and contracting out the provision of public goods can create their own problems.

In order to determine the optimal (best) allocation of public goods, the government typically calculates and compares the marginal benefits and marginal costs of provision (see Chapter 11). If the marginal social benefit (MSB) of provision is greater than the marginal social cost (MSC), then there is an under-allocation of a public good, from society's point of view. By contrast, if the MSC is greater than MSB, there is an over-allocation of the public good. Hence, the optimal quantity of a public good in an economy occurs where the demand (or marginal social benefit) is equal to the supply (or marginal social cost) curve, that is, MSB = MSC.

PAPER 1 EXAM PRACTICE QUESTION 12.2

Explain the social costs and social benefits of a highway (motorway) expansion project funded by the government. *[10 marks]*

Theory of Knowledge

How can individuals, households and firms know if government intervention is truly in their best interest?

WISE ChoICES

In the context of public goods and merit goods, to what extent do you think **intervention** improves the **economic well-being** of individuals and societies?

PAPER 2 AND 3 EXAM PRACTICE QUESTION 12.3

Explain why public goods are an example of economic goods rather than free goods. *[4 marks]*

Chapter summary

- A pure public good has two key characteristics: it is non-excludable and non-rivalrous.
- Non-excludable means that once a (public) good is provided it becomes available to everyone, including those who are unable and/or unwilling to pay for it.
- Non-rivalrous means the consumption of a (public) good by one individual does not diminish the availability of the good or service to other people.
- Public goods are not usually provided by private sector firms because they are unable to charge consumers (owing to the free rider problem as a result of the non-excludability nature of public goods).
- The free rider problem leads to public goods being under-provided or not provided at all, thus causing market failure in the production and consumption of public goods.
- If the free rider problem cannot be addressed, the government needs to intervene to provide and/or pay for the provision of public goods.
- The government provides an efficient quantity of a public good when its marginal social benefit (MSB) of provision equals its marginal social cost (MSC).

13 Market failure – asymmetric information (HL only)

Real-world issue

When are markets unable to satisfy important economic objectives – and does government intervention help?

Conceptual understandings (WISE ChoICES)

- The market mechanism may result in socially undesirable outcomes that do not achieve **efficiency**, environmental **sustainability** and/or **equity**.
- Resource overuse, resulting in challenges to environmental **sustainability**.
- Government **intervention** is effective, to varying degrees, in different real-world markets.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:
- asymmetric information, with reference to adverse selection and moral hazard (AO2)
- responses to asymmetric information with reference to government responses (legislation and regulation, and the provision of information) and private responses (signalling and screening) (AO3).

◆ **Asymmetric information** refers to missing, unbalanced or incorrect information that exists when one economic agent has more information than the other in an economic transaction.

Asymmetric information (AO2)

Accurate and unbiased information is essential for making rational and sensible economic decisions. When a market experiences imperfect information, it can lead to market inefficiencies and market failure. **Asymmetric information** is a source of market failure that exists when one economic agent (buyer or seller) has more information than the other in an economic transaction. It occurs owing to incomplete information or inaccessibility to information. Asymmetric information is common in markets such as life assurance policies, stock market products, pension fund schemes, second-hand cars, funeral services and works of art.

Figure 13.1 What information is needed to determine the value of art?

Economic theory suggests that markets are efficient in allocating resources when there is perfect knowledge (see Chapter 14). This means, the information is readily available and is evenly shared by all the economic agents in an economic transaction. The existence of asymmetric information creates a problem because one party can exploit the other with their greater

knowledge. Therefore, asymmetric (missing, unbalanced or incorrect) information is a significant cause of market failure in numerous markets because the party with access to better information has an unfair competitive advantage. This results in inefficiencies and a misallocation of resources because imperfect information results in distorted decision-making. In other words, consumers pay too much or too little, and firms produce too much or too little output. Hence, consumer and producer surplus (see Chapter 11) cannot be maximized. Essentially, asymmetric information exists if there is not perfect knowledge in a market.

Examples of asymmetric information in different markets include:

- Insurance policies – Consumers do not always understand the complexities of such policies, including the many terms and conditions associated with insurance claims.
- Financial investment products – Consumers are often unaware of the way stocks, shares, pensions, bonds and other investments work. They do not necessarily understand the risks involved or the fees and charges linked with managing their investment portfolios.
- Telecommunications service providers – Consumers may not know whether they are truly getting good value for money as there are so many service providers and monthly plans with different terms and conditions for mobile phone network and broadband internet service providers.
- Real estate – An estate agent may choose to exploit potential buyers by failing to reveal all the facts about a property, knowing that they have very little knowledge about the property.
- Job market – When applying for jobs, the applicant may not reveal (all) their areas of weakness or shortcomings, in an attempt to impress the potential employer. However, the applicant is unlikely to know the maximum salary that the employer is willing to offer in a salary negotiation.
- Second-hand cars – The seller may not reveal the full history of the car or any major problems or accidents in an attempt to sell the car for a higher price.
- Food manufacturers – Although there may be regulation on food labelling, food manufacturers are not obliged to reveal all of the ingredients. For example, what does 'Made with real fruit juice', 'Strawberry flavour' or 'Contains less sugar' really mean?

Figure 13.2 Do we know what really goes into our food?

Asymmetric information can lead to **opportunistic behaviour** in markets. This means that one party has an unfair advantage given the opportunity that the other party lacks information. Economists distinguish between two forms of opportunistic behaviour: adverse selection and moral hazard.

◆ **Opportunistic behaviour** means that one party can take advantage of the opportunity that the other party lacks information, resulting in adverse selection or moral hazard.

Adverse selection

Adverse selection refers to the undesired decisions or outcomes that occur when buyers and sellers have access to imperfect (asymmetric) information. This is because asymmetric information causes an imbalance of power between economic agents in a transaction. It distorts the process by which the price and quantity of goods or services are determined, so leads to a suboptimal selection or choice.

◆ **Adverse selection** is a form of opportunistic behaviour that refers to the undesired decisions or results that occur when buyers and sellers have access to asymmetric information. This distorts the process by which the price and quantity of goods or services are determined.

For example, in the health insurance industry, the people most likely to buy health insurance are those with underlying health issues, such as smokers and drinkers, and people of middle age upwards. Insurance companies know this information, of course, so are likely to raise the average price of insurance cover as these groups represent the highest level of risk for insurance claims and payouts. However, the adverse selection in this case is likely to risk pricing many healthy people out of the market for health insurance.

When economic agents are ignorant of certain aspects of an economic transaction, such as the quality of the good or service they are buying, they are compelled to make assumptions. For example, in the second-hand market for cars, the sale of a relatively new car may make buyers suspicious about the reasons for the sale of the car. The limited knowledge or asymmetric information means that potential buyers may think the worst and sense the car will have problems. Buyers may also assume that second-hand cars are of good quality if the price is relatively high and that cars are of low quality if the price is comparatively low.

Moral hazard

A **moral hazard** is a situation where a party protected from risk (owing to having superior information) behaves differently than if they were fully exposed to the risk. The individual or firm behaves in a less careful manner, taking unnecessary risks, because they are encouraged to act less carefully or because they believe that their lack of care will not be found out, thus imposing costs on the party that has inferior information. One party knows more about their intended actions than the other party. Thus, moral hazard is the result of asymmetric information in economic transactions.

◆ A **moral hazard** is a situation where a party protected from risk (owing to having superior information) behaves differently than if they were fully exposed to the risk. This potentially imposes costs on the party that has inferior information.

A lack of equal information causes economic imbalances that result in adverse selection and moral hazards. For example, people who drive company cars are covered by their employer's insurance so may be less careful driving the car than if they drove their own vehicle for which they have to pay. Hence, these drivers are more likely to take undue risks because the firm cannot observe all of the (careless and selfish) actions of the driver. Such moral hazards result in market failure because an individual or firm's pursuit of pure self-interest and the tendency to act less carefully (or rationally) leads to inefficient outcomes.

During the global financial crisis of 2008, many banks were controversially bailed out by the government. If banks know or expect there is a good probability of receiving emergency financial support from the government if a financial crisis should occur, bank managers and employees might be tempted to take increased and unwarranted risks.

Figure 13.3 Should banks be bailed out if they face bankruptcy?

A similar example is government provision of welfare benefits to the unemployed. Individuals who qualify for state welfare benefits are far less likely to look for employment than if government reduced or removed their benefits.

Responses to asymmetric information (AO3)

Economic theory suggests that the demand for information will increase if it is provided at a lower price or if it is available for free. In addition, promoting literacy and numeracy in schools helps to equip people with the skills to acquire knowledge and to have a greater awareness and appreciation of the value of information in making rational choices. However, firms are fully aware that asymmetric information can give them market power. Information overload can also be overwhelming for people, which can distort rational decision-making.

Asymmetric information as a source of market failure can be tackled by government intervention or by responses in the private sector. Given the inefficiencies caused by imperfect information, the government has a considerable role in reducing or removing information failures. Government responses to the problem of asymmetric information include legislation and regulation, and the provision of information. Private responses include signalling and screening.

Government responses

Governments can deal with asymmetric information as a type of market failure in a number of ways, including legislation and regulation, and the provision of information to support better, more informed decision-making.

- Legislation refers to laws stipulated by the government. For example, tobacco manufacturers in many countries are required to have health warnings on cigarette packets. Job applicants are required to declare if they have a criminal record. Regulation refers to the act of monitoring and controlling the activity of firms. For example, the UK's Advertising Standards Authority regulates the advertising industry and requires adverts to be 'legal, decent, honest and truthful'.
- **Provision of information** means the government provides additional information about goods and services, or requires firms to do so, in order to help buyers make more informed decisions. For example, governments can use public broadcasts and announcements to improve information such as informing smokers, gamblers and drinkers of the true costs of their decisions and actions. Food manufacturers are obliged to disclose nutritional information on food packaging.

◆ **Provision of information** means the government provides additional information about goods and services, or requires firms to do so, in order to help buyers make more informed decisions.

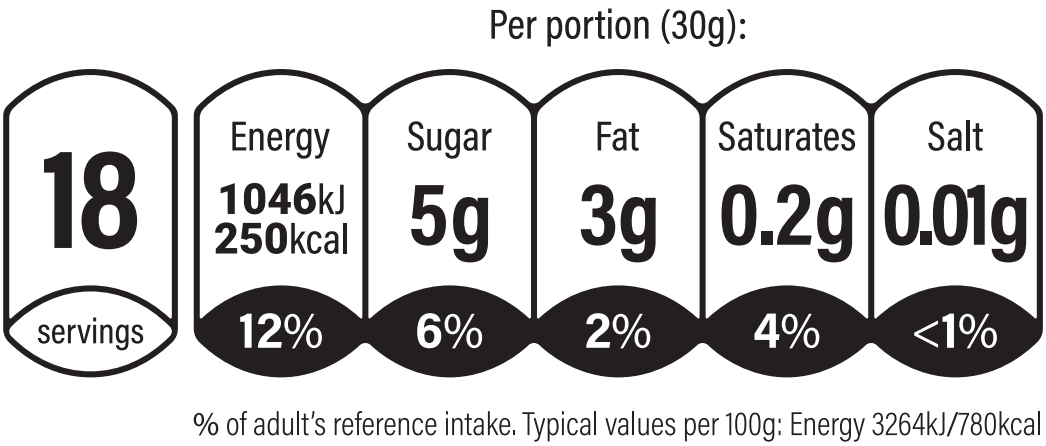


Figure 13.4 Nutritional facts on food labels is an example of provision of information

Regulation and legislation are required to tackle the potential problem of unethical marketing claims and asymmetric information. For example, cosmetics companies may provide incomplete, biased or misleading information about their beauty products. Similarly, tour operators might use brochures and adverts to make holiday resorts appear more attractive. Car manufacturers have also been known to exploit information failures (see Volkswagen scandal case study).

ATL ACTIVITY 13.1

Investigate government provision of information regarding the coronavirus pandemic. Evaluate the extent to which such provision of information was useful to the general public. Be prepared to share your findings with your teacher and peers.

CASE STUDY

The VW scandal

One of the worst cases of asymmetric information in the motor industry is the infamous Volkswagen scandal. The German carmaker had installed emissions software on more than 500,000 diesel cars in the USA and around 10.5 million more cars worldwide that allowed the vehicles to pass standards set by the likes of the USA's Environmental Protection Agency (EPA). In test mode, the VW vehicles were fully compliant with all of the EPA's emissions levels. However, when driving these cars normally, it was discovered that the emissions from VW vehicles were actually up to 40 times higher than the government's limit. Estimates vary, but the VW scandal cost the company around $35 billion in fines.

However, government responses are not always effective. For example, inefficient government regulation and legislation contribute to government failure. The *Telegraph* reported that 50,000 taxpayers in the UK missed the deadline for submitting their tax returns in 2019 (for which there is a fine of £100, or approximately $130). Improved provision of information and timely reminders to taxpayers could help to resolve this situation.

Top tip!

Note that signalling (as a response to asymmetric information) is different from the signalling function of the price mechanism (where price changes provide information to both producers and consumers about changes in market conditions).

WISE ChoICES

Using real-world examples and your knowledge of economics, discuss how it is possible for economists to determine whether government **intervention** leads to improved **economic well-being** or whether it creates more problems.

Private responses

Private responses to asymmetric information, which do not involve any form of government intervention, include signalling and screening.

Signalling

Signalling is a private response strategy undertaken by the economic agent with more or superior information in order to tackle the problem of adverse selection. Sellers strive to signal the benefits and quality of their products so that buyers can distinguish between products of high-quality or good value and low-quality or poor value products. Offering warranties (guarantees) and refunds, for example, can reassure potential customers considering purchasing the product.

◆ **Signalling** is a strategy undertaken by the party with more or superior information to tackle the problem of adverse selection with the aim of maximizing their own return or utility.

For example, in the market for classic cars, the owner is likely to have superior knowledge about the vehicle, its service history, the reliability, the condition of the engine and so forth. The potential buyer, by contrast, does not know the full knowledge or history of the vehicle. In order to sell the classic car, the seller may choose to use signalling. Buying a second-hand or classic car from a reputable car dealership that offers a warranty sends a more reassuring signal to potential buyers than buying from a one-off private individual. The advertising, promotions and physical existence of the car dealership signals to customers that the firm is genuine and has an incentive to sell quality classic cars.

■ **Figure 13.5** Sellers of classic cars use signalling to sell their vehicles

Theory of Knowledge

To what extent does possession of knowledge by a particular stakeholder group carry with it an ethical responsibility?

Screening

The existence of asymmetric information means the information available to buyers is different from the information available to sellers, so there is an incentive for one party to exploit the other in order to maximize their own return or utility. **Screening** is a private response strategy of tackling adverse selection by the economic agent with less or inferior information in order to minimize the decision-making complications caused by asymmetric information in the market. Buyers find ways to screen out biased, inaccurate, or misleading information in order to make better purchasing decisions. The notion of screening was first developed by Canadian American economist Michael Spence (1973), recipient of the Nobel Memorial Prize in Economic Sciences in 2001.

◆ **Screening** is a strategy undertaken by parties with less or inferior information to tackle the problem of adverse selection, with the aim of maximizing their own return or utility.

For example, the market for mobile (cell) phones and mobile networks is very complex, with different packages based on data usage, broadband connection speed, number of minutes available for telephone calls, and so forth. Essentially, the screener (the economic agent with less or inferior information) attempts to rectify this imbalance by learning as much as possible about the product. Screening is not required in markets for products that are simple and easily evaluated in terms of product quality and price. There is little need for this sort of response or behaviour in the market for fresh fruits and vegetables, for example.

Similarly, in the labour market, employers screen job applicants to determine their suitability and ability to do the job, giving preference to applicants who have better qualifications, skills and experience. This screening process helps to reduce the employer's risk of hiring someone not suitable for the job. Employers are also increasingly using screening processes to get feedback from their customers about the service they have received and their level of satisfaction from purchasing the product. Online review services (such as TripAdvisor, Google Reviews and Yelp) and customer reviews on social media (such as Facebook and Twitter) help to reduce misinformation for firms about the level of customer service and customer satisfaction.

■ **Figure 13.6** The market for smartphones and mobile network service plans is a very complex one

Many retailers also use screening to determine which products and prices to charge customers. For example, sales staff may try to find out as much as they can about customers and their needs or requirements in order to sell the most appropriate product to the consumer. Someone with a high willingness and ability to pay for a high-quality laptop, for example, may then choose to purchase a more expensive or superior model with better features such as a larger hard drive and faster processing ability.

Insurance companies use data to screen more careful drivers by offering 'no claims bonuses' to those who have a proven record of being of lower risk, such as those who have not been involved in traffic accidents or have been fined for speeding or reckless driving offences. Life insurance providers conduct detailed health screenings to assess the different risk profiles of their customers, enabling them to charge different insurance premiums. Dishonesty during the screening process can void the insurance policy, so this process helps tackle the problems of moral hazards.

Theory of Knowledge

Can asymmetric information ever be deemed to be ethical in the real world of economics?

PAPER 1 EXAM PRACTICE QUESTION 13.1

Explain why the existence of asymmetric information between buyers and sellers in a market is likely to result in market failure and inefficiencies. *[10 marks]*

Theory of Knowledge

To what extent does the English proverb 'Ignorance is bliss' apply to the study of economics and real-world issues?

WISE ChoICES

Some economists argue that the internet significantly reduced asymmetric information in markets, whereas others argue there is information overload, so people still make misinformed decisions (how often do you go beyond page 1 of a Google search?). Discuss the extent to which the internet has improved **choice** and **efficiency** in the economy.

Chapter summary

- Asymmetric information is a source of market failure that exists when one economic agent (buyer or seller) has more information than the other in an economic transaction.
- The existence of asymmetric information in a market results in market inefficiencies as consumer and producer surplus are not maximized.
- Asymmetric information can lead to opportunistic behaviour, that is, one party takes advantage of the opportunity that the other party lacks information.
- Adverse selection is undesired decisions or outcomes that occur when buyers or sellers have asymmetric information due to the imbalance of power.
- Imperfect information causes imbalances in decision-making that results in adverse selection and moral hazards, both of which are likely to result in market failure.
- Responses to asymmetric information include government responses (legislation, regulation and provision of information) and private responses (signalling and screening).

14 Market failure – market power (HL only)

Real-world issue

When are markets unable to satisfy important economic objectives and does government intervention help?

Conceptual understandings (WISE ChoICES)

- The market mechanism may result in socially undesirable outcomes that do not achieve **efficiency**, environmental **sustainability** and/or **equity**, that is, market failure resulting in allocative inefficiency and welfare loss.
- Resource overuse, resulting in challenges to environmental **sustainability**.
- Government **intervention** in dealing with market failure to improve **equity** and equality.

SYLLABUS CONTENT

By the end of the chapter, Higher Level students should be able to understand:

- imperfect competition – oligopoly and monopolistic competition (AO2)
- rational producer behaviour – profit maximization (AO2)
- the following from data: profit, MC, MR, AC and AR (AO4)
- degrees of market power and the meaning of market power in perfect and imperfect competition (AO3)
- the model of perfect competition: profit maximization in the short run and long run, and allocative efficiency (AO3)
- a perfect competition diagram showing the firm as a price taker, and the situation of abnormal profit, normal profit and losses, and being allocatively efficient (AO3)
- the model of monopoly: profit maximization, allocative inefficiency (market failure), welfare loss and natural monopoly (AO3)
- monopoly diagrams that shows: market power, abnormal profit, normal profit, losses, price/quantity comparison of a monopoly with perfect competition, welfare loss and natural monopoly (AO4)
- the model of oligopoly: collusive versus non-collusive, allocative inefficiency (market failure), game theory payoff matrix, price and non-price competition, and measurement of market concentration (AO3)
- a collusive oligopoly diagram (AO4)
- the model of monopolistic competition: profit maximization in the short run and long run, market power, allocative inefficiency (market failure) and product variety (AO3)
- a monopolistic competition diagram (with a more elastic demand curve compared to a monopoly) that shows: abnormal profit, normal profit and losses (AO4)
- the advantages of large firms having significant market power, including: economies of scale and abnormal profits (AO3)
- the risks in markets dominated by one or a few very large firms: risks in terms of output, price and consumer choice (AO3)
- government intervention in response to abuse of significant market power: legislation and regulation, government ownership and fines (AO3).

Perfect competition (AO2)

Economists aim to understand and predict the rational behaviour of firms in a market based on their common characteristics. One key influencing factor in the behaviour of a firm is the amount of market power that it possesses. **Market power** refers to the ability of a firm or consumer to manipulate price above or below a perfectly competitive level. The amount of market power that a firm has, determines what market structure it is allocated.

Market structure refers to the categorizing of firms in a particular industry, based on their level of market power, such as the number and size of firms, the nature of barriers to entry in the industry, and the degree and intensity of competition. If a firm is said to belong to a certain market structure, then economists can make particular assumptions relating to how that firm will behave in the market.

Factors that determine the amount of market power that a firm has, thus determining the market structure in which it operates, include:

- the number and size of firms that the business competes with
- the nature of barriers to entry to the industry, and
- the degree and intensity of price and non-price competition it engages with.

An understanding of the characteristics of various market structures is a very powerful tool in microeconomic analysis. For example, a consumer who knows how a firm is likely to behave to a change in competition can make better purchasing decisions. A firm that knows it has a large amount of market power can price its goods or services differently. Governments that understand the characteristics of a particular market structure can develop more effective economic policies.

Perfect competition is a market structure where there is an intensive degree of competition, with no individual firm being large enough to have any market power to influence the price or quantity traded. In theory, this market structure is regarded as an ideal by which economists can make comparisons. It predicts how individual firms and consumers will behave in a market where they have no market power.

Firms and consumers that are in perfect completion are called **price takers**. This means that they have no ability to influence output or price. For a firm to be in this position, there are several assumptions that must be met. The characteristics of perfect competition include the following interrelated attributes:

Many firms

Perfect competition assumes there are a very large number of buyers (customers) and sellers (firms) in the industry. This means, there is a high degree of competition in the market.

If there are a small number of firms in a market, then this can give them the opportunity to increase their price. Imagine if only two hardware shops existed in your city, located at opposite ends of town. Each vendor could be in a position to increase its price as the opportunity cost to the consumer of purchasing from the other vendor would be high. For example, the price increase may be less than the transport costs needed to go to a cheaper vendor or supplier.

This can also apply to a market with few consumers. If there are many firms and only one consumer, then the consumer is in a position to dictate the price to the firm. This situation can occur in the milk industry, where farms produce large quantities of milk to be sold to large supermarkets. If the supermarket offers a low price to these farms, it would be difficult for the farmers to sell thousands of litres of milk themselves. Therefore, the price is manipulated below the perfectly competitive

◆ **Market power** refers to the ability of a firm to manipulate the price of a product, usually above the perfectly competitive level.

◆ **Market structure** refers to categorizing of firms in a particular industry, based on their level of market power; for example, the number and size of firms, the nature of barriers to entry in the industry, and the degree and intensity of competition.

ATL ACTIVITY 14.1

With a partner, discuss how an understanding of market power might influence a Principal/Director of a fee-paying school or private tuition centre (regarding matters such as school tuition fees, uniform policy, curriculum, and so forth).

◆ **Perfect competition** is a market structure where there is intensive competition, with no individual firm being large enough to have any market power to influence the price or quantity traded.

◆ **Price takers** are firms that have no market power and are therefore unable to influence their price.

price. Hence, firms in perfect competition are price takers. This means no individual firm is large enough to have the market power to influence the equilibrium output or price.

The existence of many firms in the industry, none of which are large enough to have any market power, also implies there is **perfect information**. This means buyers and sellers have perfect knowledge of products and prices in perfectly competitive markets. This implies that buyers and sellers have equal and easy access to information about the product and prices. In perfectly competitive markets, there is no difference between what is known to the consumer and what is known to the firm. By contrast, **asymmetric information** (see Chapter 13) occurs when one party in a transaction is in possession of more information than another. For example, foreign tourists may be visiting somewhere for the first time, so they are unlikely to know what a fair price for a good or service will be. Locals, on the other hand, who have been exposed to prices in the market for a longer period of time, are less likely to pay a price that is above the equilibrium price in competitive markets.

◆ **Perfect information** means that buyers and sellers have equal and easy access to information about products and prices in perfectly competitive markets.

◆ **Asymmetric information** refers to missing, unbalanced or incorrect information that exists when one economic agent has more information than the other in an economic transaction.

■ **Figure 14.1** Farmers' markets are highly competitive, with many firms supplying similar products

Theory of Knowledge

Is there any real value to the model of perfect competition if, in the real world, there is no such thing as 'perfect knowledge'?

Barriers

Barriers to entry are non-existent in perfectly competitive markets, so there is freedom of entry into and exit from such industries. If other firms cannot enter the industry freely, then this will restrict competition. One factor that keeps prices in perfect competition at a perfectly competitive level is the fact that other firms can join the industry and take profits from incumbents. This means that if an incumbent firm increases its price, this may signal to other firms that they can operate more profitably by entering that market.

Likewise, the ability for a firm to exit a market freely (because firms in the industry may be making a loss) is also important in maintaining a perfectly competitive price. If there are only a small number of consumers in an industry, with market power, then they may manipulate price

below a perfectly competitive level. If firms cannot leave the market owing to barriers to exit, then they will continue to make a loss and the market will not be perfectly competitive. Hence, the ability of firms to easily exit an industry helps to keep the industry more robust and competitive.

Homogeneous products

A final characteristic of perfect competition is that all firms in the industry sell a homogeneous product. This means they supply products that are identical rather than differentiated. If firms sell homogeneous goods or services then the competition between them becomes more intense. If a good is differentiated from its competitor, such as in the way of size or quality, then it can command a higher price. Homogeneous goods should, in theory, not only relate to the good but also the consumption experience of the good. For example, a can of Coca-Cola sold at a convenience store should be very similar to a can of Coca-Cola in the minibar at a luxury hotel, but these items can be sold at vastly different prices because the experience of consumption is different.

In reality, there are no firms or industries that can fit these assumptions perfectly. A common example of perfect competition is market stall vendors selling fruits and vegetables (or *wet markets* as they are known in some countries in Asia Pacific), but even these fall short of classifying as perfectly competitive. For example, you can have several mango vendors in the wet market but the size of the mangoes may vary, therefore the price will differ. If there is any branding whatsoever then the mangoes are no longer homogeneous but heterogeneous. Furthermore, it is not possible to test perfect knowledge – imagine asking consumers walking out of a market whether they knew the price of the mangoes at every fruit stall in the market. Therefore, perfect competition should be seen as a representation of how competition can impact a market and change the behaviour of firms.

CASE STUDY

Fishy or perfect competition?

Sydney Fish Market (SFM) is a highly competitive market, and is the largest market of its kind in the Southern Hemisphere and the world's second largest seafood market in terms of variety outside of Japan. SFM sources products both nationally and internationally and trades more than 14,500 tonnes of seafood annually, with up to a hundred species traded every day. SFM employs around 60 staff to organize the weekday wholesale auctions, promote SFM as the centre of seafood excellence, and operate the Sydney Seafood School (SSS). Since opening in 1989, SSS has played an important part in persuading local residents to eat more fish, which is deemed to have health benefits. The price of fish products fluctuates according to the market forces of demand and supply. For example, seasonal fluctuations in demand and supply cause market prices to change accordingly.

Figure 14.2 Sydney Fish Market is a highly competitive market

Source: TripAdvisor, 2020

Top tip!

Where appropriate, try to use real-world examples in your written answers. For perfect competition, a real-world example could be markets that sell fresh fruit and vegetables, where the demand is highly price elastic because firms sell homogeneous products (carrots or oranges from one vendor are no different from those sold by other firms in the same market). No individual firm is large enough to influence market supply or market price. This means firms are price takers. It is also relatively easy for firms to enter or leave this market.

Monopoly (AO2)

A **monopoly** is a market structure where there is a single supplier of a particular good or service, thus having the power to influence the market supply and price. This means the monopolist has significant market power as it controls enough of the market supply to be able to manipulate market prices. Hence, the monopolist is a **price maker** (or **price setter**), with the ability to charge higher prices for its products. As seen in Chapter 3, the *law of demand* is held by both the income and the substitution effects. In the case of a monopoly, the substitution effect is not present, as it is the only firm in the market.

As with perfect competition, there are several assumptions that characterize a monopoly.

◆ A **monopoly** is a market structure where there is a single supplier of a particular good or service, thus having the power to influence the market supply and price.

◆ A **price maker** (or **price setter**) is a single or dominant firm that has significant market power enabling it to manipulate its prices as it has significant control over market supply.

A single or dominant firm

The term monopoly can mean different things to different people. A **pure** monopoly exists when there is only one firm in the market. For example, in the USA, this might include the United States Postal Service (the only service provider of first-class postage). However, the USPS may have a monopoly in the postage of small letters but may face considerable competition in the delivery of larger items from private courier services such as FedEx, TNT, DHL and UPS. Therefore, some economists argue that a monopoly is a dominant firm with significant power in a market.

Top tip!

Unless otherwise required or stated in the exam, it is advisable to stick with the definition of pure monopoly when referring to monopolists.

High barriers to entry

It is essential for monopolists to enjoy high barriers to entry. This prevents new firms trying to enter the industry in an attempt to compete for market share. A convenient way of remembering these entry barriers is through the acronym **BLADE**.

Branding – A monopoly, especially one that has operated for a long period of time, has such a large amount of brand awareness that if a new firm was to enter the industry, it would be very challenging and expensive for it to advertise enough to compete with the incumbent's brand. Well-established market leaders also enjoy customer and brand loyalty, creating further barriers to potential entrants.

Legal barriers – There can sometimes be legal reasons why firms cannot enter a market. These can be licences or health and safety requirements, for example. In some cases, the government grants the right for only one firm to operate in the market, as in the case of *natural monopolists* (see page 199) or legal monopolies. Monopolists can also create artificial barriers to entry by protecting their intellectual property rights through the use of copyrights, trademarks and patents.

Anti-competitive practices – Some monopolies may take measures or engage in practices that prevent other firms from entering their market. For example, the use of *limit pricing strategy* means a monopolist deliberately sets the price of their good or service below the price at which it would be possible for a new entrant to sell it. This reduces the contestability of the monopoly's market. However, such anti-competitive behaviour can be deemed to be illegal in many countries.

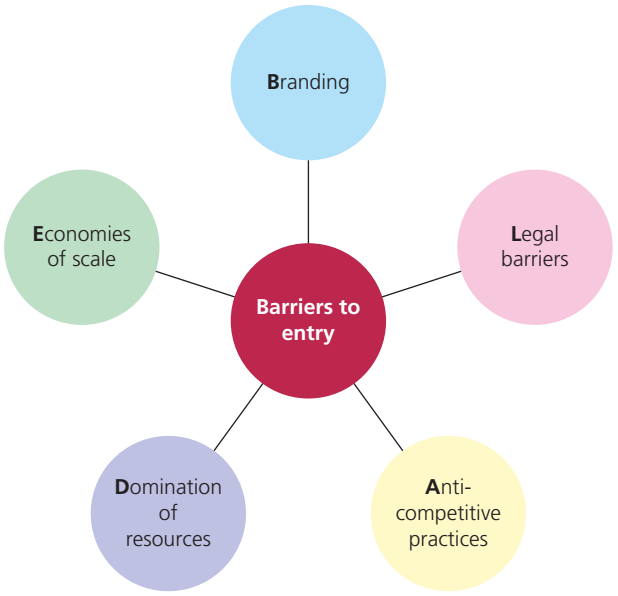


Figure 14.3 Barriers to entry in monopoly markets

Domination of resources – There are certain situations where it may not be possible for a new firm to enter a market as all the resources that are required are currently being used by the monopolist. An example of this is schools or universities, where there may not be a suitable location in a particular area where construction can take place. This means all land may be already in use by other enterprises and for other purposes. The same applies to gold or diamond mining in some countries; all areas that are rich in gold and diamonds have already been mined or are under the ownership of a dominant firm already, such as De Beers – the world's largest diamond exploration company.

Economies of scale – These are the cost-saving benefits of lower average costs of production brought about by an increase in the volume of production. There may not be enough demand in the market for a new firm to achieve the necessary economies of scale to compete. This will mean any potential entrants will have higher costs and not be able to compete with the more efficient monopoly. This will prevent a firm from joining the market in the first place. The monopolist can also create barriers to entry by investing in the latest innovations and technologies. By exploiting technical economies of scale, the higher set-up costs can deter firms from trying to enter the market.

No close substitutes

The lack of availability of close substitutes means the monopolist can enjoy its market position. Firms that are able to supply a unique good or service can charge higher prices and control market supply. For example, the demand for veterinary services is highly price inelastic owing to the high willingness to pay and because there are no close substitutes for such services (pet owners will fully understand why). Monopolists enjoy price inelastic demand for their product. Hence, by charging a higher price they can earn higher sales revenue owing to the lack of close substitutes. If you attend a fee-paying school, there are good (economic) reasons why school fees tend to increase each year …

CASE STUDY

Monopolists really do exist

In reality, a monopoly exists in a market dominated by one firm with significant market power. For example, firms with monopoly power include:

- Coca-Cola – the world's largest carbonated soft drinks producer, with more than 1.8 billion servings per day sold in over 200 countries.
- YKK – this Japanese company founded in 1934 is the world's largest manufacturer of zip fasteners, with production facilities in more than 70 countries.
- PayPal – the largest online payments platform has around 90 per cent market share in the USA, despite competition from rivals such as Apple Pay in the mobile apps market.
- Google – the largest internet search engine, which also owns YouTube, enjoys more than 60 per cent of global online advertising revenue.
- Tetra Pak – the largest food packaging company in the world, with operations in more than 160 countries, is infamous for its patented tetrahedron-shaped plastic-coated paper cartons.
- Mabuchi – this Japanese manufacturing company makes around 90 per cent of all micro-electric motors used to adjust rear-view mirrors in cars.

Common mistake

'Monopolists are harmful for the economy as they are able to charge whatever price they want because they are the only supplier in the market.'

It is incorrect to claim that monopolists can charge 'whatever' price they want because they are the single supplier of a good or service. While monopolists have the ability to control market supply, they cannot control the level of market demand. If prices rise too high, customers will switch to or seek alternatives. Hence, monopolists must lower prices if they want to sell more.

Imperfect competition (AO2)

In reality, there are few examples that can be given for perfect competition and monopoly. Recall that perfect competition is more of a representation of how a market *should* run under extremely competitive circumstances. Many firms compete intensely with other firms but are not labelled as price takers, so they could be described as operating in conditions somewhere between the two extremes of perfect competition and monopoly. Such firms are said to operate in imperfectly competitive markets.

In particular, this section of the IB Economics syllabus considers the market structures of oligopoly and monopolistic competition as forms of imperfect competition.

Oligopoly

An **oligopoly** is a market structure where a few large firms dominate the industry. Competition is very intense owing to the large market power of these dominant firms, rather than because of the number of firms in the industry. It is generally assumed that oligopolistic firms do not collude together.

◆ An **oligopoly** is a market structure where a few large firms dominate the industry.

There are three main characteristics of **non-collusive oligopoly**:

◆ **Non-collusive oligopoly** is a form of oligopolistic market in which firms act independently, rather than colluding to act as a monopolist.

1 **A few large firms**

Irrespective of the number of firms in the industry, an oligopolistic market is dominated by a few large firms. For example, there may be hundreds of small, independent supermarkets operating under a variety of names in a country, but these are likely to be dominated by a few large firms. For example, the US supermarket industry is dominated by the likes of Walmart, Costco and Publix. The market for commercial aircraft manufacturers is dominated by Airbus and Boeing. The world's largest smartphone producers include Samsung, Huawei, Apple and Xiaomi. Dominance is measured by using concentration ratios (see page 206).

Figure 14.4 The market for smartphones is dominated by a few large firms

Top tip!

In an oligopolistic market, it is the market share of the dominant firms that matters, rather than how many firms exist in the industry.

2 High barriers to entry

High barriers to entry make it very difficult for potential entrants to compete at the same scale as the large and well-established firms that dominate the market. As with monopoly, the BLADE acronym (branding, legal barriers, anti-competitive practices, dominance of resources and economies of scale) refers to the barriers to entry that are present in oligopolistic markets. These barriers also prevent prices from falling in oligopolistic markets.

3 Interdependence

Unlike in monopoly and perfect competition, firms in oligopoly are interdependent. In a market consisting of a few large firms, it is possible for an oligopolist to clearly identify its competitors. Hence, the actions of one dominant firm will impact the rest of the firms in the industry. For example, if American Airlines decides to cut prices to entice customers, it is likely that competitors such as Delta Air Lines, Southwest Airlines and United Airlines will also reduce their prices. This suggests there are no direct gains for a firm in reducing its prices. Similarly, if Sinopec decided to raise its price, other large petroleum companies like Chevron or ExxonMobil might choose not to raise their prices in order to entice customers away from Sinopec. Hence, interdependence causes **price rigidity** in non-collusive oligopolistic markets, although all dominant firms take action depending on the decisions made by their rivals.

◆ **Price rigidity** refers to the tendency of prices to remain unchanged in a non-collusive oligopoly. If a firm lowers its price, rival firms feel compelled to lower their prices, or face a loss of revenue. If a firm increases its price, then rival firms will keep their prices unchanged and take some of that firm's customers. Hence, there is a tendency for prices to stay the same.

Owing to the nature of interdependency and rigidity in oligopolistic markets, firms have to look for other ways to compete with each other. This is referred to as **non-price competition**, which involves focusing on product differentiation and advertising to try to gain a competitive edge against other firms. So, if American Multi-Cinema (AMC) introduces special promotional offers such as customer loyalty schemes to attract cinemagoers to its theatres, so will the likes of Regal Entertainment and Cinemark Theaters.

ATL ACTIVITY 14.2

Have a go at this Think-Pair-Share activity:

How do airline companies avoid price rigidity, i.e. why do airline companies charge different prices on different days for the same flight routes and destinations?

Figure 14.5 Airline companies are experts at pricing decisions

Monopolistic competition

Monopolistic competition is a market structure in which many firms exist, but each firm has only a small degree of market power. This is due to free entry (low barriers to entry) and product differentiation of competitors in the industry. The main characteristics of monopolistic competition are that there are many firms in the industry, free entry (low entry barriers) and differentiated products.

◆ **Monopolistic competition** is a market structure in which many firms exist, but each firm has only a small degree of market power.

Many firms

As with perfect competition, there are many firms in a monopolistically competitive market. However, as they do not produce or sell homogeneous products, each firm has a small degree of market power. By supplying differentiated products, monopolistically competitive firms have the ability to determine prices, although this is only to a small extent owing to the large number of firms in the market. Examples of such markets include local hairdressers, fast food outlets and restaurants, shoe repairers and key makers, and clothes retailers. For instance, a local hair salon can theoretically differentiate its services and set its own price, but of course there may be many other hairdressers that a customer can go to should the price be too high.

Figure 14.6 Hairdressers operate in monopolistically competitive markets

Free entry

The model of monopolistic competition assumes that there is an absence of barriers to entry and exit from the industry that is akin to the model of perfect competition. Any firm is free to set up in the market and compete with the existing firms. If there are more profitable industries to operate in, firms are free to leave the industry without incurring any major costs.

It is often said that firms in monopolistic competition are like mini-monopolies (or local monopolies) as they often operate separately on one street or in one small suburb. For example, in any particular suburban area, a local hairdresser, Chinese restaurant or beauty salon may be the only ones in that specific neighbourhood.

Top tip!

Free entry does not mean that there are no set-up costs to set up a business – as this would not be realistic for any business. However, many small businesses operate under conditions of monopolistic competition because entry to the market is relatively straightforward.

Product differentiation

Firms in monopolistic competition have the ability to differentiate their goods and services. In the case of hairdressers, for example, the brand name is an example of product differentiation. Some hairdressers may specialize in male haircuts, others may even go to great lengths to accommodate children. It is mainly the product differentiation that accounts for the differences in price in monopolistically competitive markets, rather than the ability to exploit consumer surplus as in the case of monopoly markets. Similarly, restaurants offer something different (such as cuisines from different regions of the world) and have some degree of uniqueness, but they are all essentially competing for the same customers in the local area.

Top tip!

Monopolistic competition is closer to perfect competition than oligopoly owing to the large number of firms in the market.

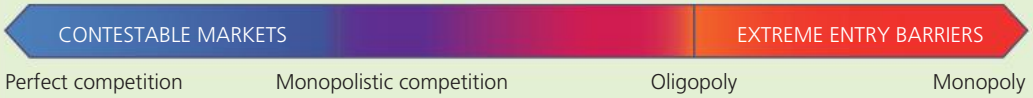


Figure 14.7 Market structures summary

Rational producer behaviour – profit maximization (AO2, AO4)

Economics attempts to solve the basic economic problem of scarcity (see Chapter 1). To achieve this, economists are concerned about the efficiency and viability of one economic activity relative to the next best alternative. In economics, it is assumed that producers act rationally by seeking to maximize profit. Profit is the positive difference between total revenue and total costs of production, and acts as an incentive or reward for entrepreneurship.

Figure 14.8 Profit is the difference between revenue and costs

Total revenue and total costs (AO2)

Revenue is the money received from the sale of a firm's output of goods or services. **Total revenue** (TR) is the overall amount of money received by a firm from selling its entire output. For example, Nike and Adidas receive most of their revenues from the sale of sports apparel and sports equipment. Total revenue is calculated using the following formula, where P = Price and Q = Quantity sold:

$$TR = P \times Q$$

Average revenue (AR) refers to the average price received from the sale of a good or service. It is calculated by using the formula:

$$AR = \frac{TR}{Q}$$

Average revenue is mathematically the same as the price per unit. This is because:

$$AR = \frac{TR}{Q} = \frac{(P \times Q)}{Q} = P$$

For example, if a cinema earns $60,000 from the sale of 7,500 tickets, the average revenue (or average price) of each cinema ticket would be:

$$AR = \frac{\$60{,}000}{7{,}500}$$

$$AR = \$8 \text{ per cinema ticket}$$

Costs are the expenses or expenditures of a firm in relation to their production. There are various categories of costs, all of which are important for understanding profit maximization.

◆ **Revenue** is the money received from the sale of a firm's output of goods and services.

◆ **Total revenue** (TR) is the value of money received by a firm from selling its output of goods and/or services.

◆ **Average revenue** (AR) refers to the price received from the sale of a good or service.

Fixed costs are expenses that do not change with the level of output, such as rent, insurance premiums, management salaries, loan interest repayments and advertising costs. By contrast, **variable costs** are those directly linked to the level of output. Hence, they continually rise if output increases, such as wages and raw material costs.

Fixed costs can be further categorized as **total fixed costs** (TFC) and **average fixed costs** (AFC). TFC refers to the sum of production costs that do not change with the level of output. AFC refers to the fixed costs per unit of output. AFC and TFC are calculated using the following formulae:

$$AFC = \frac{TFC}{Q}$$

$$TFC = AFC \times Q$$

Similarly, variable costs can be categorized as **total variable costs** (TVC) and **average variable costs** (AVC). TVC are production costs incurred directly from the output of a particular good or service. For example, variable costs for Emirates Airlines include fuel and in-flight meals. AVC refers to direct production costs incurred for each unit of output of a good or service. TVC and AVC are calculated using the following formulae:

$$AVC = \frac{TVC}{Q}$$

$$TVC = AVC \times Q$$

Total costs (TC) are the aggregate amount of production costs spent on the output of a particular good or service. TC is calculated using the formula:

$$TC = TFC + TVC$$

Average costs are the unit costs of production. AC is calculated by dividing the total costs of production (TC) by the output level (Q):

$$AC = \frac{TC}{Q}$$

Profit is calculated by measuring the difference between a firm's total revenue (TR) and total costs (TC) of production. Hence, profit maximization occurs when the positive difference between TR and TC is at its greatest. The formula for calculating profit is:

$$Profit = TR - TC$$

PAPER 3 EXAM PRACTICE QUESTION 14.1 (HL ONLY)

The table below shows the costs and revenues of Tandy Toys Ltd when operating at 5,000 units of output per month:

Item	Cost/Revenue ($)
Price	20
Raw materials per unit	8
Rent	7,000
Salaries	8,000

a Calculate the total cost of producing 5,000 units. *[2 marks]*

b Calculate the profit made by Tandy Toys Ltd if its entire output is sold. *[2 marks]*

Common mistake

Students often state that fixed costs are expenses that do not change. This is inaccurate, as fixed costs can and do change; for example, rents and salaries tend to increase over time. The definition is that fixed costs remain the same irrespective of the level of output.

◆ **Total cost** is the sum of fixed and variable costs of producing a good or for providing a service.

Top tip!

Average costs (AC) are also confusingly known as average total costs (ATC). It is acceptable for you to use these terms interchangeably.

Common mistake

In everyday language, it is common for people to use the word *cost* when referring to *price*. Economists refer to **cost** in terms of the cost of production, whereas **price** is the value paid by customers for the purchase of a good or service.

Marginal cost and marginal revenue (AO2)

Marginal revenue is the additional revenue received from the sale of an extra unit of output. It is calculated by dividing the change (Δ) in total revenue by the change (Δ) in the quantity or level of output:

$$MR = \frac{\Delta TR}{\Delta Q}$$

Marginal cost is the additional cost of producing an extra unit of output. It is calculated by dividing the change (Δ) in total costs by the change (Δ) in the quantity or level of output:

$$MC = \frac{\Delta TC}{\Delta Q}$$

Top tip!

Why do staff at the cinema allow you to buy a ticket after the movie has started while an airline would never wait for you if you called to say you were late for the flight? The reason is to do with the marginal cost of the decision. The costs of screening the movie have already been incurred; there is literally no marginal cost in allowing someone late to buy a movie ticket (if there is a spare seat). For the airline, the marginal cost would be huge as the company has to pay a penalty to the airport authorities and compensation to all the other passengers for delaying their flight.

Profit is maximized when the marginal cost (MC) of production equals the marginal revenue (MR) from selling that unit of output. Hence, the condition for profit maximization is:

Profit maximization = MC = MR

This is because:

- If MR > MC, additional revenue exceeds additional costs, so the profit maximizing firm should increase its output.
- If MC > MR, additional costs exceed additional revenue, so the profit maximizing firm should reduce its output.
- Hence, when MC = MR there is no change in profits, that is, this is the level of output for profit maximization.

PAPER 3 EXAM PRACTICE QUESTION 14.2 (HL ONLY)

Calculate the following costs from the data below. Express your answers to 2 decimal places.

Output (units)	Variable cost ($)	Total costs ($)
10	2,000	4,545
15	2,850	5,395

a fixed costs of production *[2 marks]*

b average fixed costs at 10 units and 15 units of output *[2 marks]*

c average variable costs at 10 units and 15 units of output *[2 marks]*

d average costs at 10 units and 15 units of output *[2 marks]*

e the marginal cost of production. *[2 marks]*

◆ **Fixed costs**, such as insurance premiums and advertising costs, are expenses that do not change with the level of output.

◆ **Average fixed costs** (AFC) are the production costs that do not change with the level of output.

◆ **Total variable costs** (TVC) are the overall production costs incurred directly from the output of a particular good or service. TVC continually rise with higher levels of output, such as labour costs and raw material costs.

◆ **Average variable costs** (AVC) are production costs incurred directly per unit of output for a particular good or service.

◆ **Average costs** (AC) are the unit costs of production, that is, the cost of producing one unit of output.

◆ **Profit** is the positive difference between a firm's total revenues (TR) and its total costs (TC). It is calculated using the formula: TR – TC.

◆ **Marginal cost** (MC) is the cost of producing an extra unit of output. It is calculated by dividing the change in total costs (ΔTC) by the change in the level of output (ΔQ), that is, MC = ΔTC ÷ ΔQ.

◆ **Marginal revenue** (MR) is the extra revenue received from the sale of an extra unit of output, that is, MR = ΔTR ÷ ΔQ.

PAPER 3 EXAM PRACTICE QUESTION 14.3 (HL ONLY)

a Using the diagram (Figure 14.9) below, calculate the total revenue at a price of $16. *[2 marks]*

b State the average revenue when 120 units are sold. *[1 mark]*

c Outline what the difference between area A and area B shows. *[2 marks]*

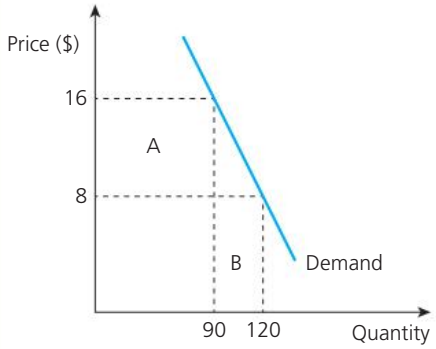


Figure 14.9 Calculating revenues of a firm

Abnormal profit (AO2, AO4)

Economists view costs in a very different way from accountants. For economists, it is important to include opportunity cost (see Chapter 1). **Economic costs** are the explicit and implicit costs of all resources used by a firm in the production process. **Explicit costs** are the identifiable and, therefore, accountable costs related to the output of a product, such as wages, raw material costs, utility bills and rent. **Implicit costs** are the opportunity costs of the output, that is, the income from the best alternative that is forgone. For example, a student might give up the opportunity to work in a job that pays $25,000 per year and choose to study at university which costs her $35,000 a year. The economic cost to the student is therefore $35,000 + $25,000 = $60,000 per year.

Abnormal profit (also called **supernormal profit** or **economic profit**) refers to the profit of a firm that is over and above its normal profit (when total costs of production equal total revenues from production). Hence, abnormal profit is profit in excess of the amount needed to generate a return on the firm's capital investment. It creates incentives for existing firms to produce and for potential rivals to enter the market. Abnormal profit occurs when AR > AC. A firm is said to earn economic profit from a particular economic activity only when it experiences a profit greater than from its next best alternative.

◆ **Abnormal profit** refers to profit that is greater than normal profit, thereby creating incentives for producers to increase output.

◆ **Normal profit** exists when a firm earns just enough revenue to cover its total costs of production and remain operational in the industry, that is, there is zero economic profit.

Normal profit (AO2, AO4)

Normal profit is the amount of profit needed to generate production of a good or provision of a service. Unlike accountants, economists regard normal profit as a cost of production because without this there is no incentive for the good or service to be provided. The condition for normal profit is AR = AC.

If a firm earns only normal profits, then it will just cover all production costs (both economic and implicit costs) from its total revenue. In other words, there is zero economic profit. For example, if a worker currently earns $45,000 and another potential employer wanted to hire this worker, the minimum salary offered must be at least $45,000 to entice the worker, ceteris paribus. Hence, normal profit is sufficient to keep the firm in business.

Top tip!

If a firm earns zero economic profit, this means it earns normal profit but does not earn any supernormal or abnormal profit. Hence, zero economic profit does not mean the firm earns zero profit.

PAPER 2 EXAM PRACTICE QUESTION 14.4 (HL ONLY)

Outline why a firm might choose to continue operating even if it earns zero economic profit. *[2 marks]*

Losses (AO2, AO4)

A firms will make a **loss** (negative profit) if its production costs exceed its total revenue, that is, TC > TR. This is because the firm's price (average revenue) is not sufficient to cover its unit costs of production (average cost). Hence, losses occur when AR < AC.

While it is possible for a loss-making firm to remain operational in the short run, this is not sustainable. Firms need to pay their suppliers, workers and financiers so must generate surplus revenues in order to continue to be operational in the long run.

◆ A **loss** occurs when a firm experiences negative economic profit, that is, its total costs of production exceed its total revenue: TC > TR.

CASE STUDY

Economic impacts of the coronavirus pandemic

The coronavirus (Covid-19) pandemic caused businesses around the globe to collapse. The outbreak of the deadly virus across the world caused countless businesses to fail and mass job losses. Major stock markets and businesses in almost every industry experienced their largest losses since the global financial crisis in 2008. This pandemic clearly showed that free markets are often unable to satisfy important economic objectives, such as employment in the economy, so requires direct government intervention, such as financial support for the unemployed and tax concessions for businesses.

Figure 14.10 The coronavirus pandemic caused major problems for all businesses

Common mistake

Students often incorrectly use the terms 'cost' and 'price' interchangeably. Remember, **costs** (of production) are paid for by firms in the production process whereas **price** is paid by the customer to purchase a good or service.

Table 14.1 Summary of costs and revenues formulae

Type of cost or revenue	Formula	Annotation
Average cost (AC)	Total cost ÷ Quantity produced	AC = TC ÷ Q
Average fixed cost (AFC)	Total fixed cost ÷ Quantity produced	AFC = TFC ÷ Q
Average variable cost (AVC)	Total variable cost ÷ Quantity produced	AVC = TVC ÷ Q
Average revenue (AR) or Price (P)	Total revenue ÷ Quantity traded	AR = TR ÷ Q
Marginal cost (MC)	Change in total cost ÷ Change in output level	ΔTC ÷ ΔQ
Marginal revenue (MR)	Change in total revenue ÷ Change in output level	ΔTR ÷ ΔQ
Total cost (TC)	Total fixed cost + Total variable cost	TC = TFC + TVC
Total revenue (TR)	Unit price × Quantity traded	TR = P × Q

Degrees of market power (AO3, AO4)

Market power refers to the ability of a firm to manipulate the price of a good or service by manipulating the level of supply and/or demand for the product. The degree of market power that a firm has determines what market structure it operates in.

Recall that **perfect competition** is a market structure where there is a very high degree of competition, with no individual firm being large enough to have any market power to influence the price or quantity traded. As firms in perfect competition do not have any market power, owing to the intensity of competition in the market, they are often referred to as **price takers** (see Figure 14.11).

In Figure 14.11, it can be seen that the firm's selling price ($10) is determined by the market forces of demand and supply, creating a market clearing price of $10. The price is not determined by any individual firm owing to the absence of market power in perfectly competitive markets. Therefore, it is only changes in market demand and/or supply that will cause a change in the price charged by perfectly competitive firms.

◆ **Market power** refers to the ability of a firm to manipulate the price of a good or service by manipulating the level of supply and/or demand for the product.

◆ **Perfect competition** is a market structure where there is intensive competition, with no individual firm being large enough to have any market power to influence the price or quantity traded.

For perfectly competitive firms, the average revenue curve (AR) and the marginal revenue curve (MR) are equal (as seen in Figure 14.11). As perfectly competitive firms are price takers, every time they sell a good or service (as there is demand for the product), they receive the same price for it. This means that the amount of revenue gained from each sale is equal to the last, which is another way of saying that MR remains constant, so MR = AR. Also, recall that AR = P, which means that P = AR = MR.

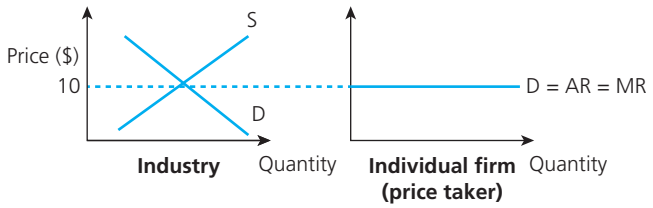


Figure 14.11 Perfectly competitive firms as price takers

Table 14.2 shows data for a firm that supplies cups of lemonade for $5 each. As the firm is a price taker, the price of each cup of lemonade will remain at $5. It can be seen that total revenue (TR) increases by increments of $5 meaning that each time a cup of lemonade is sold, revenue increases by $5, which is the price of that drink. This means that P = MR, so AR = MR = P, as demonstrated in the table.

Table 14.2 Data for a perfectly competitive firm

Quantity	1	2	3
Price	5	5	5
TR	5	10	15
AR	5	5	5
MR	5	5	5

It is assumed that firms in perfect competition will attempt to maximize their profit (or minimize their loss). Recall that profit maximization occurs when a firm's marginal cost equals its marginal revenue, that is, MC = MR.

It can be seen in Figure 14.12 that profit maximization occurs at an output of Q* (the output level where MC = MR). If the firm produces less than Q*, while it may still earn a profit, the firm would have forgone the sale of goods or services that had a greater marginal revenue than marginal cost. If the firm were to produce at an output greater than Q*, it would be producing units where the marginal cost of production would exceed the marginal revenue gained. This would result in a loss on these units. Therefore, MC = MR is the point at which all profitable units have been traded.

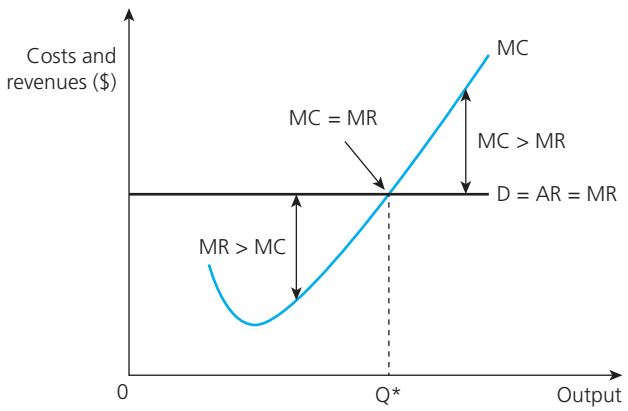


Figure 14.12 Profit maximizing position for a price taker

Top tip!

Understanding the relationship between AC and MC

To help you better understand the diagrams used in this topic, take a look at this section which explains how the average cost (AC) and marginal cost (MC) curves are constructed. The relationship between AC and MC is applicable to all market structures (perfect competition, monopolistic competition, oligopoly and monopoly).

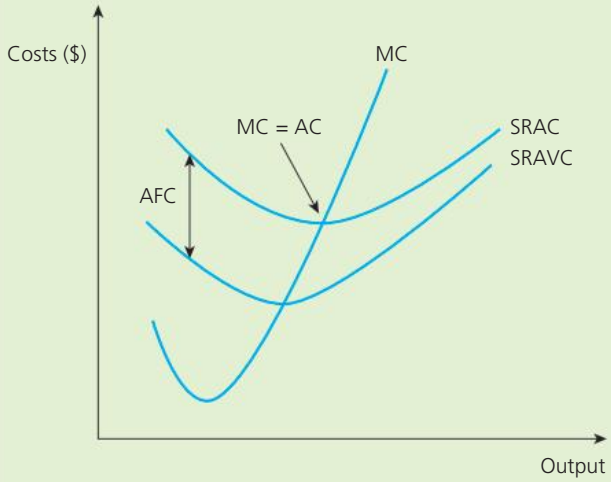


Figure 14.13 The relationship between AC and MC

- The downward sloping section of the **SRAC** curve is caused by MC being lower than AC.
- AC will reach its minimum point when MC = AC, that is, at the minimum point on the SRAC curve.
- When MC > AC, the SRAC curve will begin to rise owing to the law of diminishing marginal returns.
- In the short run, average fixed costs (AFC) will continually fall with increased levels of output because the fixed costs are spread over an increasingly larger volume of output. For example, assume fixed costs are $5,000.
- At 1,000 units of output, the AFC = (5,000 ÷ 1,000) = $5
- At 2,000 units of output, the AFC = (5,000 ÷ 2,000) = $2.5
- This relationship is shown by the short run average variable cost (SRAVC) curve. The vertical distance between SRAC and SRAVC is the AFC at each output level. Notice that this is greater at a lower level of output than at larger levels of output.

ATL ACTIVITY 14.3

It can be seen that for price takers, the demand curve is perfectly price elastic. This means that there is no consumer surplus for the good that the firm is selling.

This activity promotes discussion and reflection about one of the key characteristics of perfect competition.

With a partner, discuss the reasons why consumers in perfect competition have no consumer surplus.

◆ The **short run average cost** (SRAC) represents the cost per unit at different levels of output when at least one factor of production is fixed, such as capital resources. It is equal to the average fixed cost (AFC) plus the short run average variable cost (SRAVC).

Perfect competition – profit maximization in the short run

In the short run, it is possible for a perfectly competitive firm to earn abnormal profit (see Figure 14.14), negative economic profit (see Figure 14.15) or normal profit (see Figure 14.16).

The profit maximizing firm produces Q* output, where MC = MR (point x). As price (w) is greater than the short run average cost (z), the firm makes abnormal profit as shown by the shaded area w,x,z,y. This means that the firm is earning more profit in this market than it could in its next best alternative.

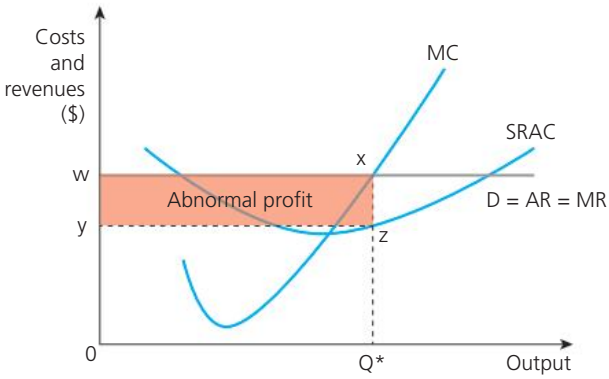


Figure 14.14 Short run abnormal profit in perfect competition

Top tip!

Notice that the MC curve intersects the AC curve at its minimum point. This is because:

- To the left of Q*, MC < AC, so the AC must be declining as output increases.
- To the right of Q*, MC > AC, so the AC must be rising as output increases.
- Hence, when MC = AC, there is no change in the AC.

Always draw your diagrams accurately in this way (showing the MC curve cuts the AC curve at its lowest point).

Alternatively, as seen in Figure 14.15, it is possible that the profit maximizing position of a firm may, in fact, be a loss-making position. The logic behind this is that, as a price taker, the position at which a firm makes the least amount of loss is technically the most profitable position for that firm. Therefore, the loss-making, profit maximizing firm will still produce where MC = MR, that is, at Q*. As price (c) is less than the average cost (b), the firm makes a loss in the short run, shown by the shaded area a,b,d,c. This means that in the short run, the firm is not making as much as its next best alternative.

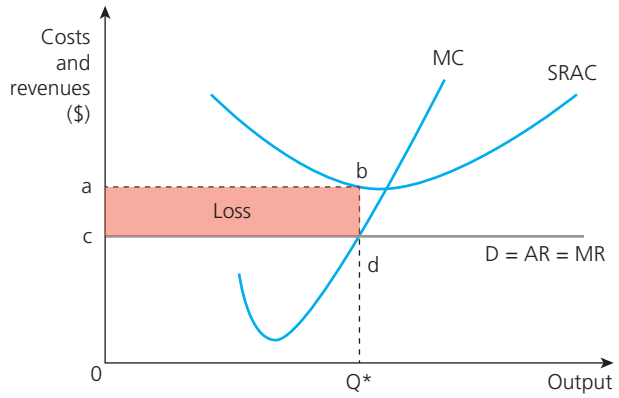


Figure 14.15 Short run losses in perfect competition

The third possible short run position for firms operating in perfect competition is earning normal profits. The corresponding diagram for this is the same as the long run position (see Figure 14.16 below).

Perfect competition – profit maximization in the long run

In the long run, perfectly competitive firms can earn only normal profits (zero economic profit), as shown in Figure 14.16.

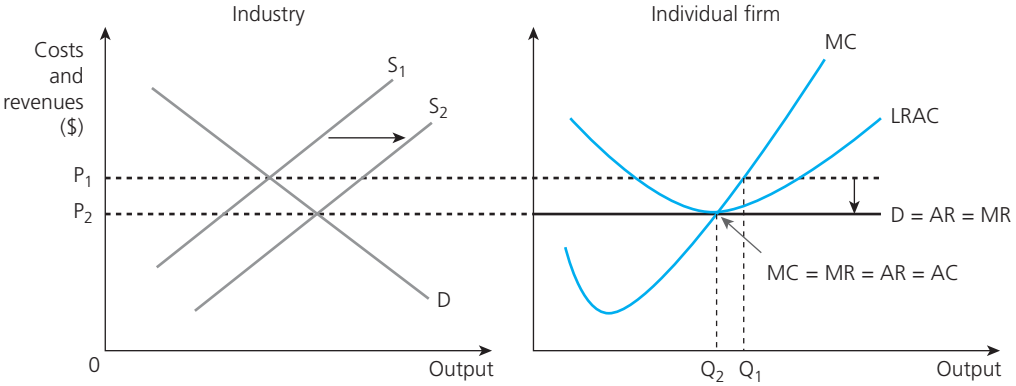


Figure 14.16 Long run normal profit in perfect competition

In Figure 14.16, the perfectly competitive firm initially operates at Q_1, earning abnormal profit at a market price of P_1. However, owing to a lack of barriers to entry and the presence of perfect knowledge, other firms will be able to enter the industry quickly. Hence, the market supply will increase, shifting the supply curve from S_1 to S_2. This will lead to a reduction in the price to P_2 (as determined by the price mechanism). Therefore, the existence of new entrants in the industry will shift the D = AR = MR curve lower because firms have to take (accept) the lower market price until they reach the minimum efficient scale (occurring at the lowest point on the **LRAC** curve). This results in a new profit maximizing output of Q_2, which occurs at the point of normal profit for the firm. Recall that if the firm earns zero economic profit, it will remain in the market as it is earning as much profit as its next best alternative.

◆ The **long run average cost** (LRAC) represents the cost per unit at different levels of output when all factors of production are variable. It is equal to the average fixed cost (AFC) plus the average variable cost (AVC).

If market price were to fall below the lowest part of the LRAC curve (as shown by Point b in Figure 14.16), then the firm would begin to make a loss. This would mean that the firm no longer enjoys as much profit as it could in its next best alternative. A rational firm would therefore leave the industry in the long run if faced with these conditions. This would then lead to a fall in the market supply, causing the D = AR = MR curve to rise owing to the resulting higher market price.

Hence, in the long run, perfectly competitive firms can earn only normal profits, that is, AR = AC. As these firms are also profit maximizers, they produce at the output level where MC = MR. This means the following long run condition must hold in perfectly competitive markets:

$$MR = MC = AR = AC$$

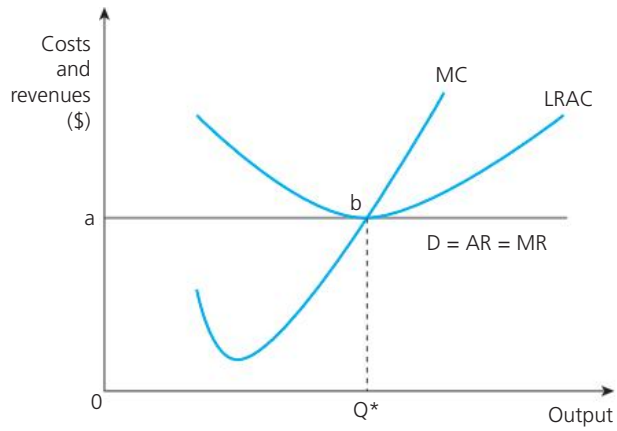


Figure 14.17 – Long run profit maximization in perfect competition

Figure 14.17 shows that the profit maximizing firm will produce where MC = MR, that is, at Q*. As price (a) is equal to the average cost (b), the firm makes zero economic profit (normal profit) in the long run.

Meaning of allocative efficiency

Efficiency is a core economic concept (one of the WIS**E** ChoICES). In economics, the term is a quantifiable concept, as determined by the ratio of total output to total input. **Allocative efficiency** refers to making the best use of an economy's scarce resources thereby ensuring the maximum possible ratio of total national output to total factor inputs, thereby minimizing or eliminating resource waste. Hence, social welfare is maximized if allocative efficiency is achieved. This implies that all factors of production are used efficiently so there are no unemployed resources.

◆ **Allocative efficiency** is the socially optimal situation that occurs when resources are distributed in such a way that consumers and producers get the maximum possible benefit, that is, no one can be made better off without making someone else worse off.

Allocative efficiency occurs when P = MC or MSB = MSC (as explained in Chapter 11), thereby generating maximum social surplus. At the socially optimal level of output, where MSB equals MSC, it is not possible to reallocate resources to make anyone better off without making others worse off. Social surplus (also known as *community surplus*) is the sum of consumer and producer surplus at the market equilibrium price. It represents the overall benefit to society from an economic transaction or activity (see Figure 11.2 on page 142).

An alternative explanation (or condition) for allocative efficiency refers to price (P) and marginal costs of production (MC). When price is set equal to marginal cost (P = MC), allocative efficiency is maximized. In other words, consumer and producer surplus are maximized. This is because the price that consumers are willing and able to pay matches that of the producers, so it is not possible to reallocate resources in order to make one party better off without making the other party worse off.

In Figure 14.17 above, allocative efficiency occurs when the perfectly competitive firm produces Q* units of output, where P = MC = MR = LRAC. Note that at this position, the following conditions hold:

- profit maximization (MC = MR)
- allocative efficiency (P = MC)
- cost minimization (MC = LRAC).

Hence, firms competing in perfectly competitive markets are unique in being able to profit maximize without compromising allocative efficiency. It is therefore understandable that some policymakers hold up perfect competition as the ideal market structure. Essentially, firms are forced to compete away their abnormal profits and operate close to the bottom of their long run average cost curves. This is arguably the most socially desirable outcome as individuals can

consume goods and services at the firm's cost of production (which makes allowances for earning normal profit). In the long run, this means that firms in perfect competition sell their goods and services at a profit, which is equal to the next best alternative forgone.

In perfectly competitive markets, firms are also said to be **productively efficient** in the long run. Productive efficiency occurs at the output level where average costs are minimized, that is, where MC = AC. Hence, there is no wastage or unemployment of scarce resources.

Therefore, a key advantage of perfect competition is **economic efficiency** because firms are allocatively efficient (charging a price that is equal to marginal cost) in both the short run and the long run. They are also productively efficient (they operate at the lowest point on the average cost curve) in the long run, although not necessarily in the short run (as they can temporarily earn economic profit).

Top tip!

The concept of **efficiency** spans the entire IB Economics syllabus. Allocative efficiency, in particular, is included in one of the six real-world issues as part of Unit 2 of the syllabus (How do consumers and producers make choices in trying to meet their economic objectives?). It is covered in Chapters 5 Competitive market equilibrium, 11 Market failure – externalities and common pool (common access) resources and, of course, 14 Market failure – market power (HL only).

Theory of Knowledge

To what extent would increased competition in the economy ever be considered undesirable?

Top tip!

When answering extended response questions, you must show evidence of critical thinking. This means you should consider the advantages and disadvantages of operating in competitive markets; for example, healthy competition can ensure consumers get the right products at the right price in an efficient way. However, there are limitations to the model – few industries, if any, actually fit the model in the real world. For instance, branding means that most products are heterogeneous in reality and it is not possible for consumers to have perfect information about all prices for all products.

Imperfect competition

Recall that under the model of perfect competition, firms are price takers as they lack market power. By contrast, firms operating in imperfectly competitive markets have varying degrees of market power, which determine the extent to which the individual firms are price makers. The market structures that are classified as imperfect competition are monopoly, oligopoly and monopolistic competition. These are covered in further detail in the next section.

Monopoly (AO3, AO4)

Recall that a **monopoly** is a market structure where there is a single supplier of a particular good or service, thus having the power to influence the market supply and price. The monopolist has significant market power as it controls enough of the market supply so can charge higher prices. Hence, it is a **price maker** (or price setter). As seen in Chapter 3, the *law of demand* is held by two factors: the income and the substitution effect. In the case of a monopoly, the substitution effect is not present, as it is the only firm in the market.

In the same way that a profit maximizing firm operating under conditions of perfect competition, the monopolist will also produce at the point where its MC = MR. As the monopolist enjoys market power, its price is set higher than the marginal costs of production (P = AR > MC), enabling the firm to enjoy abnormal profit.

◆ A **monopoly** is a market structure where there is a single supplier of a particular good or service, thus having the power to influence the market supply and price.

◆ A **price maker** (or price setter) is a single or dominant firm that has significant market power enabling it to manipulate its prices as it has significant control over market supply.

◆ **Profit maximization** is assumed to be the key goal of firms operating in the private sector. The largest positive difference between total revenue and total costs occurs at the output level where MC = MR.

Profit maximization in monopoly (AO3, AO4)

Assume, unless instructed otherwise, that a rational monopolist will seek the objective of profit maximization. As with perfectly competitive firms, this level of output will occur where MC = MR (see Figure 14.18). However, unlike firms operating in perfect competition, firms operating in imperfect competition face a downward sloping demand (D = AR) curve.

The monopoly in Figure 14.18 operates at the profit maximizing output of Q_{pm}. If this firm were to produce less than Q_{pm}, while it may still earn a profit, it would have forgone the sale of goods or services that had a greater marginal revenue (MR) than marginal cost (MC). In other words, the firm would have accumulated greater economic profit from the sale of these units. If the firm were to produce more than Q_{pm}, it would be producing units where the MC of production would exceed the MR gained from their sale. This would result in a loss from supplying these units. Therefore, it can be seen that when MC = MR, all profitable units have been produced for the monopolist.

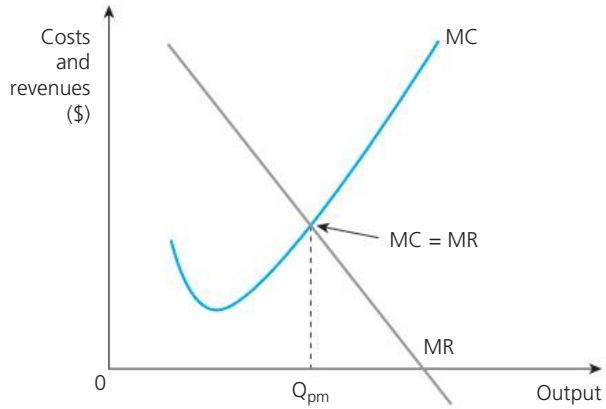


Figure 14.18 The MC = MR condition under monopoly

As seen in Figure 14.19, the average revenue curve for a monopolist is also the market demand curve (AR = D), as it is assumed there is a single supplier, that is, the firm is the industry. The AR = D curve is downward sloping because the monopolist must reduce price in order to sell more output. The gradient of the linear MR curve is twice that of the linear AR curve, because for the AR to fall continually, the MR must fall at a faster but constant rate.

The **profit maximizing monopolist** will supply at the output level where MC = MR, that is, at Q_{pm} and charge a price of P_{pm}, while average costs are AC_1. As the firm's AR > MC at this level of output, abnormal profit is earned (as shown by the red shaded area).

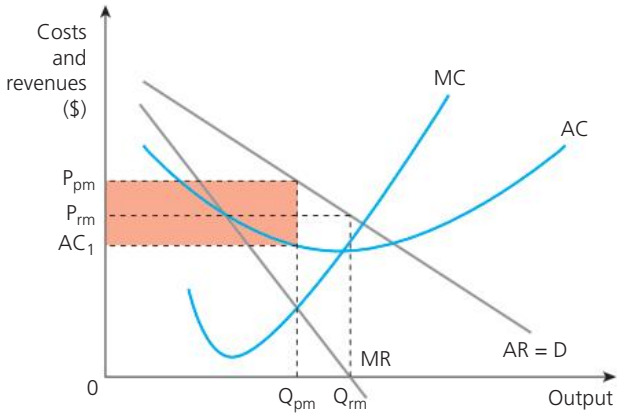


Figure 14.19 Profit maximization in monopoly (short and long run position)

It should be noted that the price elasticity of demand (PED) will change along a linear demand curve. Mathematically, MR = 0 at the mid-point of the AR curve for the monopolist. To the left of the mid-point of the AR curve, demand is relatively price elastic. The monopolist will not ordinarily choose to operate on the inelastic half of its AR curve because MR will be negative. This means profit cannot be maximized if MR is negative.

A monopolist might earn only normal profit (zero economic profit) in the short run. As shown in Figure 14.20, this occurs when AC = P. This could happen if the monopolist reduces price as a deliberate deterrent to potential entrants to the market. However, as the monopolist does not face any real competition, it is rational to assume the profit maximizing monopolist will raise its price above its average cost of production in the long run, so that P = AR > MC, and AR > AC.

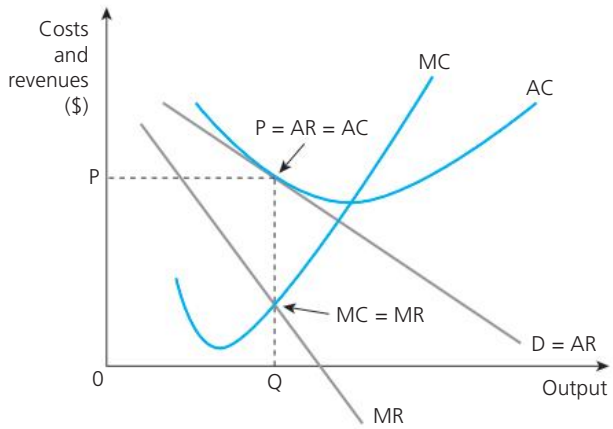


Figure 14.20 Short run normal profit position for a monopolist

As seen in Figure 14.21, it is also possible for a monopolist to make losses if its average cost curve is above its demand curve (AC > P). However, this should not persist in the long run as it would suggest that the monopolist is not covering its implicit costs, meaning that it could make more profit by pursuing the next best alternative – making it irrational to stay in this market in the long run.

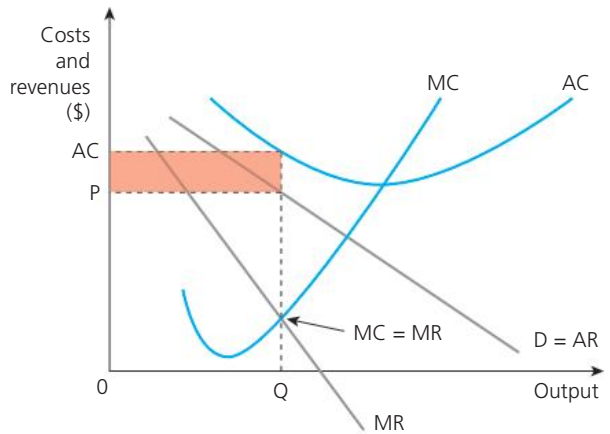


Figure 14.21 Short run loss minimization position for a monopolist

Top tip!

Students who claim that it is not possible for a monopolist to make a loss, given the lack of competition, should really consider *why* a monopolist might deliberately make a loss in the short run.

PAPER 3 EXAM PRACTICE QUESTION 14.5 (HL ONLY)

With reference to the diagram (Figure 14.22) below, identify the following:

a The price charged by the profit maximizing monopolist. *[1 mark]*

b The total costs for the profit maximizing monopolist. *[1 mark]*

c The amount of abnormal profit earned by the profit maximizing monopolist. *[1 mark]*

d The output level of the revenue maximizing monopolist. *[1 mark]*

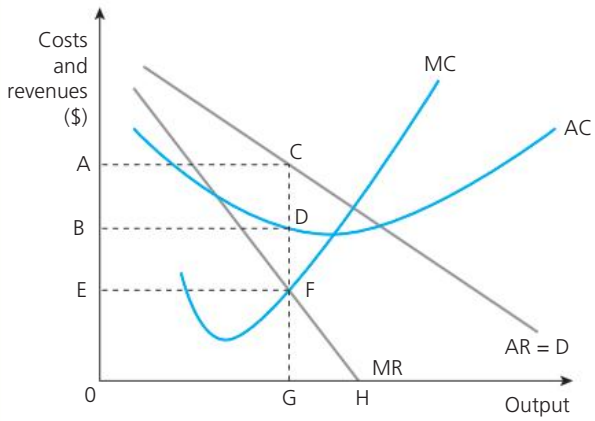


Figure 14.22 Costs and revenues for a profit maximizing monopolist

PAPER 1 EXAM PRACTICE QUESTION 14.6 (HL ONLY)

Explain why a monopolist is able to earn abnormal profits in the long run. *[10 marks]*

Allocative inefficiency (market failure) and welfare loss in a monopoly (AO3, AO4)

Unlike perfectly competitive firms, monopolies will not achieve productive efficiency. A profit maximizing monopoly lacks incentives to operate at the lowest point on its AC curve, as seen in Figure 14.23. Instead, the rational profit maximizing monopolist will limit market supply (to Q_{pm}), enabling it to charge a higher price (P_{pm}) than would be the case in perfectly competitive markets (P_{pc}).

There is also no incentive for a profit maximizing firm to seek allocative efficiency. Owing to the lack of competition, the price found in a monopoly is greater than the firm's marginal cost, that is, P > MC. For a perfectly competitive firm, the output would be higher at Q_{pc} and price would be lower at P_{pc}. The resulting welfare loss (the combined loss of consumer and producer surplus) is shown by the shaded red area in Figure 14.23. It can be seen that at the profit maximizing output (Q_{pm}, where the firm's MC = MR), consumers pay more for a lower amount of output. This is not a socially desirable outcome owing to restricted output and higher price in comparison with perfect competition.

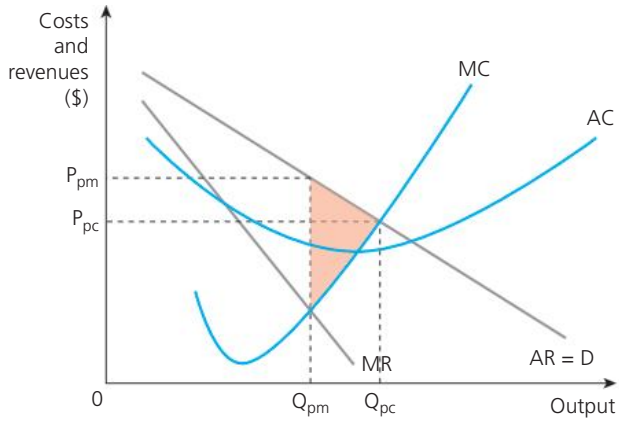


Figure 14.23 Monopoly and inefficiency

Nevertheless, there are reasons why, despite their inefficiencies, monopolies can be desirable. As a monopolist controls the market supply, it can achieve huge economies of scale. This means it can sell larger quantities *and* at lower average prices. This is particularly applicable to natural monopolies (see below). A monopolist, rather than many small but highly competitive firms, is more likely to have the financial ability to fund research and development (R&D) from its economic profits. R&D helps to create new innovations (new commercially viable plans, products and processes), thus improving the production capacity and international competitiveness of the economy.

Natural monopoly (AO3, AO4)

A **natural monopoly** occurs when the market can sustain only one firm in a profitable way. They tend to exist when the industry can tolerate only one supplier in order to avoid wasteful competition and to maximize the necessary economies of scale to operate in a sustainable way by having a single provider. This is because such industries have extremely high set-up costs, fixed costs and sunk costs (irrecoverable costs upon exit of the industry), which need to be covered by sufficient revenues to ensure supply of the good or service. This means that one provider would need all of the volume of sales to achieve such economies of scale.

◆ **Natural monopoly** occurs when only one firm can operate in a market profitably.

For example, it is more economical to have a single supplier of postal services, gas pipes, telephone lines, electricity cables and railway tracks because a monopolist could provide these with substantial cost savings compared to many smaller firms in direct competition. Therefore, trying to artificially increase competition in such industries actually creates a potential loss of efficiency as new entrants would mean a wasteful duplication of scarce resources – imagine having many competing railway lines going to the same destination or several gas, electricity and telephone companies laying cables and pipes into your home!

In the case of the state-owned Paris Métro, Europe's second busiest underground railway system, the challenges are demonstrated in Figure 14.24 where it can be seen that significant long run average costs (LRAC) exist. This is because of the high fixed and sunk costs that are involved with building an underground railway network. It is only after a considerable quantity of consumers that economies of scale (Q_1) can begin to be enjoyed. If another firm was to enter the industry and compete, Paris Métro's demand curve would shift inwards from AR = D_1 to AR = D_2 and it would no longer have enough commuters to cover its LRAC. Therefore, it is only sustainable for one firm to operate profitably in this market – hence, it is a natural monopoly.

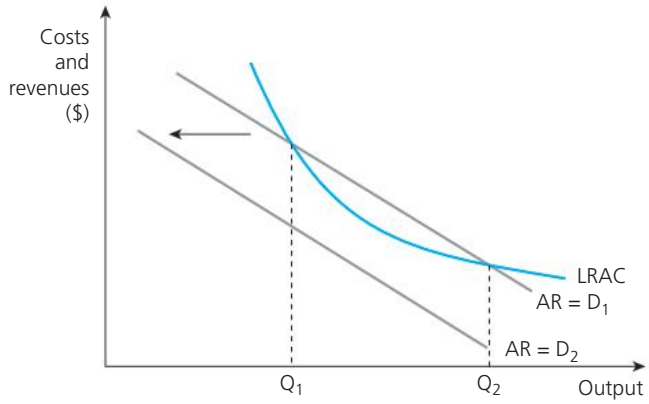


Figure 14.24 Natural monopoly

CASE STUDY

Métro, Boulot, Dodo!

Figure 14.25 The Paris Métro is a natural monopoly

Paris's underground network is run by Régie Autonome des Transports Parisiens (RAPT), a natural monpoly, and consists of two separate but linked systems: the metro and the Réseau Express Régional (RER) suburban train line. The metro has 14 numbered lines; the RER has five main lines. Paris is currently phasing out paper tickets. Since April 2019, the Navigo Easy contactless card (€2) allows infrequent transport users including visitors to prepay for journeys by topping up the card; there is no expiry date, and cards can be shared between passengers. The RAPT is a state-owned public transport operator, acting in the public interest as the free market is unlikely to be able to provide adequate public transportation service for the economy as a whole.

Source: Lonely Planet, 2020

Oligopoly (AO3, AO4)

An oligopoly is a market structure in which a few large firms, each with a high degree of market power, dominate the industry. There are two broad categories of oligopoly: collusive and non-collusive. A discerning characteristic of any oligopolistic market is the incentive for firms to collude and the incentive for firms to compete.

◆ **Collusive oligopoly** is an agreement between two or more oligopolistic firms to limit competition by using restrictive trade practices, such as price fixing or collectively limiting output.

Collusive oligopoly (AO3, AO4)

Collusive oligopoly refers to an agreement between two or more oligopolistic firms to limit competition by using restrictive trade practices, such as price fixing or collectively limiting output in order to manipulate the market price. As Adam Smith (1776) said, '*People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public or in some contrivance to raise prices.*'

◆ A **cartel** is an anti-competitive agreement between oligopolistic firms in the same industry to collude by fixing prices or to restrict the level of output, thereby effectively acting as a monopolist.

The incentive to collude is brought about by the potential to earn supernormal profits by restricting supply (thereby forcing up the market price) and creating artificial barriers for potential rivals

that are considering entering the industry. This can drive existing oligopolists to collude, which ultimately leads to the market becoming like a monopoly. Formally, such an alliance of oligopolistic firms is called a **cartel**. The most quoted example of a cartel is OPEC (Organization of Petroleum Exporting Countries), which limits the world supply of crude oil in order to keep prices high. The aim of forming a cartel is to reduce competition in the industry and to increase profit of the oligopolistic firms by acting as a group of monopolists. Cartels are illegal in many countries as they are anti-competitive. However, identifying and breaking up cartels can be difficult as it is often difficult to prove price fixing especially in the case of tacit collusion rather than open collusion.

The main methods of price fixing, an anti-competitive pricing strategy, are to collectively raise or reduce prices at the same time:

- Oligopolistic firms in the cartel may agree to simultaneously raise their prices to exploit customers by charging premium prices.
- They might agree to implement limit pricing – a pricing strategy used in a collusive oligopoly to set prices at a level that discourages new firms from entering the industry.

For collusion to be effective, there need to be high barriers to entry to prevent potential rivals from competing in the industry and mutual trust between colluding firms. It is also more effective in industries that supply or produce a homogeneous product, such as petroleum, as the cartel becomes a stronger collective monopoly selling a standardized product to customers. Collusion is illegal in many parts of the world, but tacit (implicit) collusion can be hard to detect or prove.

In Figure 14.26, the collusive oligopolistic firms have decided to fix a profit maximizing price of P_{pm} for their good or service at the output level where MC = MR. The profit maximizing quantity is referred to as Q/n with n being the number of colluding firms. This often yields them economic profits (shown by the shaded red area) as consumers have few substitutes especially as the firms have chosen to collude, thereby acting as a monopolist in the market.

CASE STUDY

Cartels in the car industry

In 2019, car manufacturers BMW, Volkswagen and Daimler (owners of Mercedes-Benz) were fined a total of €100m by the European Union after forming a cartel to buy steel. Germany's three largest carmakers had colluded with steelmakers for the purchasing prices of steel products needed to make their crankshafts, gearwheels or steering rods.

Non-collusive oligopoly (AO3)

Nevertheless, not all oligopolistic markets are collusive. Although firms in **non-collusive oligopoly** are in direct competition, they are said to have mutual interdependence, meaning that they consider the actions and reactions of their competitors when determining their pricing and non-pricing strategies (see section on page 205).

Each oligopolistic firm embarks upon a particular strategy without conspiring together with its competitors, although the firm needs to consider the possible reactions and strategies of its rivals in the industry. For example, if one firm lowers its price then other oligopolistic firms will tend to follow by lowering their prices to protect their sales revenue and market share. Naturally, this makes each of the firms worse off, despite the customer gaining consumer surplus.

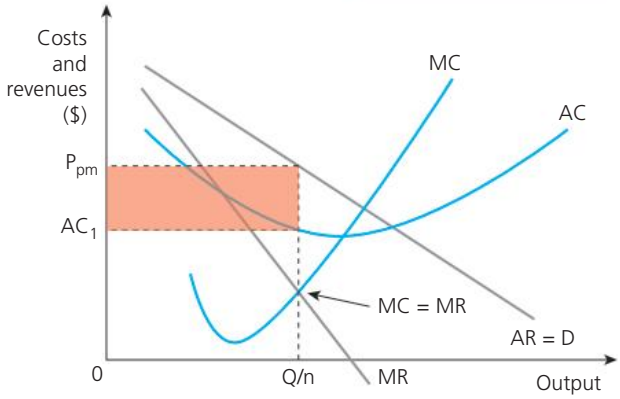


Figure 14.26 Collusive oligopoly acting as a monopoly

◆ **Non-collusive oligopoly** (or **competitive oligopoly**) refers to competing firms with mutual interdependence, This means they consider the likely or possible actions and reactions of their competitors when determining pricing and non-pricing strategies.

Interdependence, risk of price wars, incentives to collude, incentives to cheat

Some of the main characteristics or features of oligopoly are interdependence, risk of price wars, incentives to collude and incentives to cheat.

Interdependence

Interdependence means that oligopolistic firms consider the behaviours and decisions of their competitors when determining pricing and non-pricing strategies. For example, banks often

adjust their savings and mortgage interest rates by following the pricing strategies of the largest firms in the industry. The same applies to petroleum retailers, supermarket stores and car manufacturers. It is in the firm's best interest to be fully aware of the actions and reactions of their oligopolistic rivals, especially given their potentially high market power and the high intensity of competition in the industry.

Risk of price wars

Under competitive oligopoly, there is the potential for firms to engage in a price war. This occurs when firms continually reduce prices in an attempt to undercut the price being charged by rival firms so as to entice their customers to switch brands. Budget airlines, supermarkets and mobile phone network operators often use price cuts in an attempt to increase competitiveness and gain market share. However, price wars tend to be short-lived as it is not in the best interest of oligopolistic firms to continually reduce prices, especially for products with relatively price inelastic demand.

Figure 14.27 Large supermarkets often engage in temporary price wars

Incentives to collude

In the model of collusive oligopoly (see Figure 14.26), member firms can set a higher price for their good or service in order to gain more producer surplus and maximize their supernormal profits. By colluding, the individual oligopolies act as a monopoly and, as consumers have few substitutes in oligopolistic markets, are able to raise their prices. Hence, in theory, there are incentives to competing firms, all with a high degree of market power, to collude. Owing to its anti-competitive practices, collusion is illegal in most countries. Hence, it is usually undisclosed This means collusion takes place without any formal or written agreements. Price leadership, a common form of such tacit collusion, occurs when a firm (usually the market leader) sets a price that is then accepted by other firms as the going market price.

Incentives to cheat

However, there is a tendency for collusive oligopolies to collapse. This is because any individual firm willing to sell its products for less than the agreed fixed price with the other colluding firms has the prospect of gaining considerable revenue. This is commonly referred to as undercutting, and is considered 'cheating' by the other colluding members. It can therefore be said that there is an incentive for a colluding firm to 'cheat' as it seeks to act in self-interest. Most cartels often break down in the long run because individual firms tend to cheat (lower price and increasing output) to improve their own profitability at the expense of what are essentially rival firms in the market.

CASE STUDY

Disincentives to cheat

Breaking the law can often cause colluding firms to be exposed in the media and fined, thereby giving the firms bad publicity. In 2013, the European Union Commission fined electronics giants Samsung, Philips, Toshiba, Panasonic and LG a record €1.47 billion ($1.89bn) for conspiring to fix prices on some of their products between 1996 and 2006.

WISE ChoICES

Discuss the probable impacts of a price war or of price fixing on different stakeholders for a selected industry of your choice. Base your discussion on the concepts of **equity** and **choice**.

Allocative inefficiency (market failure)

Like monopolists, firms operating under conditions of collusive oligopoly are allocatively inefficient. With reference to Figure 14.26 (collusive oligopoly acting as a monopoly) on page 201, it can be seen that the oligopolistic firm is allocatively inefficient because:

- by operating at the profit maximizing level of output (Q/n), the firm charges more for its product than the marginal cost of production, that is, P = AR > MC
- the oligopoly also earns abnormal or supernormal profit, that is, P = AR > AC
- by being more competitive, the market would be able to supply more (than Q/n) and charge customers a lower price (below the profit maximizing price of P_{pm}).

Even in the case of non-collusive oligopoly, the existence of market power and asymmetric information cause market failures in oligopolistic markets. For example, car owners lack equal access to information enabling car manufacturers, car repair or service providers, and motor insurance companies to charge higher prices.

Essentially, the nature of interdependence and strategic decision-making means that socially undesirable prices and quantities are common in oligopolistic markets. Therefore, oligopolies are an example of market failure and their suboptimal output and anti-competitive behaviour can be regarded as allocative inefficiency.

Top tip!

Make sure you practise drawing the collusive oligopoly diagram and use this to explain allocative inefficiency of operating in oligopolistic markets.

The game theory payoff matrix (AO3, AO4)

Game theory is an economic model that attempts to explain the nature of strategic interdependence in oligopolistic markets. The theory considers the actions and behaviour of competitors when making a decision, based on various measured probable outcomes. Game theory marked a considerable shift in the way economists viewed competition.

For several hundred years, many economists believed in a collection of theories written by Adam Smith in his book *The Wealth of Nations* (1776). In Chapter 2 of Book 1 of his work, Smith wrote:

> *It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest.*
>
> *Smith, A,* The Wealth of Nations, *Book I, Chapter 2*

◆ **Game theory** attempts to explain the nature of strategic interdependence when making a decision in oligopolistic markets by considering the actions of competitors and based on probable outcomes.

Hence, Smith suggested that it is in the nature of firms to act in their own interest (such as profit maximization). He qualified this from a societal perspective later on, by writing:

> *… every individual necessarily labours to render the annual revenue of the society as great as he can. He generally, indeed, neither intends to promote the public interest, nor knows how much he is promoting it… and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention… By pursuing his own interest, he frequently promotes that of the society more effectually than when he really intends to promote it.*
>
> *Smith, A,* The Wealth of Nations, *Book IV, Chapter 2*

Hence, economists were in consensus for many years that competition and self-interest would make not only firms better off, but also the market and society as a whole.

However, in the 1950s, Professor John Nash (not yet a professor) developed his theories around Nash Equilibrium, which demonstrated that there are circumstances where firms acting in their best interest can make themselves and the market worse off. This was later seen as a major shift in economic thought, and made way for a popular aspect of economics called *game theory*.

In game theory, the outcome of a firm depends on the actions taken by other firms in the industry, with each firm having incomplete information about the competitor's intentions. Game theory, or the **prisoner's dilemma**, shows that oligopolistic firms can increase their profits through collusion rather than competing independently. A large part of this analysis hinges on symmetrical information (the assumption that each firm can predict the behaviour of the other). If we assume that these two firms are not colluding, then we need to determine which strategy is best for each firm and which is best for the market. Remember that Adam Smith argues that these two strategies should be the same. However, the tendency for rival firms to cheat results in a suboptimal outcome for all firms. Consider the scenario in Figure 14.28, where two firms, Green and Blue, are deciding whether to collude or compete.

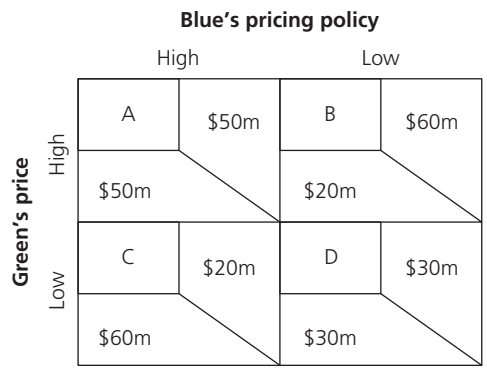


Figure 14.28 A simple game theory payoff matrix

With reference to Figure 14.28, if Green knew or expected that Blue was going to sell its products at a high price, then Green should sell at a low price (as the payoff is $60m – shown in quadrant C – rather than $50m from a high price – shown in quadrant A). Similarly, if Green knew or expected that Blue was going to sell at a low price, then it would also sell at a low price (because the payoff of $30m in quadrant D is more than the $20m they would receive selling at the high price in quadrant B). The payoffs work in the same way for Blue, that is, a low-price strategy will pay off more than a high price (quadrant B is a better outcome for Blue than quadrant A, and quadrant D gives a better payoff than quadrant C). In this case, a low-price strategy is therefore the strongest strategy that can be played by both firms if they are competing (non-collusive oligopoly).

However, if we assume that both oligopolistic firms choose a low price as their strategy (as it is the strongest approach), this will yield them a payoff of $30m each. Had both firms colluded and sold at a high price by using a price fixing strategy, they would have received a payoff of $50m each (quadrant A). Therefore, the low-price strategy of both firms is referred to as a **Nash equilibrium**, a suboptimal situation that arises due to oligopolistic firms competing, which ultimately results in both firms making themselves worse off.

Nevertheless, the incentive for oligopolies to cheat is still present. If Green had an agreement with Blue to charge a high price, but then cheated and sold at a low price they would get a payoff of $60m and Blue would receive only $20m (quadrant C). Blue will know this is a probable outcome (for its rival to cheat), so there is a tendency for the market to revert to a Nash equilibrium (quadrant D).

Theory of Knowledge

Economists assume that firms operating in oligopolistic industries are interdependent. In the context of market failures, what assumptions do economists make about the roles of reason and emotion? Are these assumptions justified?

Price and non-price competition (AO3, AO4)

Price competition is the use of pricing strategies to compete in an industry. For example, some firms adopt a *going-rate* pricing strategy where the pricing decision is based on the average price charged by rival firms in the industry. *Loss-leader* pricing can also be used by oligopolists such as large supermarkets, and involves setting the price of a product below its cost price in order to entice customers, who might also buy other, often related, products with higher profit margins sold by the firm. Some firms adopt *penetration pricing* when launching new products by setting a low introductory price to establish brand awareness and market share. *Promotional pricing* involves charging a low price to attract customers to encourage brand switching and to develop customer loyalty.

However, the market power of oligopolists often means there is **price rigidity** (price stability) in non-collusive oligopolistic markets, simply because competitors would simply retaliate by matching any price reduction from a rival firm. By contrast, firms in non-collusive oligopoly would not necessarily follow an increase in price in order to gain a competitive advantage.

Figure 14.29 Supermarkets often compete using various pricing strategies

Furthermore, as oligopolists can produce differentiated products, non-price competition is also a key feature in such markets. **Non-price competition** refers to all other forms of competition. For example, *advertising* is used extensively to differentiate between the products of competing firms. *Branding* is used in an attempt to gain customer loyalty. *Packaging* can also be used to distinguish products from one another and to act as a distinctive selling point; for example, the infamous glass bottle designs of Coca-Cola and Tabasco sauce, or Toblerone's prismatic shaped packaging.

◆ **Price competition** is the use of pricing strategies to compete in an industry.

◆ **Price rigidity** refers to the tendency of prices to remain unchanged in non-collusive oligopoly.

CASE STUDY

Loyalty card schemes

A popular way that firms in oligopoly compete is through the use of loyalty cards. It is a strange paradox that petrol stations around the world post their prices for all to see on the roadside, yet have loyalty cards which then change that price. The idea behind this is called *price-masking*. If the petrol station were to lower its price on its external boards, its competitors would just do the same. However, through the use of loyalty cards where the discount depends on the number of points accrued, it is impossible for competitors to see the real price that is charged to each customer.

Figure 14.30 Petrol companies use loyalty card schemes to entice customers

Measurement of market concentration (AO3, AO4)

Market concentration measures the extent to which sales revenue in an industry is dominated by one or more of the largest firms. This means market concentration is indicative of the degree of market power of these firms in the market. For example, the online social media industry is dominated by the likes of Facebook, Twitter, Pinterest and Instagram. Concentration ratios are used to measure the degree of market concentration (market power) in a particular industry.

A **concentration ratio** measures the degree of market power in an industry by adding the combined market shares of the largest few firms. For example, a *4-firm concentration ratio* measures the combined market shares of the four largest firms in an industry – the UK supermarket industry is dominated by Tesco, Sainsbury's, Asda and Morrisons with a combined 4-firm concentration ratio of over 69 per cent. The higher the concentration ratio for an industry, the more concentrated (oligopolistic) the market tends to be.

The most common method of calculating market concentration is the **Herfindahl–Hirschman Index** (HHI), named after economists Orris Clemens Herfindahl (1918–72) and Albert Otto Hirschman (1915–2012). The HHI gives a greater weighting to the market share of larger firms simply by squaring the value of their market share. Hence, the higher HHI number (for a pure monopoly) is 10,000 as the firm has 100 per cent market share. Therefore, the higher the HHI, the more concentrated the industry. For example, assume that the three-firm concentration ratio for Industry A and Industry B are both 65%:

- Industry A: Firm 1 = 30%, Firm 2 = 25% and Firm 3 = 10% market share
- Industry B: Firm 1 = 25%, Firm 2 = 20% and Firm 3 = 20% market share
- The HHI Index for Industry A = $30^2 + 25^2 + 10^2 = 1{,}625$
- The HHI Index for Industry B = $25^2 + 20^2 + 20^2 = 1{,}425$

Industry A has the higher Herfindahl–Hirschman Index value (1,625), so is said to be more oligopolistic than Industry B (1,425), even though they both have the same 3-firm concentration ratio (65%).

◆ Market concentration measures the extent to which sales revenue in an industry is dominated by one or more of the largest firms.

◆ Concentration ratio measures the degree of market power (or market concentration) in an industry by adding the combined market share of the largest few firms.

◆ Herfindahl Index is a measure of market concentration that gives greater weighting to the market power of larger firms by squaring the value of their market share.

◆ Monopolistic competition is a market structure where many firms exist, but each firm has a small degree of market power as they produce differentiated products.

ACTIVITY 14.4

With a partner, have a brief discussion to see whether you are able to name five different suppliers or brands of the following products. What do your answers reveal about market power in these industries? Be prepared to share your responses and reflections with the rest of the class.

1. Chewing gum
2. Toothpaste
3. Games consoles
4. Sports shoes
5. Cola drinks

Figure 14.31 Can you name 5 different brands of cola?

Top tip!

Whether branding and advertising represent inefficiencies depends on perspective and context. They can help oligopolistic and monopolistic firms to boost their sales, thereby achieving some productive efficiency gains through economies of scale. However, branding and advertising can also create inefficiencies if they act as a major artificial barrier to entry, giving certain brands significant monopoly power.

Monopolistic competition (AO3, AO4)

Recall that **monopolistic competition** is a market structure where there are many firms but each firm has only a small degree of market power. Hence, they are able to earn some economic profit, although it is also possible for firms in monopolistic competition to earn normal profit or losses (in the short run). Firms operating in monopolistic competition have less market power due to many substitutes being available, and, hence, face a more price elastic demand curve compared with firms operating in monopoly industries.

Figure 14.32 shows a firm in monopolistic competition that earns economic profit in the short run. As a profit maximizer, the firm operates at the output level (Q*) where MC = MR. Its price (P*) at this level of output is greater than the average cost of production (AR = P > AC), so the firm earns abnormal profit as shown by the shaded area. However, in the long run, the absence of barriers to entry mean that more firms enter the contestable market, attracted by the prospect of earning economic profit. However, this simply reduces the market price until AR = P = AC, thereby reverting all firms to a position of normal profit.

Alternatively, the firm could be in a loss-making position in the short run, as shown in Figure 14.33. It can be seen that the average cost (AC) at the profit maximizing level of output is higher than the selling price (P*). Losses will cause firms to exit the industry, thereby reducing the market supply and raising the market price until AR = P = AC, This means the industry will revert to a position of normal profit in the long run.

As can be seen in Figure 14.34, the existence of free entry to and exit from the industry means firms in monopolistic competition will earn zero economic profit in the long run. They still remain in the industry because if firms are achieving normal profit (as AR = P = AC) then they are making the same amount of profit as they could in their next best alternative. Note that the demand curve (D = AR) and marginal revenue curve (MR) are now more price elastic owing to the presence of more firms in the market (due to a greater substitution effect). This is also because these firms have less market power due to the availability of many substitutes.

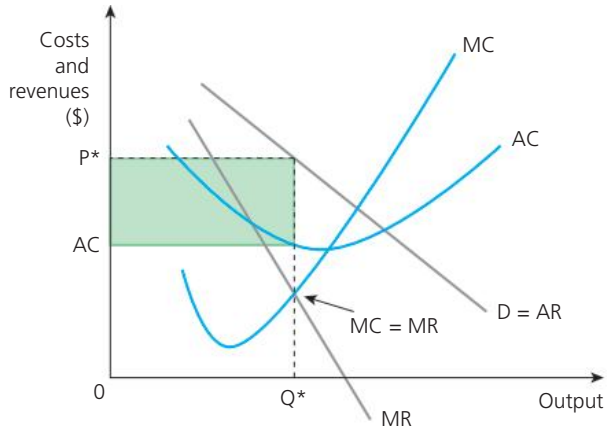


Figure 14.32 Short run abnormal profit in monopolistic competition

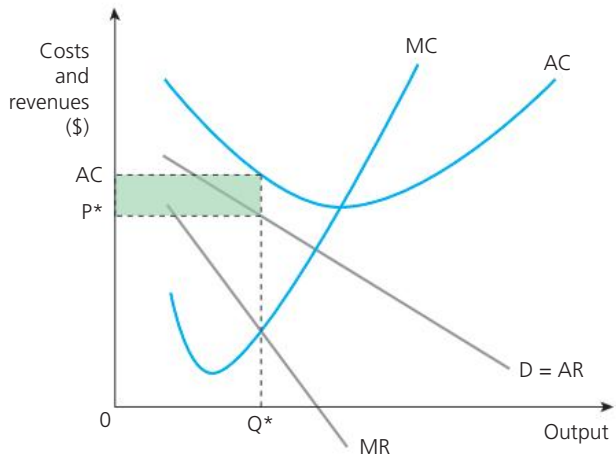


Figure 14.33 Short run loss position in monopolistic competition

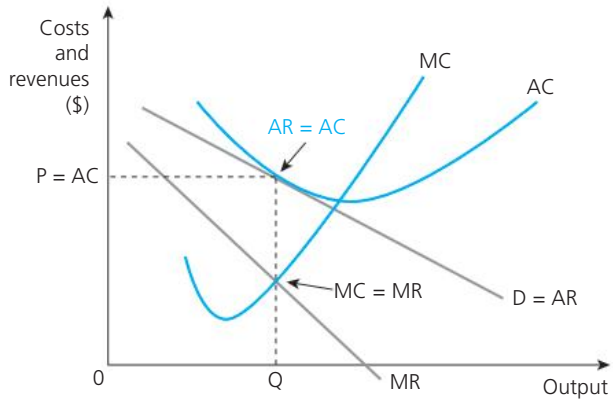


Figure 14.34 Long run normal profit in monopolistic competition

Top tip!

Note that the diagrams for firms operating in monopolistic competition look almost identical to those for a monopoly. However, the demand curve of a firm in monopolistic competition will be more price elastic as there are many firms operating in this market structure, whereas there is only one firm in a monopolistic industry.

Monopolistic competition and allocative inefficiency (market failure) (AO3, AO4)

Like monopolists and oligopolists, monopolistically competitive firms are not allocatively efficient in both the short run and long run. This means that they are effectively gaining consumer surplus thereby making consumers worse off. As can be seen in Figure 14.34, the monopolistically competitive firm does not operate at the lowest point on its AC curve (where P = MC = AC), because there are limited opportunities to exploit economies of scale. This means they struggle to get the demand needed to achieve these cost savings. Hence, consumers are deprived of enjoying a larger volume of output at a lower average price.

Inefficiencies and market failure can also occur in monopolistically competitive markets owing to the vast degree of choice that consumers have. Competition can be wasteful, such as advertising clutter and junk mail. Receiving too much marketing information from competing firms in the same industry can cause information overload or misinformation, which could lead to market failures.

Like a monopolist, the monopolistically competitive firm is a price maker (although to a lesser extent), so the price exceeds its marginal cost (P > MC), as seen in Figure 14.34. Again, this suggests that the firm could operate at a larger scale, up to the point where P = MC thereby maximizing efficiency gains.

Nevertheless, the monopolistically competitive firm faces far more rivalry than those operating in oligopoly or monopoly markets so is less inefficient. Also, because there are many firms in monopolistic competition, consumers could benefit from more product variety. For example, many shopping malls have a large range of restaurants or food vendors operating in a 'food hall' or 'food court' area of the shopping precinct. The vendors sell a variety of different products, creating more choice for consumers.

Figure 14.35 Fast food vendors in a food court provide product variety for consumers

Some firms might differentiate themselves by focusing on the quality of service (the level of customer service), giving hairdressers, clothes retailers, restaurateurs and other retailers a competitive advantage over their rivals. Similarly, providing good after-sales care can also create competitive advantages for monopolistically competitive firms, such as guarantees (warranties), technical support and delivery services.

PAPER 1 EXAM PRACTICE QUESTION 14.7 (HL ONLY)

Explain why product differentiation leads to a negatively sloped demand curve in monopolistically competitive markets. *[10 marks]*

PAPER 1 EXAM PRACTICE QUESTION 14.8 (HL ONLY)

Using real-world examples, compare and contrast the model of monopolistic competition with perfect competition. *[15 marks]*

Advantages of large firms having significant market power (AO3)

While perfect competition can create greater efficiency gains, large firms with significant market power can benefit from economies of scale and abnormal profits.

Economies of scale

Economies of scale refer to cost-saving benefits from lower average costs of production brought about by an increase in the volume of production. Economies of scale is a long run concept, because to increase the scale of production a firm must be able to increase all of its factors of production.

◆ **Economies of scale** are lower average costs brought about by an increase in the long run scale of production.

In section 2.2 of the IB Economics syllabus, it is discussed that firms in the short run face at least one factor of production that is fixed (leading to diminishing marginal returns).

When a firm grows in size, it will be able to enjoy cost savings that smaller firms may not be able to achieve. Some of these cost savings are due to the firm itself growing in size (internal economies of scale) and some are due to the industry growing in size (external economies of scale).

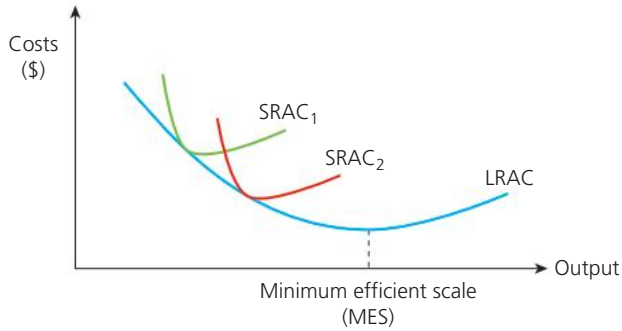


Figure 14.36 The long run average cost (LRAC) curve

The long run average cost (LRAC) curve is 'U' shaped owing to economies and diseconomies of scale, as shown in Figure 14.36. The LRAC curve encompasses all SRAC curves, representing the minimum average costs at each level of output. The lowest point on the LRAC curve represents maximum efficiency, known as the **minimum efficient scale** (MES) of production.

◆ The **minimum efficient scale** occurs at the point when all economies of scale have been exploited. This is shown at the lowest point of the LRAC curve.

Top tip!

It is not precise enough to refer to economies of scale as falling average costs as output increases. Remember that economies of scale is a long run concept, with a decline in the LRAC curve as a firm increases the size or scale of its production.

Internal economies of scale are lower average costs brought about by an increase in the size of a firm itself. Factors that give rise to internal economies of scale include:

◆ **Internal economies of scale** are lower average costs brought about by an increase in the size of a firm itself.

Specialization

Specialized labour (such as divorce or criminal lawyers, professors of economics and aircraft pilots) are highly productive in the output of goods and services, thus helping to reduce the average costs of production. These advantages are referred to as *managerial economies of scale*.

Efficiency

Specialized machinery, equipment and tools can be used to their full potential to raise output (such as in car manufacturing or the production of ball bearings), again helping to reduce the average costs of production. These gains are referred to as *technical economies of scale*.

Marketing

Effective advertising and branding helps to boost sales, so firms are able to produce on a larger scale. For example, Nike and McDonald's are able to market their brands to a global audience, thus reducing their marketing costs per unit of sales. These benefits are referred to as *marketing economies of scale*.

Purchasing

Large plants or factories can function efficiently only if large volumes of output are generated, such as mass-produced cars, soft drinks, canned products, cigarettes, bottled water or smartphones. By contrast, highly differentiated or unique output, such as bespoke wedding dresses, have higher unit costs of production. These cost savings are referred to as *purchasing economies of scale*.

◆ **External economies of scale** are lower average costs of production brought about by an increase in the overall size of an industry.

External economies of scale are the cost-saving benefits from lower average costs or production brought about by an increase in the size of the industry in which the firm operates. Hence, external economies are typically enjoyed by all firms in the industry. Examples of these benefits include:

Lower recruitment costs

If the industry that the firm is operating in is large, it is likely to have a large pool of qualified employees. For example, Singapore is a global financial hub, and many experienced and qualified finance professionals live and work there, making Singapore a cost-effective and time-efficient place for firms to recruit staff. Silicon Valley, California, is the global centre that hosts major high-technology, innovation and social media companies. Again, this helps to attract highly skilled and suitable workers to the area, helping firms in the San Francisco Bay Area to reduce their recruitment and search costs.

■ **Figure 14.37** Silicon Valley has a large pool of well-qualified and skilled high-tech workers

Ancillary services

Ancillary services refer to the services provided by other firms that support an industry. These can include legal, financial, courier and shipping services. If the industry that a firm operates in is large, then it is likely that these ancillary services will already exist and be located relatively close to that firm. The competition between these firms (or suppliers) helps to keep their prices lower and results in lower average costs for all firms in the area.

Top tip!

Natural monopolies enjoy internal and external economies of scale more than any other market structure. This should not come as a surprise given the significant scale of their operations and the fact that the industry can sustain only one firm in the market.

The upward sloping part of the LRAC curve (see Figure 14.36) shows **diseconomies of scale**. This means an increase in average costs caused by an increase in the scale of production. This means the firm is too large to operate efficiently. Diseconomies of scale occur when a firm becomes too large to manage efficiently so its unit costs begin to increase. For example, a firm increases its factor inputs (and, hence, production costs) by 20 per cent but the subsequent output increases by only 10 per cent. Diseconomies occur at all output levels beyond the MES point as shown in Figure 14.36.

◆ **Diseconomies of scale** are increasing average costs of production caused by an increase in the scale of production.

Internal diseconomies of scale are caused by problems of coordination, control and communication within a firm as it grows in size. For example, as firms get larger, managers find it more difficult to coordinate, control and communicate with a larger workforce, thereby resulting in higher unit costs.

Causes of **external** diseconomies of scale include problems such as traffic congestion, causing an increase in the average costs of production for all firms in the industry. Higher rental costs due to the high demand for firms locating in a certain area can also be seen in larger and growing markets. Labour shortages in a particular area can lead to increased labour costs, especially if the industry is growing rapidly.

Top tip!

Whereas internal economies of scale involve a movement down the long run average cost (LRAC) curve, external economies of scale will involve a downwards shift of the whole LRAC curve. Likewise, internal diseconomies of scale are shown by a movement up along the LRAC curve while external diseconomies shift the whole LRAC curve upwards.

Abnormal profits

Abnormal profits may help to finance investments in research and development (R&D) and, hence, innovation. This is a second advantage of large firms having significant market power. As a monopolist controls the market supply, it can achieve huge economies of scale. This means it can sell larger quantities *and* at lower average prices. While a perfectly competitive firm might be allocatively efficient, it is unlikely to be able to achieve the minimum efficient scale as its demand will not be sufficient to exploit economies of scale. Therefore, it can be argued that the profit maximizing price in monopoly may be lower than in long run perfect competition.

A monopolist, rather than many small highly competitive firms, is more likely to have the financial ability to finance R&D from its economic profits. R&D helps to create new innovations (commercial plans, products and processes), thus improves the production capacity and international competitiveness of the economy. It can also be argued that oligopolistic firms, such as large supermarket chains, can afford to offer greater consumer choice, thereby leading to an improvement in economic well-being. Perfectly competitive firms are unlikely to achieve the economic profits needed in the long run to be able to invest in the R&D needed to create innovative products.

Top tip!

You need to demonstrate that you clearly understand the difference between *internal* economies of scale (which relate to a specific firm) and *external* economies of scale (which relate to the whole industry).

PAPER 3 EXAM PRACTICE QUESTION 14.9 (HL ONLY)

Crispian's Candies has monthly fixed costs of $4,800 and average variable costs of $1.50. Demand for its candies is 4,000 units each month. The average unit price is $4.

a Calculate the average costs for *Crispian's Candies* each month. *[2 marks]*

b Calculate the units of output required for *Crispian's Candies* to break-even. *[2 marks]*

c Calculate the profit made by *Crispian's Candies* each month. *[2 marks]*

Common mistake

'Economies of scale mean that a firm's costs of production will fall as it increases the scale of its operations.'

This is not quite correct – clearly, it is cheaper to produce 1,000 cans of Coca-Cola than it is to produce 50,000 cans. However, it is cheaper to produce *each* can on a larger scale. This means economies of scale reduce the long run *average* costs of production.

Risks in markets dominated by one or a few very large firms (AO3)

There are risks (or disadvantages) in markets dominated by one or a few very large firms (monopolies and oligopolies). These can be categorized as risks in terms of output, price and consumer choice.

Risks in terms of output

A lack of competition means that profit maximizing monopolists are likely to supply less output than the social optimum (see Figures 14.20 and 14.35), that is, there is allocative inefficiency. This results in a welfare loss for society due to the under-provision and under-consumption of the product. In other words, the abuse of monopoly power causes a market failure. For example, as the world's largest cartel, OPEC deliberately limits the supply of petroleum in order to have control in the market.

Risks in terms of price

Firms with significant market power are able to use price discrimination. This means they charge different customers different prices for essentially the same product owing to differences in their price elasticity of demand (PED) for the good or service. Essentially, this enables a dominant firm to extract consumer surplus, leaving these consumers worse off.

Anti-competitive practices are more likely to be used by firms with significant market power. Firms that enjoy a monopoly (not a natural monopoly) will often purposefully act in ways that will maintain the monopoly, including the manipulation of market prices such as anti-competitive pricing strategies to prevent potential entrants from joining the market.

Common mistake

'Monopolists can charge whatever price they want, as they are the only supplier of a particular good or service so have no competition.'

While monopolists have the ability to control market supply, they do not control the level of market demand. As per the law of demand, customers will buy less or go without if the price rises too high. Hence, monopolists must lower their prices if they want to sell more.

Risks in terms of consumer choice

A lack of competition in a market dominated by monopolies or oligopolies can also lead to a lack of incentive for a firm to innovate and to provide choice for consumers. While a firm may wish to release new products to induce its customers to buy more, this is more likely to occur in highly competitive markets. A notable example would be the smartphone industry, where society would have much less choice of smartphones had there been only one supplier in the market. Barriers to entry deliberately set up by the dominant firm, such as marketing campaigns to build brand loyalty, also limit the market offering consumers a wider choice.

Top tip!

Consumer choice is not only limited in highly imperfectly competitive markets, but imperfect knowledge can also lead to consumers making irrational decisions and cause disincentives for firms to be innovative (due to the absence of competition).

Government intervention in response to abuse of significant market power (AO3)

Given the potential risks of market structures dominated by one or a few very large firms, government intervention may be needed in response to prevent or correct the abuse of firms with significant market power. For example, if monopolies or oligopolies exploit their market power and act against the public interest (perhaps by deliberately charging unreasonably high prices, restricting competition or by engaging in price fixing), then the government can intervene to break up their monopoly powers. By increasing the competitiveness of markets, the government reduces the opportunities for larger firms to abuse their market power. Government responses to the abuse of significant market power include legislation and regulation, government ownership (nationalization) and fines.

Legislation and regulation

Legislation and regulation are used to control firms with significant market power. Many countries have clear guidelines set by governments to regulate the operations of firms with monopoly power. For example, European Competition Law prohibits anti-competitive acts and the abuse of monopoly power. Rail regulators check the safety records of rail operators. Regulators are able to use significant powers to prevent, warn or penalize firms that abuse their power and treat customers unfairly. Examples of legislation and regulation include:

- **The prevention of mergers –** If a government believes that a merger between two very large firms in an industry can lead to socially undesirable market dominance, they can block (prevent) the merger from taking place. For example, the Competition and Markets Authority in the UK blocked the merger of two large supermarkets, Asda and Tesco, in 2019 owing to fears this would lead to higher prices for consumers.

- **Promoting competition** – A government may also look for ways to promote competition in an oligopolistic market. Strategies can centre around the removal of barriers to entry (especially overly bureaucratic and legal barriers). Governments can also give tax breaks or free advice for firms wishing to join an industry to make it more contestable.
- **Forced pricing strategies** – A government may also decide to force a monopoly to sell at a predetermined price such as its marginal cost (*marginal cost pricing*, or P = AR = MC) or at its average cost (*average cost pricing*, or P = AR = AC). In the case of marginal cost pricing, this would theoretically make the monopolist allocatively efficient.
- **Sale of existing assets** – When a firm is deemed to have become too large, and the market becomes incontestable (uncompetitive), the government may insist that the monopolist sells part of the firm or demerges.

Top tip!

Whether the government decides to break up a firm with significant market power depends on the perception or evidence of whether the monopolist or oligopolists are acting in the interest of the public (society as a whole).

ATL ACTIVITY 14.5

Have a go at this Think-Pair-Share activity:

Consider the potential benefits of a monopoly that operates in your country (or a country of your choice). To what extent is regulation of this monopoly needed? Pair up with a classmate a share your thoughts.

Government ownership

Government ownership of a firm with significant market power can help to ensure the business is not run primarily for profit, but to provide an important service to the general public. This process is called nationalization. **Nationalization** is the purchase of privately owned assets or industries by the government. This usually occurs when the government takes control of an industry previously in the private sector in order to run it in the best interest of the public. Government ownership and control of certain industries prevent the potential exploitation of firms with significant monopoly power. In such cases, the government will purchase the firm/industry at a predetermined market rate. This can be a costly exercise as the government will not only incur an opportunity cost in purchasing the private monopoly, it will also then have to invest resources in running the industry.

◆ **Nationalization** is the purchase of privately owned assets or industries by the government. This usually occurs when the government takes control of an industry previously in the private sector in order to run it in the best interest of the public.

Fines

There is a range of punitive measures that governments can use to penalize firms caught engaging in anti-competitive practices. Fines are financial penalties issued by governments for law infringements. For example, in 2019, Google was fined €1.49 billion ($1.66bn) by the European Commission for antitrust laws relating to advertising. These fines can be coupled with lengthy jail terms for executives at firms caught breaking competition laws.

The impact of these fines is a potentially large increase in the average costs of the firm, which will likely make it easier for new entrants to join the market and compete with the monopolist. The fines and exposure also damage the brand reputation of the firm, which may also lead to improved competition in the market.

CASE STUDY

Trouble for Tesco in Thailand

Figure 14.38 Tesco Lotus is a supermarket chain in Thailand

British supermarket giant Tesco owns nearly 2,000 outlets in Thailand and Malaysia, which three Thai companies are hoping to buy for potentially $10 billion. However, the deal is at risk because of Thailand's competition laws, which state the winning bidder cannot own more than half of the sector in the country. Any buyer that breaks the law could be fined up to 0.5 per cent of the transaction value.

Top tip!

Do not assume that monopolists are 'bad' for the economy. Instead, consider the extent to which their behaviour is in the best interest of the general public. Abuse of monopoly power is a type of market failure, so governments intervene to protect the interest of society as a whole.

WISE ChoICES

Consider the risks of increasing monopoly power and potential abuse in the technology industry. Discuss how such risks can influence **economic well-being**, **change** and **efficiency**.

Chapter summary

- Market structure refers to the key characteristics of a particular market, in terms of: the number and size of firms in the market, the degree and intensity of price and non-price competition, and the nature of barriers to entry.
- Market power refers to the ability of a firm or consumer to manipulate price above or below a perfectly competitive level.
- Barriers to entry are the obstacles that prevent firms from entering the market. They can be remembered with the acronym, **BLADE**, which stands for Branding, Legal, Anti-competitive measures, Domination of resources, Economies of Scale.
- The two extreme market structures in economics are highly competitive markets and monopoly. Imperfect competition exists in the case of monopoly, oligopoly and monopolistic competition.
- Characteristics of perfect competition include no barriers to entry or exit, homogeneous products, price takers and perfect knowledge. Perfectly competitive firms are unique in so far as they achieve economic efficiency while being able to profit maximize at the same time.
- Monopoly is a market structure where there is only one supplier of a good or service, with the power to affect market supply and prices. Characteristics of monopoly include: a single firm in the market, price makers, imperfect (or asymmetric) knowledge and very high barriers to entry.
- Advantages of monopoly include: the potential to supply larger quantities *and* at lower prices (due to economies of scale), having the financial resources for research and development, and the elimination of wasteful competition (in the case of a natural monopoly).

- Disadvantages of monopoly include: the potential to exploit customers through higher prices and limited supply (output), high entry barriers to limit competition and consumer choice.
- Oligopoly is a market structure in which a few large firms, each with a high degree of market power, dominate the industry. Characteristics of oligopoly include interdependence, the risk of price wars, incentives to collude and incentives to cheat.
- Collusive oligopoly is the agreement between two or more firms to limit competition by using restrictive trade practice; for example, price fixing or collectively limiting output. The main goal of a cartel is to restrict competition in order to maximize the profits for the colluding firms that act as if they were a collective monopoly.
- Concentration ratios are used to measure market concentration (the degree of market power of the largest few firms in an oligopoly).
- Game theory is an economic model that attempts to explain the nature of strategic interdependence in oligopolistic markets by considering the probable actions of competitors when making decisions.
- Monopolistic competition is where many firms exist, but each firm has a small degree of market power. Characteristics of monopolistic competition include product differentiation, negligible barriers to entry and exit, and a more price elastic demand curve compared with monopoly.
- The absence of barriers to entry in monopolistic competition means the market is contestable. If potential rival firms are aware that incumbent firms are earning economic profit, they will enter this market.
- Government responses to potentially anti-competitive behaviours of firms with high market power include legislation and regulation, government ownership (nationalization) and fines (financial penalties).

Table 14.3 Market structures summary

Criteria	Perfect competition	Monopolistic competition	Oligopoly	Monopoly
Nature of product	Homogeneous	Differentiated	Homogeneous/Differentiated	Unique
Customer loyalty	Does not exist	To a small extent	To a large extent	To a large extent
Knowledge of market conditions	Perfect knowledge	Relative knowledge	Relative knowledge Some trade secrets	Specialized information Trade secrets
Profit/break-even	Normal profit	Normal profit	Abnormal profits	Abnormal profits
Price taker	✓	–	–	–
Price maker	–	✓	✓	✓
Price competition	✓	✓	✓	–
Non-price competition	–	✓	✓	✓
Allocation of resources	Efficient allocation	Limited efficiency	No efficiency in allocation	Inefficient allocation
Price determination	Demand/supply mechanisms	Independent decisions of firms	Interdependent prices based on rivals	Deduces price independently
Competitiveness	Most competitive	2nd most competitive	3rd most competitive	Least competitive of all
Collusion	–	–	✓	–
Number of firms	Many (countless)	Many small firms	Few	One
Level of market control	Limited (if any)	Small degree	Relatively large degree	Complete control
Freedom of entry	Complete freedom	A few restrictions	Barriers to entry	Extreme barriers to entry
Demand curve	Perfectly price elastic	Relatively elastic	Relatively price inelastic	Relatively price inelastic

15 The market's inability to achieve equity (HL only)

Real-world issue

When are markets unable to satisfy important economic objectives – and does government intervention help?

Conceptual understandings (WISE ChoICES)

- The market mechanism may result in socially undesirable outcomes that do not achieve **efficiency**, environmental **sustainability** and/or **equity**.
- Resource overuse, resulting in challenges to environmental **sustainability**.
- Government **intervention** is effective, to varying degrees, in different real-world markets.

SYLLABUS CONTENT

By the end of the chapter, Higher Level students should be able to understand:

- why the workings of the free market economy may result in an unequal distribution of income and wealth (AO2)
- a diagram that shows the circular flow model to illustrate why the free market results in inequalities (AO3).

Workings of the free market economy may result in an unequal distribution of income and wealth (AO2)

The operations of the free market may result in an unequal distribution of income and wealth in the economy. Income can be measured for individuals (household income) or the economy as a whole (national income). Economists use income and wealth as indicators of living standards for individuals, households and the economy as a whole. However, income and wealth are not distributed equally in a free market economy. Most economists and politicians believe that inequalities in economy are detrimental to the well-being of society (see Chapter 19).

The degree of inequality also depends on what kinds of income are taken into account. Recall from Chapter 1 that households earn income in the form of wages and salaries from employment, although some individuals are able to also earn other factor incomes from rent, interest and profit for supplying land, capital and entrepreneurial resources in the production process. This means income is the sum of wages (and salaries), rent, interest and profit. Hence, it is inevitable that there will be an unequal distribution of income and wealth in a market economy, as the rich get richer and the poor get relatively poorer.

Recall that the circular flow of income is a macroeconomic model used to explain how economic activity and national income are determined. A person's **income** is a flow concept referring to money received on a regular basis from the production process, such as wages and

◆ **Income** is a flow concept referring to the money a person receives from the production process, for example, wages and salaries.

salaries from employment. When income is unevenly distributed in a region or country, **income inequality** is said to exist. A person's **wealth** is a stock concept referring to the value of that individual's stock (or accumulation) of assets, such as property, land, motor vehicles, stocks and shares, and savings. **Wealth inequality** exists when there is an unequal distribution of these assets across the economy.

◆ **Income inequality** refers to income being unevenly distributed in a community or country.

◆ **Wealth** is a stock concept referring to the value of accumulated assets, for example, property, land and savings.

◆ **Wealth inequality** refers to the unequal distribution of assets across the economy.

Top tip!

Inequalities in income and wealth occur at both a micro and macro level, from disparities between individuals within a community and inequalities between countries across the globe. When writing your exam answers, make sure you write in context and fully explain your arguments.

Figure 15.1 Inequalities exist in every country

The **circular flow of income model** (see Chapter 1) can be used to illustrate how the workings of the free market economy may result in an unequal distribution of income and wealth. The circular flow of income model forms the basis for understanding the macroeconomy and for explaining how national income, output and expenditure is generated over time.

In a closed economy, households supply factors of production (land, labour, capital and enterprise) to firms in order to produce or supply goods and services. In return, firms provide factor incomes to households. With this income, households spend their money on goods and services produced by firms, thus creating incomes for firms. In a more open economy of the circular flow of income model (see Figure 15.2), there is also consideration of the public sector, financial sector and international trade sector. This shows that not all income is spent on domestically produced goods and services because some of it is withdrawn in the form of taxation (T), savings (S) and import expenditure (M). Similarly, income can be injected into the economy in the form of government spending (G), export earnings (X) and investment expenditure (I).

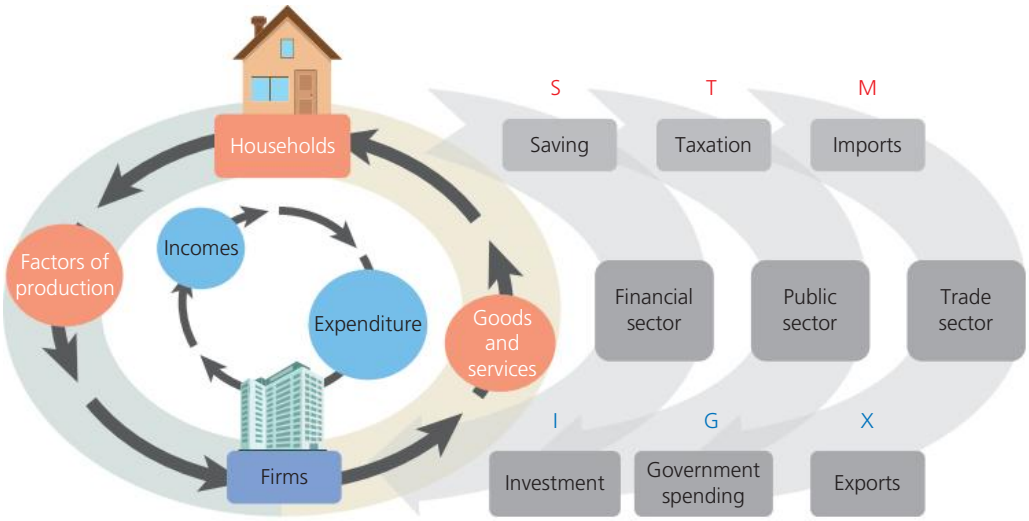


Figure 15.2 The circular flow of income model

As can be seen in Figure 15.2, an injection into the circular flow of income is money which enters the economy. This is in the form of government spending, investment and exports. A withdrawal from the circular flow of income results in money leaving the economy. This can be from taxes, saving and imports. Any imbalance in injections and leakages can result in different income flows resulting in inequalities, as explained in the following section.

Savings (S) and investments (I)

Households do not necessarily spend all their factor incomes, as they may choose to save (S) some of their income (Y) rather than spend it on consumption (C). This would reduce the circular flow of income in the economy, ceteris paribus. This is because savings represent a leakage from the circular flow. However, there are differences between how much individuals or households can save, which is dependent on several interrelated factors, such as:

- Differences in the level of human capital, such as skills, qualifications and experience. This is acquired and accumulates through education, training, professional development and work experience. Individuals with a high degree of human capital will tend to earn more, creating an income and wealth gap in the economy over time.
- Parental wealth and inheritance of income and wealth can give some individuals a major financial advantage in life, such as greater access to educational and healthcare services. Differences in the level of income and wealth owned by the previous generation create inequalities in the current and future generations.
- Differences in social attitudes towards women, including gender equality, can cause disparities in income and wealth. There tends to be less income inequality in countries where women are empowered and have equal rights, such as in Norway, Sweden, Denmark, Finland and New Zealand. The opposite tends to be true in countries where women's social and economic rights are suppressed, such as in Yemen, Afghanistan, South Sudan and Syria (see Table 15.1). The Fairtrade Foundation reported in 2020 that the average female cocoa farmer is paid just $0.30 a day, which is significantly less than the World Bank's extreme global poverty line of $1.90 a day.
- Similarly, wider disparities in income and wealth occur in countries that do not have or do not consistently enforce equal opportunities legislation and anti-discrimination laws. This applies beyond gender equality and covers social and economic rights of people irrespective of race, ethnicity, age, faith, sexual orientation, disability, marital status and so forth.

The Peace Research Institute Oslo (PRIO)'s annual The Women, Peace and Security Index reports on the world's best and worst places to live. The index is based on three broad criteria under the categories of inclusion, justice and security for women. Table 15.1 shows PRIO's Women, Peace and Security Index for 2019–20 for 167 countries.

Table 15.1 Gender equality based on inclusion, justice and security for women

Best performers		Worst performers	
Rank	***Country***	***Rank***	***Country***
1	Norway	158	Libya
2	Switzerland	159	Mali
3	Finland	160	Central African Rep.
4	Denmark	161	Dem. Rep. of Congo
5	Iceland	162	Iraq
6	Austria	163	South Sudan
7	United Kingdom	164	Pakistan
8	Luxembourg	165	Syria
9	Sweden	166	Afghanistan
10	Netherlands	167	Yemen

Source: PRIO (https://gps.prio.org/Publications/Publication/?x=1288)

ATL ACTIVITY 15.1

Take a look at the United Nation's website about the empowerment of women. Investigate why a government might choose to improve gender equalities and the empowerment of women.

https://www.unwomen.org/en/what-we-do/economic-empowerment/facts-and-figures

Although savings are a withdrawal out of the circular flow of income, firms also spend their income on fixed assets, such as machinery, vehicles, buildings and other capital goods. This expenditure is called **investment** (I) and represents an injection of income into the circular flow. Investment expenditure occurs partly because firms need to replace or upgrade existing machinery and capital equipment, and also because firms may wish to increase their production capacity.

Nevertheless, it is large firms with market power (see Chapter 14), such as multinational companies, that are most likely to invest in newer and better technologies. Smaller firms are unlikely to have the financial power to invest in the latest innovations, which creates greater inequalities between small and large firms.

Essentially, income inequalities can arise because labour and capital resources are compensated differently. This has a direct impact on the value of savings and investments in the economy, thereby influencing differences in income and wealth between individuals and firms.

Taxation (T) and government spending (G)

In a mixed economy with a public sector, households and firms need to pay taxes (T) to fund government expenditure (G). These opposing forces withdraw money out of the circular flow of income (in the case of taxation) and inject money (via government spending) into it.

Taxation refers to government charges or levies on income and expenditure. They can be categorized as direct taxes (such as personal income tax or corporate tax on business profits) or indirect taxes (such as sales taxes on goods and services). The higher the taxes are in a country, the lower the disposable income that remains for consumption expenditure and savings. Personal disposable income is generally more representative of an individual's standard of living rather than gross income (before tax deduction) because this amount determines the quantity of actual goods and services that can be bought.

Tax systems differ across the world. In general, a more progressive tax system (see Chapter 19) helps to redistribute income and wealth from high-income earners in a country to those on low or no incomes. Hence, progressive taxes help to reduce income and wealth inequalities. By contrast, in economies such as Hong Kong with a narrow tax structure (where there is only a relatively

small difference between the top and lowest rates of tax) there are likely to be large disparities in income and wealth (see Table 15.2). However, governments have a challenging task in establishing optimal tax rates – marginal tax rates (the additional tax paid on each extra dollar of income), that are too high create disincentives to work and invest, whereas marginal tax rates that are too low do little to tackle the economy's problems of inequalities in income and wealth.

Table 15.2 Tax rates, selected countries (2020)

Country	Top rate personal income tax (%)	Sales tax rate (%)	Corporate tax rate (%)
Sweden	57.2	25.0	21.4
Belgium	53.7	21.0	29.0
Australia	45.0	10.0	30.0
South Korea	42.0	10.0	25.0
Venezuela	34.0	16.0	34.0
Bangladesh	30.0	15.0	25.0
Singapore	22.0	7.0	17.0
Hong Kong	17.0	0.0	16.5
Russia	13.0	20.0	20.0
Saudi Arabia	0.0	5.0	20.0

Sources: https://tradingeconomics.com/country-list/personal-income-tax-rate *and* https://home.kpmg/xx/en/home/services/tax/tax-tools-and-resources/tax-rates-online/corporate-tax-rates-table.html

The circular flow of income model also shows another important source of money for some households that comes from the government in the form of *transfer payments*. For example, some low-income households may receive state welfare benefits such as unemployment benefit, pensions or child benefit. Such government spending is funded from direct and indirect taxes, thereby redistributing income in the economy.

The government also uses its tax revenues to inject income back into the economy by spending on merit and public goods (see Chapters 11 and 12 respectively), such as education, healthcare, the emergency services (police, ambulance and fire services) and public housing. Thus, such government spending provides further support for low-income earners and those unable to work.

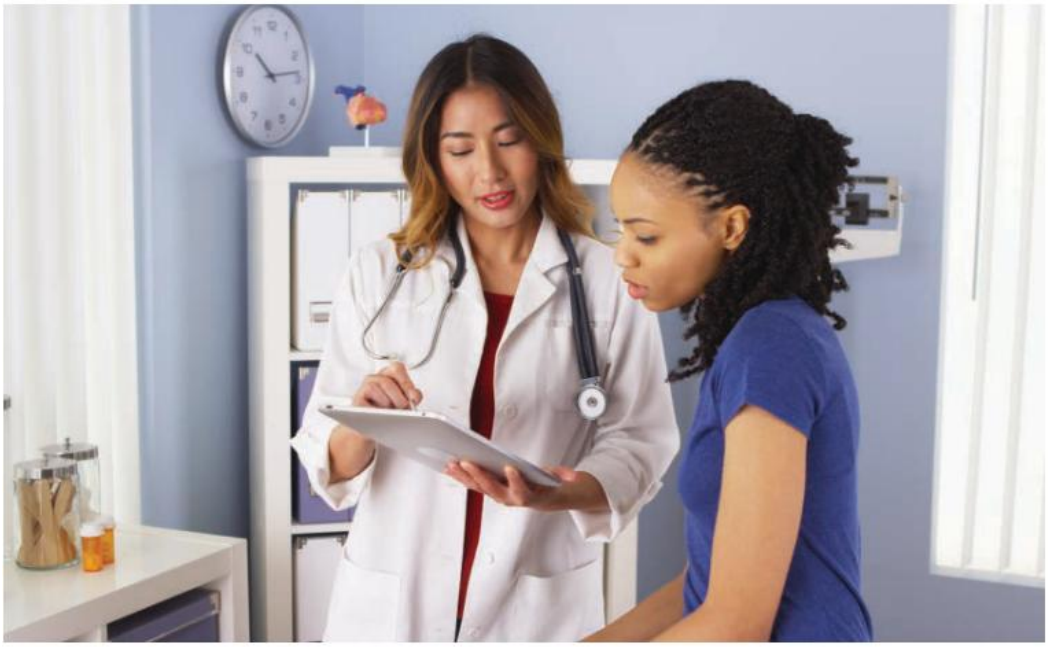


Figure 15.3 Government provision of healthcare services helps to rebalance inequalities in the economy

WISE ChoICES

During the coronavirus pandemic, governments around the world intervene to stabilize their economies. For example, the USA injected over $2 trillion to support the around 40 million unemployed people. However, such expenditure is not sustainable in the long run.

Discuss the extent to which **equity** can be achieved without government **intervention** in markets.

Imports (M) and exports (X)

In the open economy, the circular flow of income model is adjusted to include international trade. Domestic households and firms in an open economy will spend some of their income on goods and services from abroad, which are called **imports** (M). Import expenditure is a leakage from the economy's circular flow. By contrast, foreign consumers and firms will wish to buy domestic products, called **exports** (X). Export earnings are an injection into the economy's circular flow. In general, a country that exports more than it imports will tend to have higher real gross domestic product (GDP) and, hence, a higher standard of living in the long run, ceteris paribus. Income disparities arise between countries that enjoy net export earnings compared with those countries that have net import expenditure.

Figure 15.4 Index of coffee prices (2010–20), base year = 2015

Source: Trading Economics **https://tradingeconomics.com/commodity/coffee**

Figure 15.5 Honduras needs to sell a huge volume of coffee to purchase iPads from the USA

WISE ChoICES

To what extent does **intervention** in markets always bring about improved **economic well-being**?

One feature of economically less developed countries (ELDCs) is the over-reliance on production and employment in the primary sector, such as agricultural farming. The low value-added in the production of primary sector products, such as coffee beans, is a major source of poverty in ELDCs

(see Chapter 34). For example, the world price of coffee fell throughout 2010–20 (see Figure 15.4). In low-income countries that rely on coffee exports, such as Honduras (where nominal GDP per capita is around $2,200), this means a fall in the real incomes of farmers and their families who rely on this source of production for their livelihoods. By contrast, the output and export of manufactured high value-added electronic products and the provision of high value-added services in more economically developed countries (MEDCs) generates greater national income for these economies. Hence, there is a widening of the income and wealth gap between high-income and low-income countries.

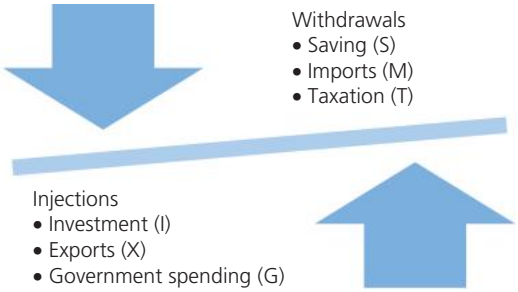


Figure 15.6 Changes in net injections and withdrawals directly affect the level of national income

Essentially, if the total value of imports exceeds that of exports (M > X), the net leakage of income from the economy causes greater inequalities between low-income and high-income countries. In the long run, this will cause negative economic growth and a rise in unemployment and poverty.

Overall, if the sum of leakages (S + T + M) is greater than the total value of injections in the economy (I + G + X), then the level of national income and employment will fall, resulting a recession in the country. If prolonged, this will cause poverty and greater income and wealth inequalities. By contrast, when the total value of injection is greater than the sum of leakages, national income and employment will rise, resulting in an economic boom and greater economic prosperity in the country.

Top tip!

Note that the circular flow of income diagram appears *three* separate times in the IB Economics syllabus – in Units 1.1, 2.12 and 3.1. Therefore, make sure you learn to draw this diagram accurately in preparation for the final exams.

WISE ChoICES

The World Bank reports that nearly half the world's population lives on less than $5.50 a day. Can the concepts of **economic well-being** and **intervention** be mutually exclusive in the study of welfare and development economics?

PAPER 1 EXAM PRACTICE QUESTION 15.1 (HL ONLY)

Explain why the workings of the free market economy may result in an unequal distribution of income and wealth. *[10 marks]*

Theory of Knowledge

How can economists claim to know when an economic problem is sufficiently large to warrant government intervention?

ATL ACTIVITY 15.2

For a country of your choice, use real-world data to show that women, on average, are paid less than men. Explain three economic reasons for the disparity between the average pay for men and women in the country.

Chapter summary

- The operations of the free market may result in an unequal distribution of income and wealth in the economy.
- Economists use income and wealth as indicators of living standards for individuals, households and the economy as a whole.
- The circular flow of income model can be used to explain the free market economy's ability to achieve equity in the distribution of income and wealth.
- Income inequality refers to income being unevenly distributed in a community or country, while wealth inequality is the unequal distribution of fixed assets across the economy.
- Essentially, income inequalities arise because labour and capital resources are compensated differently.
- If the sum of leakages (S + T + M) exceeds the sum of injections (I + G + X), then national income and employment will fall, resulting a recession in the country.
- If the total value of imports exceeds that of exports (M > X), the net leakage of income from the economy causes greater inequalities between low-income and high-income countries.

Macroeconomics

16 Measuring economic activity and illustrating its variations

Real-world issue

Why does economic activity vary over time and why does this matter?

Conceptual understandings (WISE ChoICES)

- **Change** in the conditions of the demand and supply sides of the economy cause economic activity to vary over time.
- Fluctuations in economic activity impact the **economic well-being** of individuals and societies.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- national income accounting as a measure of economic activity (AO2)
- the equivalence of the income, output and expenditure approaches to national income accounting, with reference to the circular flow model (AO2)
- a circular flow of income model diagram to show the interactions between decision makers, leakages and injections (AO4)
- nominal gross domestic product (GDP) as a measure of national output (AO2)
- nominal GDP from sets of national income data, using the expenditure approach (AO4)
- nominal gross national income (GNI) as a measure of national output (AO2)
- nominal GNI from data (AO4)
- the difference between real GDP and real GNI (AO2)
- real GDP and real GNI, using a price deflator (AO4)
- the difference between real GDP/GNI per capita and real GDP/GNI per capita at purchasing power parity (PPP) (AO2)
- real GDP per capita and real GNI per capita (AO4)
- the business cycle, including reference to short-term fluctuations and the long-term growth trend (potential output) (AO2)
- a business cycle diagram to show short-term fluctuations and the long-term growth trend (AO4)
- the appropriateness of using GDP or GNI statistics to measure economic well-being (AO3)
- alternative measures of well-being: OECD Better Life Index, Happiness Index and Happy Planet Index (AO2).

National income accounting as a measure of economic activity (AO2, AO4)

Macroeconomics focuses on the allocation of a country's scarce resources, with specific emphasis on the level of economic activity in relation to four macroeconomic objectives:

- a steady rate of increase in national income, that is, economic growth (see Chapter 18)
- a low and sustainable rate of unemployment in the economy (see Chapter 19)
- low and stable rate of inflation, that is, price stability (see Chapter 20)
- sustainable level of government or national debt (HL only) (see Chapter 21).

National income accounting is used to measure the level of economic activity in the economy. It refers to the money value of all the goods and services produced in a country during a year. The most common measure of national income is to calculate the country's gross domestic product (GDP). Note that GDP includes the value of the output of foreign-owned businesses operating within a country, such as the sales revenue from McDonald's restaurants operating outside of the USA. Measuring the level and rate of growth of national income is important for determining the country's rate of economic growth and changes to the standard of living in the economy.

◆ **National income** can be defined as the value of all the goods and services produced in a country during a year, thereby indicating the level of economic activity in the country.

CASE STUDY

McDonald's – the world's largest fast food chain

McDonald's is the world's largest restaurant chain as measured by sales revenue. According to its website, the US company serves more than 69 million customers every day in over 100 countries and employs more than 210,000 workers. The monetary value of the output from McDonald's therefore contributes to the gross domestic product of the countries that the company operates in.

■ **Figure 16.1** McDonald's has stores around the world, including this one in Dubai

There are three different methods that are used to calculate this:

1 **The output method**

The **output method** measures the actual value of all final (finished) goods and services produced within the economy each year. Only the final output is included in the calculation of national output to prevent double counting, such as the value of manufactured products used in the production process. Hence, this method of measuring the level of economic activity is also referred to as national output (O).

2 **The income method**

The **income method** of measuring the level of economic activity calculates the value of all factor incomes earned in the economy, that is, it is the sum of wages and salaries (labour), rent (land), interest (capital) and profits (enterprise). Households receive factor incomes for the output produced. Hence, this method is also referred to as national income (Y).

3 **The expenditure method**

Alternatively, the level of economic activity can be measured by the total value of all spending on goods and services in the economy (see Chapter 17). The **expenditure method** measures economic activity by calculating the value of total spending on newly produced goods and services during the year, comprising consumption (C), investment (I), government (G) and net exports (X – M).

- Consumption expenditure (C) refers to spending by individuals and households on goods and services. It is the single largest component of national expenditure.
- Investment expenditure (I) refers to spending by all firms within the economy in order to increase their capital stock and production capacity. This is sometimes referred to as *gross fixed capital formation*.

Top tip!

Note that **transfer payments**, such as state pensions and income support for low-income households, are not included in the calculation of national income as there is no corresponding output generated.

◆ The **output method** measures economic activity by calculating the actual value of all final (finished) goods and services produced within the economy each year. This method of measuring the level of economic activity is also referred to as national output (O).

- Government spending (G) refers to spending of the public sector.
- Net export expenditure (X – M) refers to the difference between the country's export earnings and its import expenditure in the year.

The expenditure method is expressed by the formula GDP = C + I + G + (X – M). Hence, this method of measuring the level of economic activity is also referred to as **national expenditure** (E).

Essentially, each of the three approaches of measuring the level of economic activity (national output, national income and national expenditure) measure the same thing, but by looking at different sets of data. Therefore, in theory, national income accounting will result in equal final figures or values, that is, O = Y = E. In reality, the figures may be slightly different due to inaccuracies and variations in the method of data collection.

◆ The **income method** measures economic activity by calculating the value of all factor incomes earned in the economy, that is, it is the sum of wages, rent, interest and profits.

◆ The **expenditure method** measures economic activity by calculating the value of total spending on newly produced goods and services during the year, comprising consumption (C), investment (I), government (G), and net exports (X – M).

National income accounting and the circular flow model

As a social science, economics uses theories and models (such as the circular flow of income) to explain the workings of an economy, such as variables that affect the level of economic activity. Recall that the **circular flow of income** (see Chapter 1) is a macroeconomic model used to explain how national income and economic activity are determined, based on the interactions of economic decision makers. As illustrated in this model (see Figure 16.2), the economic decision makers are *households* (private individuals in society), *firms* (that operate in the private sector of an economy), and the *government* (the public sector of an economy).

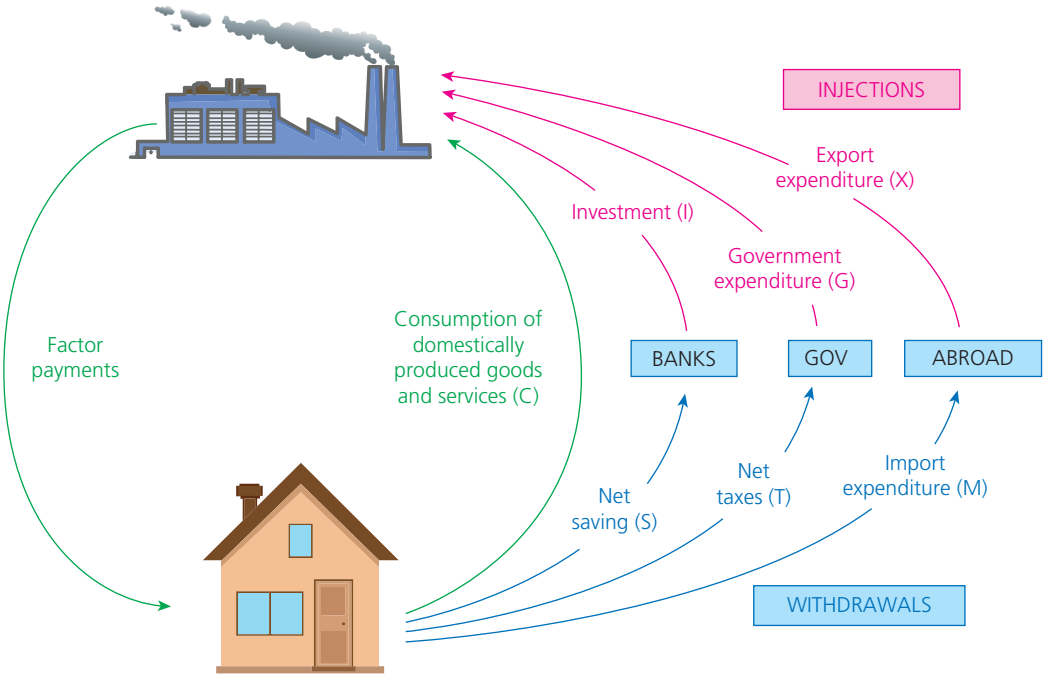


Figure 16.2 The circular flow of income model

As explained in Chapter 1, decision makers (or economic agents) interact with each other in the economy thereby generating economic activity. For example, households comprise individuals who provide their labour services to firms, in return for income (factor payments). With this income, households consume goods and services with the aim of maximizing their utility or satisfaction. Firms use factors of production to generate or supply the output of goods and services. Traditional economic theory assumes that firms do this in order to maximize profit (the reward to entrepreneurs for risk-taking and running a business). The government exists to maximize social welfare for the society as a whole. It taxes households and firms (on incomes

and profits) thereby generating tax revenues to fund government spending on goods and services, such as public and merit goods (see Chapters 11 and 12 respectively).

As shown in Figure 16.2, the circular flow of income model includes both **leakages** (or **withdrawals**) and **injections**. Leakages remove money from the circular flow of income whereas injections add money to it.

There are three types of leakages or withdrawals (W): savings (S), taxes (T) and import expenditure (M) as money leaves the circular flow of income and expenditure. There are three types of injections (J): investment (I), government spending (G), and export earnings (X) as money enters the circular flow and expenditure thereby stimulating economic activity. The model shows that the level of economic activity depends on the relative size of injections (J) and withdrawals (W). If W > J then economic activity declines, whereas economic activity increases if J > W. Hence, national income equilibrium exists when W = J, that is, S + T + M = I + G + X.

Note that the model shows equivalence of the three measures of economic activity: national income (Y), national output (O) and national expenditure (E). This is because national income is the return from generating national output that has been purchased by national expenditure. Hence, the circular flow of income model shows that Y = O = E.

◆ **Leakages** or **withdrawals** (W) take money out of the circular flow of the income. They comprise savings (S), taxation (T) and imports (M), that is, W = S + T + M.

◆ **Injections** (J) put money into the circular flow of the income. They comprise government spending (G), export earnings (X) and investment expenditure (I).

Top tip!

Practise drawing the circular flow of income diagram, which can be examined explicitly in the following sections of the IB Economics syllabus:

- Unit 1.1 – draw a circular flow of income model diagram, with leakages and injections to explain the interdependence between economic decision makers interacting and making choices in an economy.
- Unit 2.12 *(HL only)* – draw the circular flow model diagram to illustrate why the free market results in inequalities.
- Unit 3.1 – draw a circular flow of income model diagram to show the interactions between decision makers, as well as leakages and injections.

PAPER 1 EXAM PRACTICE QUESTION 16.1

Explain the three ways that national income can be measured. *[10 marks]*

PAPER 1 EXAM PRACTICE QUESTION 16.2

a Explain the impact that a cut in taxation and an increase in government spending might have on the circular flow of income. *[10 marks]*

b Using real-world examples, discuss whether the real gross national income (GNI) per capita of a country is a good indicator of its standard of living. *[15 marks]*

◆ **Gross domestic product** (GDP) is the value of all final output of goods and services produced within a country in a year.

◆ **Nominal GDP** measures national output using current market prices, that is, the value of GDP at the time of measurement. It measures the monetary value (or face value) of all goods and services produced within a country for a given period of time, usually a year.

Nominal gross domestic product as a measure of national output (AO2, AO4)

Gross domestic product (GDP) is the value of all final output of goods and services produced within a country in a year. **Nominal GDP** measures national output using current prices, that is, the value of GDP at the time of measurement. Therefore, it does not make any adjustment for the effects of inflation in the calculation of prices of goods and services.

Recall that there are three ways to measure the level of economic activity: the value of national output (O), national income (Y) and national expenditure (E). Irrespective of the method chosen, all three approaches should give the same numerical result, that is, O = Y = E because the value of national output equals the value of what is spent on the output (national expenditure). This expenditure becomes the (national) income to households and firms that produced the output.

Using the expenditure method of calculating nominal GDP, economists add the totals of consumption expenditure (C), investment expenditure (I), government expenditure (G) and net export expenditure (X – M), that is, nominal GDP = C + I + G + (X – M).

When calculating the nominal GDP, there is no need to discount (or deflate) the impact of price fluctuations caused by inflation. Hence, economists measure nominal GDP by using **current prices** – the actual prices of goods and services at the time of collecting the data.

◆ **Current prices** refer to the actual prices of goods and services at the time of collecting the data, which are used to report nominal GDP and nominal GNI.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 16.3

Calculate the value of nominal gross domestic product (GDP) from the given information: Consumption = $150bn, Investment expenditure = $60bn, Government spending = $55bn, Export earnings = $31bn, Import expenditures = $28bn, Net income earned abroad = -$8bn. *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 16.4

The figures below show a particular country's national income statistics for last year (all figures in $ billion).

- Gross domestic product = $4,770
- Consumption = $3,720
- Government expenditure = $668
- Investment expenditure = $524

Calculate the value of the country's net exports from the figures above. *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 16.5

The figures below show a particular country's nominal GDP for last year (all figures in $billion).

- Private consumer expenditure = $1,205
- Government consumption expenditure = $680
- Gross domestic fixed capital formation = $1,150
- Domestic exports of goods = $850
- Imports of goods = $940
- Domestic exports of services = $1,100
- Imports of services = $1,700

Calculate the value of the country's gross domestic product (GDP) at market prices. *[2 marks]*

Nominal gross national income (GNI) as a measure of national output (AO2, AO4)

An alternative to using *nominal GDP* (or *GDP at market prices*) as a measure of national output is nominal gross national income (GNI). Recall that nominal GDP refers to the value of all final goods and services produced within a country per year. By contrast, nominal GNI calculations include the country's GDP along with net factor income earned from abroad (such as net interest, profits and dividends). Hence, **GNI** is the country's total income earned from its factors of production irrespective of where they are located, known as **net factor income from abroad** (or *net property income from abroad*).

◆ **Gross national income (GNI)** is the value of a country's final output of all goods and services plus net factor income earned from abroad.

◆ **Net factor income from abroad** is the net income gained from other nations, that is, the difference between a country's income earned from abroad and the income earned by foreign individuals and firms in the domestic economy.

Nominal GNI = Nominal GDP + Net factor income from abroad

For example, Japanese carmakers Honda, Nissan and Toyota have production plants in the UK. The output of their cars in the UK is counted as part of the country's GDP. However, if the Japanese firms send profits back to their headquarters in Japan, then this withdrawal of income is subtracted from the UK's GNI. Hence, a country hosting many multinational corporations that repatriate profits from local production and sales mean that its GNI will be lower than its GDP.

Common mistake

You may confuse GDP with GNI. Although they both measure the level of economic activity in terms of national output, GDP is geographically based whereas GNI considers a country's factor incomes irrespective of location.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 16.6

Calculate the value of nominal gross national income (GNI) from the given information: Consumption = $150bn, Export earnings = $31bn, Government spending = $55bn, Import expenditures = $28bn, Investment expenditure = $60bn, Net income earned abroad = -$8bn. *[2 marks]*

Real GDP and real GNI (AO2, AO4)

Real GDP takes account of fluctuations in prices (inflation) over time. Using real GDP (or real GNI) figures allows better comparisons of economic activity over time by using **constant prices** (the values of real GDP and real GNI that are adjusted for inflation over time). For example, if nominal GDP increases by 4 per cent while the general price level rises by 3 per cent, then the average person's income increases by only 1 per cent in real terms.

◆ **Real GDP** (and real GNI) takes account of fluctuations in prices that affect the value of nominal national output. Instead, the values are adjusted for inflation by using a GDP deflator.

◆ **Constant prices** refer to the values of real GDP and real GNI as they have been adjusted for inflation over time.

Comparing the nominal GNI or nominal GDP from year to year is a less meaningful measure of changes in national output or the level of economic activity because prices change over time. To calculate real GDP (or real GNI), the values of nominal GDP (or nominal GNI) are adjusted for inflation by using a **GDP deflator** (or **price deflator**). This is used to convert GDP or GNI at current prices to GDP or GNI at constant prices. The GDP deflator can be viewed as a measure of the general level of inflation in the economy, showing the extent of changes in price levels over

◆ The **GDP deflator** is a measure of the general level of inflation in the economy, showing the extent to which prices have changed over time. It is used to calculate real GDP.

time. For example, a GDP deflator of 106.5 means that prices have risen, on average, by around 6.5 per cent since the base year. To deflate the value of nominal GDP to get the real GDP figure, the following formula is used:

Real GDP = Nominal GDP ÷ GDP deflator

PAPER 2 AND 3 EXAM PRACTICE QUESTION 16.7

Use the data below to calculate the real gross domestic product (real GDP) in 2019 and in 2020 and explain your results. *[4 marks]*

Year	Nominal GDP ($bn)	GDP deflator
2019	260.0	106.7
2020	262.4	108.5

Top tip!

Make sure you know the difference between *current* and *constant prices*. National income statistics reported in current prices include the impact of inflation while those expressed as constant prices have the effects of inflation removed (allowing for better comparisons of GDP over time and between countries).

Real GDP/GNI per person (per capita) (AO2, AO4)

Economists find it useful to calculate real GDP and real GNI expressed in terms of per capita as this averages out total GDP per person in the country. For example, China has the second largest GDP in the world (around $14.2 trillion in 2020) whereas Luxembourg's GDP is much lower (around $72.1 billion). This would suggest that China is significantly wealthier than Luxembourg. However, when the population size is accounted for, the figures are drastically different. With a population of over 1.4 billion people in China, this equates to a GDP per capita of about $7,755 whereas the much smaller population of around 610,000 people in Luxembourg means that the GDP per capita is $107,243. This would suggest that the average person in Luxembourg has a much higher level of income that the average person in China. Hence, **real GDP per capita** is a way of measuring the standard of living for the citizens of a country. Real GDP per capita is calculated by using the formula:

◆ **Real GDP per capita** means expressing the real GDP of a country in terms of its population size in order to determine the value of national income per person.

Real Gross Domestic Product per capita = Real GDP ÷ Population size

Similarly, real GNI per capita is calculated in the same way, by dividing the total real GNI by the country's population size:

Gross National Income per capita = Real GNI ÷ Population size

Real GDP/GNI per capita at purchasing power parity (PPP)

Taking this one step further, it is possible to calculate the real GDP or real GNI at **purchasing power parity** (PPP). As GDP for different countries is measured in their local currency, trying to make international comparisons can be difficult. Even if a common currency is used for international comparisons (usually the US dollar) exchange rates can be very volatile from month to month and from year to year, and do not reveal differences in purchasing power.

◆ **Purchasing power parity** (PPP) refers to the exchange rate that enables residents to purchase a common basket of goods and services in different countries.

For example, if $1 = 23,000 Vietnamese dong (VND), it does not mean that something in the USA that is $1 will be priced at VND23,000. Purchasing power parity is the exchange rate needed for people to buy the same quantity or basket of goods and services in different countries using the same amount of money. This means real PPP equalizes the purchasing power of two currencies by considering the differences in the cost of living and inflation. This means that currencies are at par when the exchange rate means the same basket of goods and services is priced the same in these countries. So, if a cup of cappuccino is priced at $5 in the USA, the exchange rate suggested the product should be VND115,000. However, if the cup of cappuccino is priced at VND90,000 then this equates to only $3.91. This means the VND is undervalued against the USD. Therefore, as $5 = VND90,000, purchasing power parity exists when $1 = VND18,000. This is known as the *PPP adjusted exchange rate* (rather than the *market exchange rate* of VND23,000).

Top tip!

PPP adjusted exchange rates enable economists to see if a currency is under or over-valued, it is a rather arbitrary measure. Only certain products are selected but the basket of goods and services is not necessarily representative of what the average local resident purchases in different countries. The products, such as rice or potatoes, are not necessarily homogeneous nor are they necessarily of the same quality.

ATL ACTIVITY 16.1

Economists often refer to the Big Mac Index when discussing purchasing power parity. This crude method of measuring PPP was introduced by *The Economist* back in 1986. Visit www.bigmacindex.org to investigate how the Big Mac Index works. As a light-hearted way to measure whether currencies are under- or over-valued, discuss the limitations of using the Big Mac Index to measure purchasing power parity.

PAPER 2 AND PAPER 3 EXAM PRACTICE QUESTION 16.8

From the data below for Taiwan, calculate the percentage change in the country's gross domestic product (GDP) per capita between 2018 and 2019. *[2 marks]*

Year	GDP (US$ bn)	Population (million)
2018	590	23.59
2019	605	23.60

PAPER 2 AND PAPER 3 EXAM PRACTICE QUESTION 16.9

Year	Nominal GDP ($bn)	GDP deflator
2018	228.0	106.0
2019	230.2	107.8
2020	232.4	109.8

a Using the data above, calculate the real GDP for Country G in 2019. *[2 marks]*

b In Country G, the average annual salary in 2020 is $24,000. Calculate the average real income for the average worker in Country G. *[2 marks]*

c Explain why despite the nominal GDP increasing in the given time period, the real value of GDP in Country G has actually fallen. *[3 marks]*

Business cycle (AO2, AO4)

The **business cycle** (also known as the **trade cycle**) describes the fluctuations in the level of economic activity in a country over time. The short-term fluctuations show the actual level of economic activity during the usual cycle of economic booms, recessions, slumps and recoveries. The business cycle also shows the long-term growth trend (or potential output) by establishing the trend from the actual data of economic activity (see Figure 16.3).

◆ The **business cycle** (or **trade cycle**) is a model that describes the fluctuations in the level of economic activity of a country over time, thus creating a long-term trend of economic growth in the economy.

With reference to Figure 16.3, the various stages in the business cycle are outlined below:

- A **boom** is a phase in the business cycle when the level of economic activity rises, caused by an increase in any component of aggregate demand, that is, consumption, investment, government spending and net export earnings. The general price level rises during an economic boom due to increased expenditure in the economy.
- The **peak** of the business cycle occurs when economic activity is at its highest level, so unemployment is low, and business and consumer confidence levels are very high.
- A **recession** is a phase in the business cycle that occurs when there is a fall in gross domestic product (GDP) for two consecutive quarters. The decline in the level of economic activity eventually causes business failure and the level of unemployment to rise. A recession creates uncertainty for households and firms, thereby reducing both consumer and business confidence levels.

◆ A **boom** is a phase in the business cycle when the level of economic activity rises, caused by an increase in aggregate demand.

◆ The **peak** of the business cycle occurs when economic activity is at its highest level.

◆ A **recession** is a phase in the business cycle that occurs when there is a fall in gross domestic product (GDP) for two consecutive quarters.

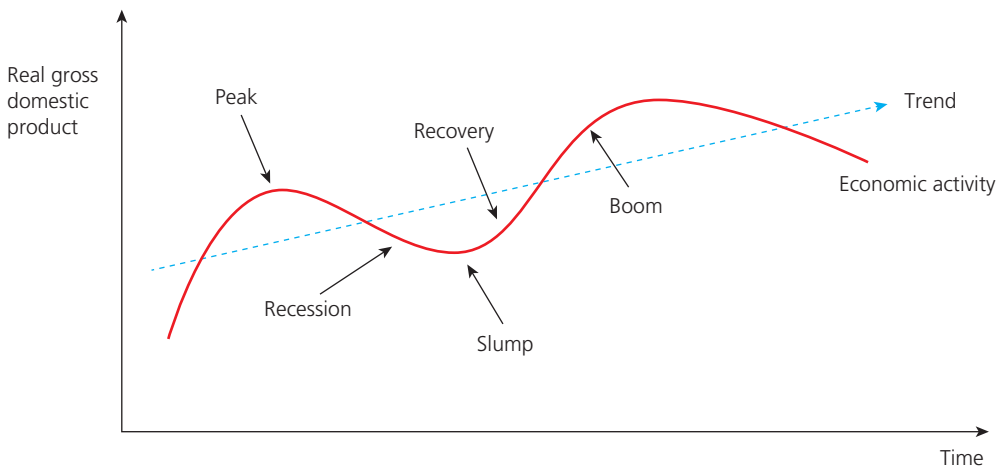


Figure 16.3 The business cycle

- A **slump** (or **trough**) occurs at the bottom of a recession in the business cycle when consumption, investment and net exports are at a low point, resulting in negative economic growth, business closures and cyclical unemployment (mass-scale unemployment across the economy). This is because the level of consumption, investment, and net export earnings all remain low.
- A **recovery** occurs when GDP starts to rise after the trough in the business cycle. Consumption, investment and net exports gradually rise in value, thus creating employment opportunities and increasing both consumer and business confidence levels. Hence, recovery eventually leads to economic growth and employment opportunities.

◆ A **slump** (or trough) occurs at the bottom of a recession in the business cycle when aggregate demand remains low.

◆ A **recovery** occurs when GDP starts to rise after the trough in the business cycle, eventually leading to economic growth.

Economic growth occurs when the level of economic activity rises, as measured by the value of real GDP. Technically, growth occurs when the economy's real GDP increases for two consecutive quarters. Hence, growth leads to lower unemployment although it also causes inflation to rise. Nevertheless, economic growth is generally regarded as desirable as it suggests the average person earns more income and the economy is more prosperous.

Figure 16.3 also illustrates the potential national output (potential real GDP) of an economy, as indicated by the long-term trend line in the business cycle. **Potential output** refers to the highest level of real gross domestic product that can be sustained over the long-term. However, severe exogenous shocks can distort the trend line of an economy's business cycle. Such events include the outbreak of major infectious diseases (such as the worldwide outbreak of the novel coronavirus beginning in 2019), natural disasters (such as earthquakes, tsunamis and severe flooding or droughts), and global financial crises.

◆ **Potential output** refers to the possible level of real GDP of an economy, as shown on its production possibility curve, if all resources are used efficiently. It is a long run phenomenon represented by an outwards shift of the PPC.

CASE STUDY

Recession in Europe

Following the global financial crisis of 2008, the economies of Portugal, Ireland, Greece and Spain faced a major economic downturn in their business cycles. By 2013, both Greece and Spain experienced record unemployment rates of 26.8 per cent and 26.6 per cent respectively. Spain's youth unemployment rate (those aged 21 and below) reached 55 per cent – the highest in the history of the European Union. Under such extreme economic circumstances, the government needs to intervene to help the economies to recover. All four economies received financial bailouts of the following amounts:

- Portugal = $114bn
- Ireland = $113bn
- Greece = $159bn
- Spain = $130bn.

Top tip!

Not all components of aggregate demand necessarily fall during a recession – a boost in government spending (G) may be needed to help the economy recover from the recession (such as the fiscal stimulus policies used by many governments during financial crisis of 2008 and the coronavirus pandemic).

Top tip!

The best-performing students are able to show skills of evaluation and critical thinking. Not all businesses suffer during a recession. Counter-cyclical businesses are those that do well and survive during an economic downturn, such as suppliers of inferior goods (for example, fast food restaurants and discount stores) and pawnbrokers.

Common mistake

You should be aware of the difference between a *fall in GDP* and a *fall in GDP growth* – a fall in GDP (over two consecutive quarters) causes a recession whereas a fall in GDP growth means the economy is still growing, only at a slower rate than before.

Theory of Knowledge

Given the number of variables that contribute to economic activity, how confident can we be in the use of economic models? How do we decide (know) what to include or exclude in economic models?

Appropriateness of using GDP or GNI statistics to measure economic well-being (AO3)

The use of GDP and/or GNI statistics enables economists to use national income statistics for making comparisons of economic well-being over time, as well as making comparisons between countries. National income statistics give an indication of standards of living in a country, because a higher real GDP per capita is associated with higher standards of living for the average person.

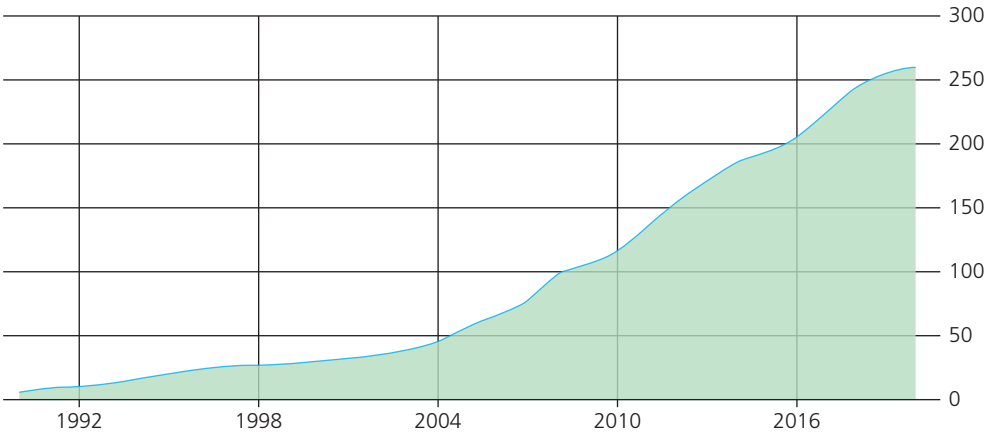
Figure 16.4 shows that Vietnam's GDP has risen from around $6.7bn in 1990 to over $260bn by 2020, that is, an increase of about 3,780 per cent in just 30 years. This suggests that the average person in Vietnam has enjoyed a significantly better standard of living in the past three decades.

This is because a person's standard of living is directly dependent upon the amount of goods and services that individual is able to consume.

International comparisons of nominal and real GDP and/or GNI per capita figures can also reveal varying standards of living between countries. Figure 16.5 compares the GDP per capita in Iceland and Honduras, between 2000 and 2020. Note that although the trend is very similar, the absolute difference is significant. Iceland, one of the world's most economically developed countries, had a GDP per capita of $37,465 in 2000, which increased to around $52,105 in 2020. By contrast, Honduras is a low-income country and had a GDP per capita of only $1,597 in 2000 and climbing up to just above $2,200 twenty years later. Note that although the GDP per capita in both countries rose by approximately the same percentage, the GDP is significantly higher in Iceland. This suggests the average person in Iceland enjoyed a higher standard of living throughout the time period under consideration.

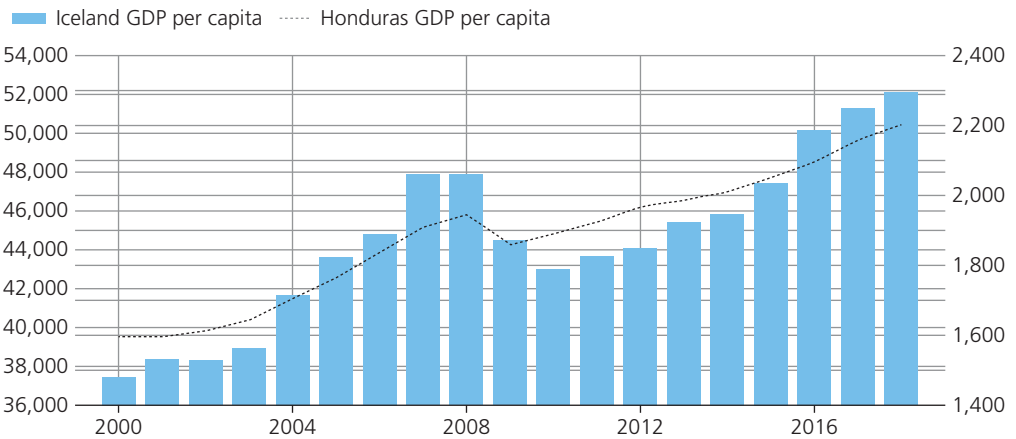
Theory of Knowledge

To what extent can standard of living be meaningfully measured according to the amount of goods and services consumed?



Source: Trading Economics **https://tradingeconomics.com/vietnam/gdp**

Figure 16.4 Vietnam's GDP, $bn (1990–2020)



Source: Trading Economics **https://tradingeconomics.com/honduras/gdp-per-capita** *and* **https://tradingeconomics.com/iceland/gdp-per-capita**

Figure 16.5 GDP per capita in Iceland and Honduras, $ (2000–20)

However, there are several limitations in using national income statistics to measure standards of living within a country and to make international comparisons with the data. These limitations include:

- GDP and GNI figures do not show the distribution of income and wealth in the country. This applies whichever measure is used (nominal, real, or per capita).
- The data does not reveal differences in the varying rates of direct and indirect taxes between countries. For example, Saudi Arabia, the UAE, Oman, Qatar and the Cayman Islands have a zero rate of income tax, whereas some countries impose more than 53 per cent income tax, such as in Belgium, Denmark, Japan and Sweden.
- Differences in the cost of living are ignored. For example, while real GDP per capita in Singapore ($58,248) and Hong Kong ($38,785) are quite high by international standards, the costs of housing, education and healthcare are extremely high. Hong Kong, for example, has the highest real estate prices in the world, with the average property price exceeding the average annual salary by 22 times.
- Exchange rate fluctuations can make international comparisons of GDP and GNI over time less meaningful. This is especially the case when comparing the data between high-income and low-income countries because the latter tend to have weaker and more volatile currencies. Such fluctuations in the exchange rate can distort historical comparisons and interpretations of well-being between countries.
- The extent of social welfare benefits will vary between countries. For example, the UK's National Health Service ensures every citizen has access to healthcare services, even if they cannot pay for it. By contrast, healthcare is not necessarily affordable in other countries, although is it vital for physical and mental well-being. The Scandinavian countries typically spend in excess of 25 per cent of their GDP on public social welfare programmes, thus boosting well-being in these nations.
- Similarly, the GDP and GNI figures of different countries do not reveal the composition of national output. For example, some countries such as North Korea spend a significant proportion of their national income on armaments and national defence whereas other countries such as New Zealand, Norway and Luxembourg prefer to spend a larger proportion on education and their social welfare system.
- The figures for GDP and GNI can be inaccurate (unreported or underreported) owing to economic activities in the hidden or shadow economy. Variations in the size of the shadow economy (for goods and services that are not officially traded) make it challenging to compare national output data between different countries. According to the IMF, Mexico, Italy and the Republic of Korea have large hidden economies, estimated to account for over 25 per cent of each of these country's GDP. This is, however, insignificant compared to some other countries (see Table 16.1).
- GDP and GNI data do not show the extent to which national output generates negative externalities (external costs) on society, thereby limiting the sustainability of economic activity and the quality of life in the country. Chapters 11 and 12 show how greater levels of economic activity can be detrimental to social, economic and ecological sustainability in the long run.

Table 16.1 Shadow economy output as a percentage of GDP (1991 to 2015)*

Country	*Average percentage
Georgia	64.87
Bolivia	62.28
Nigeria	56.67
Guatemala	54.74
Benin	53.66
Haiti	53.28
Gabon	52.43
Peru	52.40
Azerbaijan	52.19
Myanmar	51.39

Source: IMF (www.imf.org/en/Publications/WP/Issues/2018/01/25/Shadow-Economies-Around-the-World-What-Did-We-Learn-Over-the-Last-20-Years-45583)

Figure 16.6 GDP and GNI data ignore the external costs of economic activity and growth

Hence, alternatives to nominal or real GDP and GNI are needed to calculate the external costs of economic activity and growth, such as the use of **Green GDP**. This is a measure of GDP (or GNI) that accounts for environmental destruction of economic activity by deducting the environmental costs associated with the output of goods and services. Assuming that external costs can be accurately measured, green GDP is calculated by using the formula:

◆ **Green GDP** adjusts the country's GDP to take account of the value of environmental degradation, damage and destruction associated with economic growth.

Green GDP = Nominal GDP – Environmental production costs

Owing to the limitations of using GDP and GNI figures to measure a society's well-being, alternative measures are needed. These alternatives are explained in the section below.

WISE ChoICES

To what extent does increased economic activity improve **economic well-being** at the cost of **sustainability**?

Alternative measures of well-being (AO2)

Nobel Prize-winning American economist Joseph E Stiglitz (born 1943) is among several economists who argue that GDP fails to account for the impact of well-being beyond economics. For example, economic growth often comes at the opportunity cost of greater inequalities in income and wealth, or damage to the environment in terms of pollution, destruction of natural ecosystems and climate change – all of which are detrimental to human well-being. The IB Economics syllabus refers to three alternative composite indicators of measuring well-being: the OECD Better Life Index, the Happiness Index and the Happy Planet Index.

The OECD Better Life Index (BLI)

In May 2011, the Organization for Economic Co-operation and Development (OECD) launched the **Better Life Index** (BLI) to compare well-being across countries. The BLI is based on 11 topics identified by the OECD to be essential in terms of *material living conditions* (items 1–3) and *quality of life* (items 4–11):

1. Housing – housing conditions and expenditure (for example, the cost of real estate).
2. Income – household disposable income and net financial wealth.
3. Jobs – earnings, job security and unemployment.
4. Community – quality of social support network.
5. Education – quality of education.
6. Environment – quality of the environment (environmental health).
7. Governance – involvement in democracy.
8. Health – quality of healthcare.
9. Life satisfaction – level of happiness.
10. Safety – murder and assault rates.
11. Work–life balance – hours of work versus hours of leisure.

◆ The **Better Life Index** (BLI) is an alternative to national income as a measure of well-being, based on a set of 11 criteria identified by the OECD to be essential in terms of material living conditions and quality of life.

The BLI does not rank these 11 measures of well-being in any particular way, so making like-for-like comparisons can be challenging. Although rather complex and somewhat subjective, the BLI enables the policymaking process to be more informed about the various factors that shape our lives and human well-being.

The BLI covers the 35 members of the OECD (which accounts for most of the world's economically developed economies) plus a few emerging economies and key partners. However, this means the BLI is rather narrow in its coverage, which limits the use of the BLI when trying to make international comparisons of well-being.

Theory of Knowledge

Humans have to work to live, but do not generally aspire to live to work. How can we know what true work–life balance really is?

The Happiness Index

The **Happiness Index** was created in April 2012 by the United Nations Sustainable Development Solutions Network, which is published in the annual World Happiness Report (WHR). Compilation of the index requires the use of representative samples of respondents in a country, who are asked to rate their own current lives using different criteria on a scale of 0 (least happy) to 10 (most happy). The survey uses 14 indicators of well-being: (1) business and economic, (2) citizen engagement, (3) communications and technology, (4) diversity (social issues), (5) education and families, (6) emotional well-being, (7) environment and energy, (8) food and shelter, (9) government and politics, (10) law and order (safety), (11) health, (12) religion and ethics, (13) transportation, and (14) work (employment). Economists and social scientists believe that happiness leads to greater productivity, performance and profits as well as peace, progress and purpose in life.

◆ The **Happiness Index** is an alternative to national income as a measure of well-being, by considering how information technology, governance and social norms influence communities and their level of well-being.

Table 16.2 Happiness Index, World Happiness Report (2019)

Rank	Country	HI score
1	Finland	7.769
2	Denmark	7.600
3	Norway	7.554
4	Iceland	7.494
5	Netherlands	7.488
6	Switzerland	7.480
7	Sweden	7.343
8	New Zealand	7.307
9	Canada	7.278
10	Austria	7.246

Source: https://worldhappiness.report/ed/2019/

One strength of the Happiness Index is that the index is updated to reflect changes in what contributes to well-being. Happiness changes over time, so the Happiness Index currently reflects how changes in information technology, governance (links between government and happiness) and social norms (pro-social behaviour) influence communities and their level of well-being. However, this also means that it can be difficult to make historical comparisons of well-being by using this measure.

Theory of Knowledge

What is happiness? Use the concepts of emotion and sense perception to discuss this question.

Figure 16.7 How do we know what makes us happy?

The Happy Planet Index

The **Happy Planet Index** (HPI) is a measure of human well-being that was introduced by the New Economics Foundation in 2006. The HPI is used to measure the extent to which individuals and countries are able to achieve long, happy and sustainable lives, so emphasizes the impact of economic activity on the environment by giving higher scores to countries with lower ecological footprints. The HPI is an important measure of human well-being as it considers both the current and future prosperity of individuals and societies. Hence, it emphasizes the importance of happy, healthy and sustainable lives without jeopardizing the planet.

◆ The **Happy Planet Index** is a measure of sustainable human well-being, that is, how individuals and countries are able to achieve long, happy and sustainable lives.

The HPI is calculated by combining four elements to show how efficiently individuals and countries use their environmental resources to achieve long, happy lives:

1. **Well-being** – how citizens of a country feel about their own life overall, on a scale of 0 (worst) to 10 (best), based on representative research data conducted as part of the Gallup World Poll.
2. **Life expectancy** – the average number of years a person is expected to live for in each country, based on data gathered from the United Nations.
3. **Inequality of outcomes** – a measure of the inequalities of people within a country in terms of how long they live and how happy they feel, expressed as a percentage figure. This is based on the distribution of the life expectancy and well-being data for each country.
4. **Ecological footprint** – a measure of the impact that the average person in the country has on the environment, based on data from the Global Footprint Network. The measure is expressed in terms of global hectares (gha) per person.

Source: **https://happyplanetindex.org/about**

Figure 16.8 The Happy Planet Index (HPI)

Hence, the HPI is calculated as being equal to (Well-being × Life expectancy × Inequality) ÷ Ecological footprint.

Data show that high-income countries underperform in the HPI rankings. Although they may succeed in terms of high real GDP per capita, several countries in Latin America and Asia Pacific achieve higher HPI scores owing to their higher life expectancy, improved human well-being, and smaller ecological footprints (such as carbon emissions and damage to the natural environment). Nevertheless, critics argue that the HPI ignores some major problems affecting people's human well-being and happiness across the world, such as abuse and violation of human rights in some of the countries that rank high in the Happy Planet Index (see Table 16.3).

Table 16.3 The happiest countries according to the HPI, 2016

HPI rank	Country
1	Costa Rica
2	Mexico
3	Colombia
4	Vanuatu
5	Vietnam
6	Panama
7	Nicaragua
8	Bangladesh
9	Thailand
10	Ecuador

Source: **http://happyplanetindex.org/s/hpi-data-2016.xlsx**

CASE STUDY

Well-being and sustainable growth

Iceland, Scotland and New Zealand are among the first countries to include well-being as part of the official government budget, with emphasis on environmental and family-friendly priorities, rather than focusing only on traditional GDP data. Scientists argue that Iceland's Okjokull glaciers, which were declared 'dead' in 2014, provides a clear example and evidence of global warming, caused mainly by excessive carbon emissions. Government intervention and international co-operation are vital to tackle the global problem of climate change. As Joseph E Stiglitz said, '*The reason that the invisible hand often seems invisible is that it is often not there.*' Such a shift in thinking and behaviour would enable countries to focus on an alternative and brighter future, based on well-being and sustainable and inclusive growth.

PAPER 1 EXAM PRACTICE QUESTION 16.10

Explain the difference between the Better Life Index (BLI) and the Happy Planet Index (HPI). *[10 marks]*

Chapter summary

- Economists use national income accounting to measure the level of economic activity in a country over time.
- The output method of calculating economic activity adds up the final value of newly produced goods and services during the year.
- The expenditure method of calculating economic activity adds up the total spending on newly produced goods and services during the year.
- The income method of calculating economic activity adds up the total value of all factor incomes earned during the year.
- The business cycle (trade cycle) describes the fluctuations in economic activity in a country over time. These fluctuations create a long-term trend of growth in the economy.
- Economic growth occurs when there is a sustained increase in the level of economic activity in a country.
- National income statistics, which reveal the level of economic activity in the country, directly affect the average person's social and economic well-being.
- GNI and GDP both reflect the national output and income of an economy. The main difference is that GNI considers net factor incomes from abroad, that is, GNI = GDP + Net factor income from abroad.
- The values of real GDP and real GNI have inflation adjusted for by using a GDP deflator, thereby expressing real values at constant market prices.
- A GDP deflator (or price deflator) is used to convert GDP at current prices to GDP at constant prices. This also applies to GNI calculations.

Theory of Knowledge

How do we value things other than financially? How can we compare different types of value? Can we put a price on happiness?

17 Variations in economic activity – aggregate demand and aggregate supply

Real-world issue

Why does economic activity vary over time and why does this matter?

Conceptual understandings (WISE ChoICES)

- **Change** in the conditions of the demand and supply sides of the economy cause economic activity to vary over time.
- Fluctuations in economic activity impact the **economic well-being** of individuals and societies.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- aggregate demand (AD) and the components of AD (AO2)
- an aggregate demand curve diagram (AO4)
- the determinants of the different components of AD (AO2)
- shifts of the AD curve caused by changes in the determinants of AD (AO2)
- diagrams that show shifts of the AD curve (AO4)
- short run aggregate supply (SRAS) and the determinants of SRAS (AO2)
- a short run aggregate supply diagram (AO4)
- shifts of the SRAS curve (AO2)
- diagrams that show shifts of the SRAS curve (AO4)
- alternative views of aggregate supply (AS) with reference to the monetarist/new classical view and Keynesian view of the AS curve (AO2)
- the difference between inflationary and deflationary/recessionary gaps (AO2)
- shifts of the AS curve over the long run or over the long-term (AO2)
- diagrams that show shifts of the LRAS or Keynesian AS curves (AO4)
- macroeconomic equilibrium, with reference to short run equilibrium, equilibrium in the monetarist/new classical model, and equilibrium in the Keynesian model (AO2)
- appropriate diagrams to show macroeconomic equilibrium in both the short run and long run (AO4)
- the assumptions and implications of the monetarist/new classical and Keynesian models (AO3).

Aggregate demand (AO2, AO4)

In microeconomics, **demand** refers to the willingness and ability of consumers to pay for a particular product at each price level (see Chapter 3). In macroeconomics, **aggregate demand** (AD) refers to the value of total demand for all goods and services in the economy, per time period. It is calculated as the total amount of spending in the economy, usually per year. The AD curve shows the real national output that is purchased at each price level, per time period. It has a negative slope (that is, it is downward sloping) because when the general level of prices is high, the level of aggregate demand tends to be low (see Figure 17.1).

◆ **Aggregate demand** is the value of all goods and services demanded in the economy, per time period. It is calculated as the total amount of spending in the economy, usually per year.

Figure 17.1 shows that the higher the general price level, the lower the level of aggregate demand tends to be, ceteris paribus. There are three main reasons for the inverse relationship between the general price level and the real GDP of the economy:

- **The Pigou wealth effect** – English economist Arthur Pigou said that for any given nominal value of income, a lower price level enables households, firms and the government to have greater purchasing power, which results in greater consumption, investment and government spending.
- **Keynes's interest rate effect** – A fall in the general price level causes interest rates to drop, thus boosts the demand for money, ceteris paribus. This results in greater consumption, investment expenditure and government spending, that is, higher aggregate demand in the economy.
- **Mundell-Fleming's exchange rate effect** – As the general price level falls, the interest rate also tends to fall resulting in a depreciation of the exchange rate (see Chapter 31). This will tend to increase the demand for net exports because domestic products are cheaper. Hence, higher net exports will raise the level of AD, ceteris paribus.

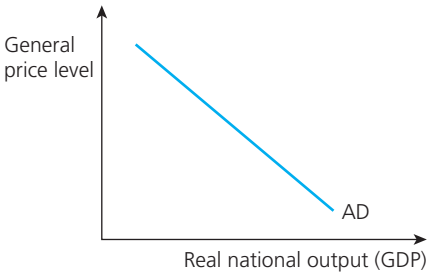


Figure 17.1 The aggregate demand curve

Top tip!

Mundell and Fleming are not explicitly mentioned in the IB Economics Guide, so you may prefer to refer to the three key reasons for the downward sloping AD curve as: the wealth effect, the interest rate effect and the exchange rate effect.

Components of aggregate demand (AO2)

The components of aggregate demand are consumption expenditure (C), investment expenditure (I), government spending (G), exports earnings (X) and import expenditure (M). Therefore:

$$AD = C + I + G + (X - M)$$

- **Consumption** (C) is the total spending on goods and services from households in the domestic economy, per time period. It is the single largest component of aggregate demand.
- **Investment** (I) is the expenditure of all firms in the economy in order to increase their capital stock and production capacity; for example, the purchase of fixed assets such as machinery, commercial vehicles and buildings. This results in the economy having a larger productive capacity in the long run.
- **Government spending** (G) is the total expenditure on goods and services by the public sector; for example, government expenditure on education, healthcare, national security and social welfare schemes. For example, in 2018, the governments of mainland China, Hong Kong and Macau opened the Hong Kong-Zhuhai-Macau Bridge to facilitate greater trade opportunities in the Pearl River Delta. The 55 kilometres (34 miles) bridge is designed to last 120 years, at a cost of 127 billion yuan ($18.8 billion) to build.
- **Net exports** (X – M) measures the difference between the value of the country's export earnings and its import expenditure in the year.

Common mistake

Gross domestic product (GDP) is often confused with aggregate demand (AD). While they are similar concepts, GDP refers to the value of *actual* expenditure in the economy over the year, whereas AD refers to the *planned* (or the expected) expenditure in a year.

Determinants of AD components (AO2) and shifts of the AD curve caused by changes in the determinants (AO2, AO4)

The components of aggregate demand are affected by different but interrelated factors, as outlined in this section. Factors that cause a change in any component of aggregate demand (AD = C + I + G + X – M) will cause a shift in the AD curve, ceteris paribus. The determinants of aggregate demand refer to the factors that cause a shift in the AD curve.

Consumption (C)

Factors that cause a change in consumption include: consumer confidence, interest rates, wealth, income taxes, level of household indebtedness and expectations of the future price level.

Consumer confidence

The more confident consumers are about the economy, the greater the level of consumption will be. Consumer confidence in the economy is low during a recession (when unemployment increases) and high during a boom (when there is economic growth).

In Figure 17.2, an increase in consumer confidence shifts the aggregate demand curve rightwards, from AD_1 to AD_2, thereby raising national output (real GDP) at all price levels, ceteris paribus.

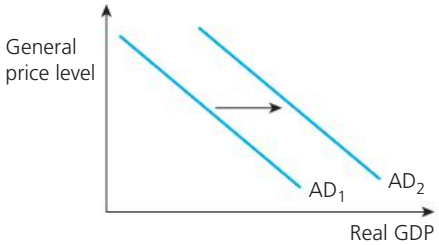


Figure 17.2 Rightwards shift of the aggregate demand curve

Interest rates

Interest rates refer to the price of money, in terms of the cost of borrowing money or the return for saving money. An increase in interest rates tend to reduce the level of consumption because households will find it less affordable to borrow money to finance their expenditure. In addition, households with existing loans and mortgages have higher loan repayments when interest rates increase, so have less income at their disposal for consumption. Furthermore, if interest rates rise, it could give an incentive for some people to save more money. Hence, higher interest rates tend to reduce the level of consumption as shown by a leftwards shift of the AD curve, and vice versa.

Wealth

Changes in household wealth have a direct impact on the level of consumption in the economy. There is a positive correlation between household wealth and the value of real GDP, that is, the wealthier households are, the more they tend to consume. Hence, an increase in the wealth of the average household in the economy will shift the aggregate demand curve to the right, from AD_1 to AD_2 in Figure 17.2.

Income taxes

If the level of disposable income falls due to higher income taxes, the level of consumption will also fall, ceteris paribus. Disposable income refers to the earnings after taxes have been accounted for, that is, the actual take-home income that workers are able to spend. For example, a reduction in income tax rates will stimulate consumption expenditure and shift the AD outwards from AD_1 to AD_2 in Figure 17.2, ceteris paribus.

Level of household indebtedness

Indebtedness is the condition of owing money, that is, households with debts. A lower level of household indebtedness can be brought about by reduced interest rates (which reduces the amount of debt interest owed by households) or by higher household incomes (perhaps due to better employment opportunities in the economy). Hence, a lower level of household indebtedness will shift the AD curve to the right, ceteris paribus.

Expectations of the future price level

Expectations refer to future expected price movements of goods and services. Household expectations about the future price level will also affect domestic consumption levels. For example, if households expect the price level to increase in the future, they will tend to increase consumption today, especially on the purchase of durable goods. These are products that are not perishable but can be stored (such as batteries or canned foods) or used repeatedly over a long period of time (such as furniture or home appliances). This would tend to shift the AD curve outwards to the right. By contrast, if households expect price levels to fall in the near future, they will tend to postpone consumption today as they expect goods and services will be less expensive in the future. Hence, this would cause the AD curve to shift to the left, ceteris paribus.

Investment (I)

Factors that cause a change in investment expenditure in the economy include: interest rates, business confidence, technology, business taxes and level of corporate indebtedness.

Interest rates

Higher interest rates make it more expensive for firms to borrow money to finance their growth, so investment expenditure will tend to fall. Firms with existing loans will usually face larger debt repayments when interest rates increase due to higher debt interest. Therefore, higher interest rates will reduce investment expenditure, thereby shifting the AD curve to the left (see Figure 17.3). By contrast, if interest rates fall, the AD curve will shift rightwards owing to the fall in the cost of borrowing funds to invest (see Figure 17.2).

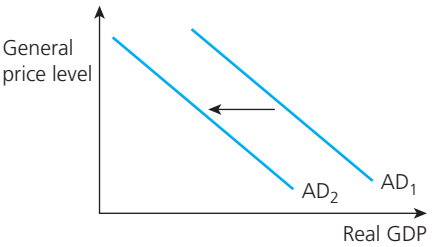


Figure 17.3 Leftwards shift of the aggregate demand curve

Business confidence

In general, the greater the level of business confidence in the economy, the higher the level of investment will be. Business confidence is high when the economy is growing, as national income and employment are both rising. Hence, this would shift the AD outwards to the right. By contrast, business confidence is low when the economy is in a recession or a slump, so investment will fall. All things being equal, this would cause the AD curve to shift inwards.

Technology

Technological progress and the associated productivity gains from this will tend to boost the level of investment expenditure. This is because in highly competitive markets, both domestically and internationally, firms need to keep up to date with advances in technology. Hence, technological progress such as 5G technologies will increase the level of investment in the economy. This will increase aggregate demand, ceteris paribus.

Business taxes

The lower the rate of business taxes in the economy, the more attractive investment becomes because firms are more likely to be able to earn a return on their investment. Some countries such

Figure 17.4 Investments in 5G technologies will help to boost productivity

as The Bahamas and Estonia have a zero rate of corporation tax to attract foreign direct investment. Hence, lowering business taxes will tend to shift the AD curve to the right, and vice versa.

Level of corporate indebtedness

Like households, the more debts that businesses have the less money they have available for investment expenditure. Indebtedness tends to increase during periods of rising interest rates (because firms with existing loans have higher debt interest repayments) or during an economic downturn when firms struggle to survive (so may have borrowed more money to finance their operations). Hence, lower corporate indebtedness will shift the AD curve to the right, ceteris paribus.

Government spending (G)

Factors that cause a change in government expenditure in the economy include political and economic priorities.

Political priorities

Government spending will vary depending on the political priorities. For example, there may be more public sector spending on education and healthcare prior to a general election (in order to win political votes). A government will also increase its expenditure on social welfare during periods of high unemployment. Such measures will tend to increase aggregate demand (see Case Study on page 248).

Economic priorities

The global financial crisis of 2008 and the coronavirus (Covid-19) outbreak that began in 2019 required government intervention and immense amounts of government spending to support the economy. Hence, this would boost the level of aggregate demand. Following such major economic crises, austerity measures may be needed to reduce the government budget deficit

ATL ACTIVITY 17.1

Investigate the priorities that your government or a government of your choice is currently facing. Discuss the extent to which these priorities are reflected in the government's annual budget. Use both quantitative and qualitative analysis to present your answer.

and national debt, such as through cuts in government spending and tax increases. These austerity measures would then reduce the level of aggregate demand in the economy, ceteris paribus.

CASE STUDY

The UK 'coronavirus' budget 2020

In an attempt to prevent a major economic slowdown caused by the global coronavirus outbreak, the UK government announced a £350 billion ($390bn) stimulus package to support the British economy. The National Health Service was given £5 billion ($6.5bn) emergency funding with a guarantee of additional funding if needed. Employees forced to self-isolate qualified for statutory sick pay from the first day away from work. Businesses were also reimbursed for their losses during the crisis and the government made it easier for firms to access bank loans. The Bank of England also cut interest rates from 0.75 per cent to a historical low of 0.25 per cent as part of the emergency measures in response to the economic impacts of the coronavirus crisis. A week later, this was cut to an all-time low of 0.1 per cent. Despite going into further debt, the government's priority at the time was to spend money to help households and firms and to support confidence in the economy during the worldwide pandemic.

Net exports (X – M)

Factors that cause a change in net export expenditure include: income of trading partners, exchange rates and trade policies.

Income of trading partners

Owing to the forces of globalization and interdependence, when a country such as the USA or China suffers from an economic downturn, there are negative impacts on its trading partners. Essentially, a fall in the income of trading partners leads to fewer export sales and therefore a decline in aggregate demand, ceteris paribus. The opposite is true if the income of trading partners increases.

Exchange rates

Exchange rates refer to the price of one currency in terms of the currencies of its trading partners (see Chapter 31). A higher exchange rate tends to increase the demand for imports as the local currency can buy more units of foreign currencies, so imported products become relatively more affordable. By contrast, it reduces the demand for exports as they become more expensive for foreign buyers. Hence, higher exchange rates will usually reduce aggregate demand as import expenditure increases while export earnings fall. The opposite applies when exchange rates fall.

Trade policies

Trade policies refer to the measures a government takes to regulate or liberalize international trade. For example, changes in the level of trade protection (see Chapter 28) have a direct impact on the volume of exports and imports. Imposing more trade barriers, such as tariffs and/or quotas will cause the price of imports to rise, so will tend to reduce the demand for foreign goods and services but makes exports more appealing. Hence, restrictive trade policies, such as raising tariffs and quotas, will tend to increase aggregate demand in the domestic economy (since X > M). By contrast, the AD curve in the economy will generally shift to the left if import expenditure exceeds export earnings (M > X).

Short run aggregate supply (SRAS) curve and determinants of the SRAS curve (AO2, AO4)

Aggregate supply (AS) is the total amount of output of goods and services that firms within an economy are willing and able to supply at a given time and at an overall price level. It is therefore a measure of the economy's potential output (total supply). The **short run aggregate supply** (SRAS) curve shows the total planned national output at different price levels, holding productivity and costs of factor inputs constant. This means that in the short run, wage rates and the state of technology are assumed to be fixed.

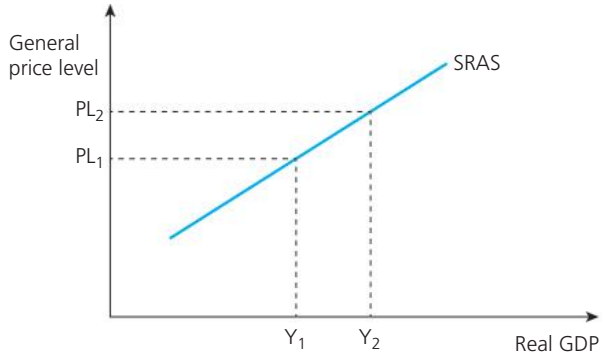


Figure 17.5 The SRAS curve

The SRAS curve is upward sloping, as shown in Figure 17.5, because higher prices attract more firms in the economy to raise their output level (recall from microeconomics that the supply curve for a single firm is upward sloping for the same reason). In the short run, it is assumed that capital is fixed, so firms can only alter variable factors of production, such as labour in order to increase output. The SRAS is therefore relatively price elastic in the short run as firms can increase output by getting employees to work overtime.

◆ **Aggregate supply** (AS) is the total amount of output of goods and services that firms within an economy are willing and able to supply at a given time and at an overall price level.

◆ The **short run aggregate supply** (SRAS) curve shows the total planned national output at different price levels, ceteris paribus.

Top tip!

The slope of the SRAS curve depends on the extent to which there is spare capacity in the economy. The flatter the SRAS curve, the greater that degree of spare capacity (underutilized resources). This means the SRAS curve is relatively price elastic, vice versa.

The main determinants of the SRAS curve can be classified as either being costs of factors of production or indirect taxes.

Costs of factors of production

Costs of production are the expenses that firms face in producing goods and services, such as the price of raw materials and costs of labour. Examples of variables that affect the costs of factors of production, include:

- **Labour costs** – Wages and salaries account for a significant proportion of costs for most firms, especially those that are labour intensive. Hence, a relatively small change in wage costs (such as an increase in the national minimum wage) will tend to have a larger impact on the overall production costs of firms. In general, higher labour costs will reduce the level of production in the short run owing to higher production costs for firms.
- **Raw material costs** – An increase in the cost of raw materials or components (such as crude oil, plastics, timber or metal ores) will increase the cost of production for all firms in the industry or even across the economy. For example, an oil crisis which results in considerably higher oil prices will increase the cost of transportation throughout the economy.
- **Exchange rate** – A rise in the exchange rate (currency appreciation) means that domestic firms can buy imports at a lower price. If domestic firms need to import raw materials,

for example, the currency appreciation helps to reduce their costs of production, thereby helping to increase short run aggregate supply. Note however that a currency appreciation will also make exports less competitive (although this directly impacts on the demand for exports). The opposite is true for a currency depreciation, that is, a fall in the value of the exchange rate causes higher production costs for firms reliant on purchasing imported materials and components for their own production.

- **Interest rates** – Firms with existing borrowing will have fewer funds to invest if interest rates increase, as the cost of debt interest will rise. Similarly, firms will find it less affordable to take out new loans to finance their investment projects. Hence, higher interest rate will tend to reduce the level of aggregate supply. The opposite applies if interest rates are reduced as this cuts the costs of loan repayments.
- **Bureaucracy and administration** – Strict legal procedures and policies mean that firms have to comply with rigorous and complicated bureaucratic and administrative matters. Domestically, this could mean abiding by stringent health and safety laws, while firms that trade internationally may have to deal with non-tariff barriers (see Chapter 28). Hence, increased levels of bureaucracy and administration costs will tend to reduce short run aggregate supply. By contrast, deregulation of such bureaucratic and administrative barriers can encourage firms to produce more output, thereby increasing the level of SRAS.

Indirect taxes

Indirect taxes are government levies or charges on expenditure, rather than on incomes (see Chapter 4). Most countries impose Value Added Tax (VAT) or a Goods and Services Tax (GST). As they are imposed on consumption or expenditure, these indirect taxes inevitably increase the price paid by customers. Indirect taxes can also include environmental taxes such as carbon duties (see Chapter 11). Note that indirect taxes are imposed on the supplier of a product, which basically adds to the costs of production, even if the producer can pass on some of the tax to consumers in the form of higher prices for goods and services. The indirect taxes reduce the profitability of firms, so tend to reduce aggregate supply. This is illustrated by an inwards (leftwards) shift of the SRAS curve at given general price levels.

Country	Sales tax rate (%)
Bhutan	50%
Hungary	27%
Denmark	25%
Cuba	20%
India	18%
Mexico	16%
South Africa	15%
Ghana	12.5%
Indonesia	10%
Sri Lanka	8%

Source: **Trading Economics https://tradingeconomics.com/country-list/sales-tax-rate**

Top tip!

Subsidies are not explicitly mentioned in the syllabus as a determinant of the SRAS curve, partly because most subsidies have a microeconomic impact, rather than a macroeconomic effect on SRAS. Nevertheless, they can have the opposite effect of indirect taxes. Subsidies offered to producers in various industries help to reduce their production costs, which can help to shift the SRAS curve outwards to the right. For example, governments around the world offered subsidies for businesses in an attempt to save them from the lockdowns due to the coronavirus pandemic in 2020. In theory, the impact of this is a lower general price level and an increase in national output, ceteris paribus.

Shifts of the SRAS curve (AO2, AO4)

A change in any of the main determinants of short run aggregate supply will cause the SRAS curve to adjust, that is, a change in the costs of factors of production or indirect taxes will shift the SRAS curve.

In the short run, adverse changes in non-price factors that affect aggregate supply will shift the SRAS curve to the left, from $SRAS_1$ to $SRAS_2$ (see Figure 17.6), such as higher indirect taxes or higher

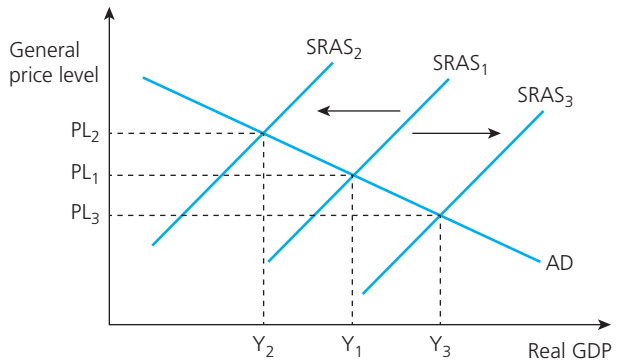


Figure 17.6 Shifts in the SRAS curve

production costs. This would then lead to higher prices in the economy, from PL_1 to PL_2, and cause a contraction in aggregate demand from Y_1 to Y_2.

By contrast, favourable changes in non-price factors that affect aggregate supply will shift the SRAS curve to the right, from $SRAS_1$ to $SRAS_3$, as shown in Figure 17.6. This occurs in the case of lower indirect taxes and/or lower production costs. This would then lead to lower average prices in the economy, from PL_1 to PL_3, and subsequently cause an expansion in aggregate demand from Y_1 to Y_3.

Note that the slope of SRAS curve depends on the degree of spare capacity within the economy – in general, the steeper the SRAS curve, the lower the degree of spare capacity (underutilized resources) there is in the economy. For example, as national output expands and the economy heads towards its productive capacity, supply bottlenecks start to appear, so SRAS becomes more price inelastic.

Top tip!

Although not explicitly referred to in the IB Economics syllabus, **adverse supply-side shocks** can also shift the SRAS curve to the left, thus causing higher prices (inflation) and lower national output. Adverse supply-side shocks are unexpected increases in costs or major disruptions to production, such as the outbreak of wars, financial crises, natural disasters, major oil shortages or the outbreak of an epidemic (such as swine flu, malaria, HIV/Aids, and dengue fever) will shift the SRAS curve to the left.

WISE ChoICES

Discuss how **changes** in the conditions of aggregate demand and aggregate supply have affected the level of economic activity in a country of your choice, and determine the extent to which it has impacted the **economic well-being** of individuals and societies in the country.

Top tip!

Make sure you can distinguish between a **shift** and a **movement** in the SRAS curve. A shift is caused by any change in the determinants of SRAS (costs of factors of production and indirect costs).

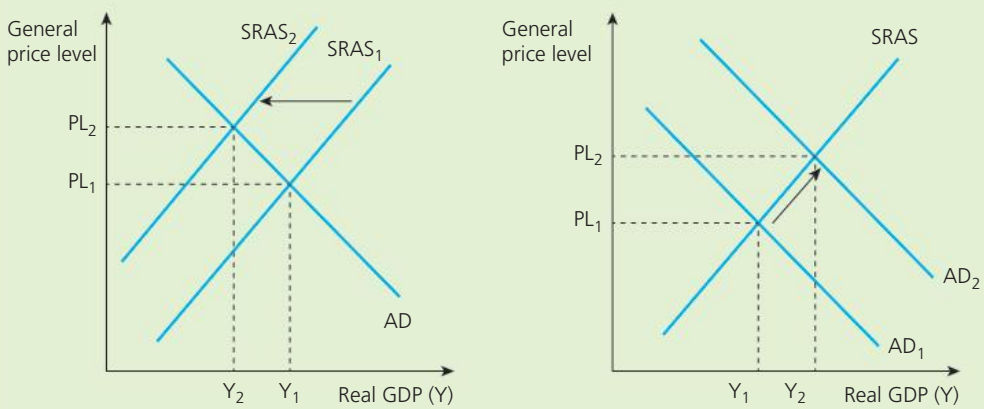


Figure 17.7 a) A shift in the SRAS curve to the left; b) movement along the SRAS curve

In diagram a), the SRAS has shifted to the left from $SRAS_1$ to $SRAS_2$ owing to higher costs of production and/or indirect taxes. In diagram b), a higher level of aggregate demand, perhaps due to higher household incomes, causes a shift of the AD curve from AD_1 to AD_2. This leads to a higher price level (from PL_1 to PL_2) and therefore causes a movement along the SRAS curve.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 17.1

Explain **two** factors that can shift the short run aggregate supply (SRAS) curve. *[4 marks]*

PAPER 2 EXAM PRACTICE QUESTION 17.2

Using an appropriate diagram, explain the impact of an increase in business taxes on short run aggregate supply. *[4 marks]*

PAPER 2 EXAM PRACTICE QUESTION 17.3

a With the use of an appropriate AD–AS diagram, explain the impacts on the domestic economy of an increase in net exports. *[4 marks]*

b With the use of an appropriate AD–AS diagram, explain the impact of a major global financial crisis on the domestic economy. *[4 marks]*

Top tip!

Different schools of macroeconomic thought, such as Keynesians and monetarists, identify different causes and offer different solutions for macroeconomic problems. Make sure you are able to understand how these groups of economists vary in their views on the causes and solutions for macroeconomic problems.

Alternative views of aggregate supply (AS) (AO2, AO4)

While economists agree on the shape of the short run aggregate supply (SRAS) curve, there is disagreement about the long run aggregate supply (LRAS).

Monetarist/new classical view of the long run aggregate supply (LRAS) curve

Monetarists (or new classical economists) differ in their view of the long run aggregate supply (LRAS) curve from that of Keynesian economists. Monetarists believe that the LRAS curve is vertical at the full employment level of output in the long run (see Figure 17.8). Hence, according to the new classical model of LRAS, aggregate supply is independent of the price level in the long run, that is, it is perfectly price inelastic at the full employment level of output.

◆ The **long run aggregate supply** (LRAS) curve, according to the new classical model, is the maximum level of real GDP at the full employment level of national output. It shows that aggregate supply is independent of the price level, that is, LRAS is perfectly price inelastic.

Figure 17.8 shows that the economy cannot produce beyond its productive capacity of Y_f. Monetarists argue that long run equilibrium exists at the potential (full employment) level of real national output, that is, where AD_1 intersects the LRAS curve. Any attempt to increase aggregate demand beyond the capacity of the economy (Y_f), from AD_1 to AD_2, will only be inflationary, with the general price level increasing from PL_1 to PL_2. Therefore, monetarist/new classical economists argue potential output is based on the economy's quantity and quality factors of production, so the price level does not affect the LRAS.

In the long run, the aggregate supply is independent of the price level as this represents the maximum level of output that the economy can produce. Hence, the LRAS curve is vertical and is dependent of non-price variables such as the state of technology or the quantity and quality of factors of production. The potential output of the economy (Y_f) is determined by the quantity and quality of its factors of production, that is, the quantity and quality of natural resources (land), human capital (labour), physical capital (for example, buildings and infrastructure) and entrepreneurship.

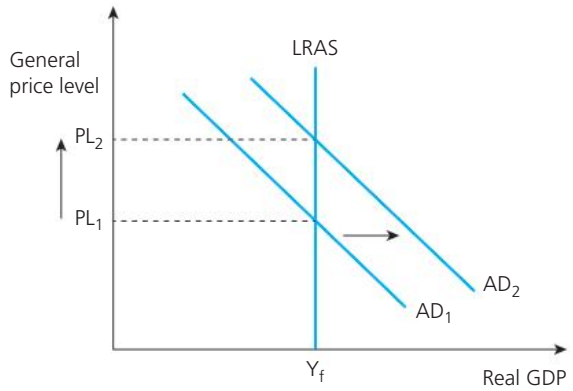


Figure 17.8 The LRAS curve (monetarist/new classical view)

Keynesian view of the AS curve

By contrast, Keynesians believe that the AS curve has three distinct sections (see Figure 17.9), owing to the varying degrees of spare capacity in the economy. They argue that the economy can be below the full employment level (as shown by Y_f in Figure 17.9), even in the long run. For example, following the global financial crisis, Spain suffered a recession from 2008 to 2014.

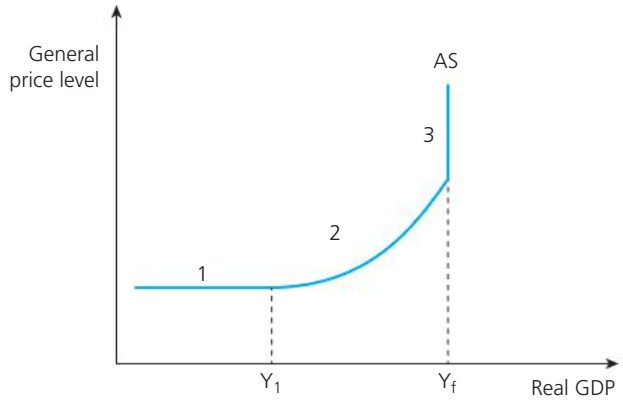


Figure 17.9 The Keynesian AS curve (Keynesian view)

In section 1 of the Keynesian AS curve, aggregate supply is perfectly price elastic (horizontal) as there is plenty of spare capacity (unemployed resources) in the economy. Any increase in aggregate demand has no direct impact on the general price level.

In section 2 of the Keynesian AS curve, aggregate supply is relatively price elastic (upward sloping between Y_1 and Y_f in Figure 17.9) as there is pressure of scarce resources as the economy grows. As with the law of supply (see Chapter 4), an increase in the general price level will incentivize firms to supply more output.

In section 3 of the Keynesian AS curve, aggregate supply is perfectly price inelastic (vertical) as there is no longer any spare capacity in the economy, that is, all factor resources are fully employed. Any increases in aggregate demand beyond the full employment level of output is simply inflationary.

Top tip!

When referring to the Keynesian AS curve, do not use the term 'long run aggregate supply' (LRAS), but **Keynesian aggregate supply**. After all, Keynes famously said that *'in the long run, we are all dead'*.

Inflationary and deflationary/recessionary gaps

An **inflationary gap** (or **positive output gap**) exists if actual national output exceeds the full employment level of output. This means that the economy operates beyond the full employment level of national output due to the excess levels of aggregate demand for the consumption of goods and services. It is called an inflationary gap because the actual GDP exceeds potential GDP so the higher levels of expenditure throughout the economy increases the general price level over time (see Figure 17.10).

◆ An **inflationary gap** exists when an economy's actual real GDP exceeds the potential output at the full employment level of national output.

By contrast, a **deflationary gap** (also known as a **recessionary gap** or **negative output gap**) exists when the real national output equilibrium (shown by P_e and Y_e in Figure 17.11) is below the full employment level of output (Y_f). Hence, actual economic growth is below the average trend rate of growth (see Figure 17.11).

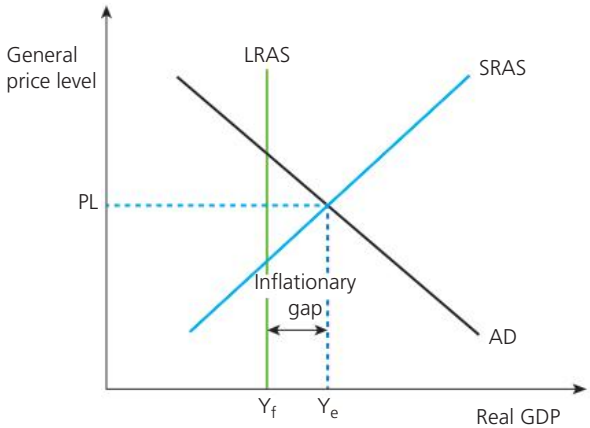


Figure 17.10 The inflationary gap

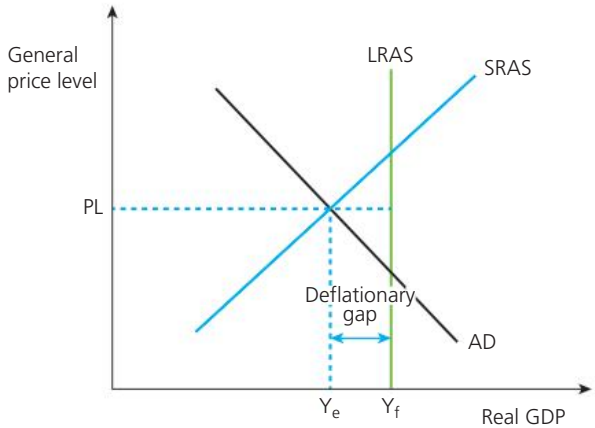


Figure 17.11 The deflationary gap

Deflationary gaps are caused by lower levels of aggregate demand (AD) due to any combination of a fall in:

- consumer spending caused by lower household income or higher interest rates, for example
- investment spending due to lower business confidence or a financial crisis, for example
- net export earnings due to a global recession, for example.

Top tip!

A deflationary gap can also be seen from a PPC diagram (see Chapter 1). Any point within the PPC shows there is a negative output gap as actual output is less than potential output (as depicted on the PPC).

ATL ACTIVITY 17.2

Investigate how a government of a chosen country has responded to a prolonged deflationary gap, such as the global financial crisis (2007–09) or the coronavirus outbreak. Refer to both aggregate demand and aggregate supply, and the WISE ChoICES concepts of intervention and economic well-being in your responses.

Shifts of the AS curve over the long run (AO2, AO4)

◆ A **deflationary gap** (also called a **recessionary gap** or **negative output gap**) occurs when macroeconomic equilibrium is below the full employment equilibrium, that is, AD is insufficient to create full employment.

Irrespective of which AS curve is used, economists agree that aggregate supply increases over the long run (using the monetarist/new classical LRAS) or over the long-term (using the Keynesian AS model). The LRAS or Keynesian AS curves can shift outwards to the right under any combination of the following interrelated factors:

1. **Changes in the economy's quantity and/or quality of factors of production** – The more resources and the better these resources that a country has, the greater its aggregate supply will tend to be. The quantity of factors of production can be increased by discovering new supplies (such as oil or other natural resources), the reclamation of land, or by a larger and more skilled workforce (perhaps due to migration of workers). The quality of labour resources can be improved by better education, training and work practices.
2. **Improvements in technology** – Higher productivity of factor resources brought about by improvements in technology (technological progress and improved production techniques) and innovations can increase the LRAS. Innovation of new products and services will help to increase national output in the long-term. For example, the digital world continues to revolutionize the corporate world, creating economic opportunities across almost all industries, helping to increase the productive capacity of the economy. The Case Study below shows some examples of how digital disruptions have created economic opportunities in various industries.

CASE STUDY

Examples of digital disruptions of major industries

- Alibaba, the world's most valuable retailer, does not have any inventory.
- Airbnb, the world's largest accommodation provider, does not own any real estate.
- Uber Eats, Grubhub, Just Eat and Deliveroo are some of the world's largest food ordering and delivery companies but do not own any restaurants.
- Netflix, the world's largest movie house, does not own any cinemas.
- Skype, Viber and WhatsApp (owned by Facebook) are some of the world's largest telecommunications companies but do not rely on having any physical infrastructure.
- Uber, the world's largest taxi company, does not own any taxis.

3 **Increases in efficiency** – Efficiency refers to making the best possible use of an economy's scarce resources to produce optimal combinations of goods and services for society, determined by the input–output ratio, that is, the input of scarce factor resources to generate national output of the economy. It requires all factor resources (such as land, labour and capital) to be put to their best use and that there is no wastage in the production process. Hence, the most output is generated from the available resources of the economy. Investments in technology and research and development (R&D) can lead to efficiency gains, including improved factor mobility and labour productivity. Note that in microeconomics, the concept of efficiency was explained by using the PPC model (see Chapter 1), whereas in macroeconomics it can also be represented by an outwards shift of the AS curve.

4 **Changes in institutions** – LRAS is also affected by the institutional structure of the economy, such as its legal and financial systems. For example, many governments intervened after the 2008 global financial crisis to rescue commercial banks and again following the global outbreak of the novel coronavirus (Covid-19 in 2019–2020). Governments can further boost LRAS by creating economic incentives and promote greater entrepreneurial risk-taking. Such institutional factors include improved education and training of the workforce, and better healthcare provision, which help to increase the economy's potential output in the long-term. Better infrastructure such as improved road, rail and air networks also enhances the economy's productive capacity, thus increasing the AS in the long-term. Finally, promoting and facilitating freer and fairer international trade deals (see Chapter 28) can also increase the economy's AS in the long-term.

Keynesians and monetarists differ in their view about the shape of the LRAS curve. Figure 17.13 shows a shift of the LRAS curve from the monetarist/new classical perspective, while Figure 17.12 shows the Keynesian view.

A favourable change in the productive capacity of an economy (such as major technological advances or the discovery of important resources such as oil) can shift the AS curve outwards to the right (from Y_{f1} to Y_{f2} in Figure 17.12), due to the economy's increased potential output.

In Figure 17.13, the rightwards shift of the LRAS curve from $LRAS_1$ to $LRAS_2$ results in an increase in the productive capacity of the economy from Y_{f1} to Y_{f2}. This is caused by an increase in the quantity and/or quality of the economy's factors of production, perhaps due to technological advances such as 5G wireless broadband technologies.

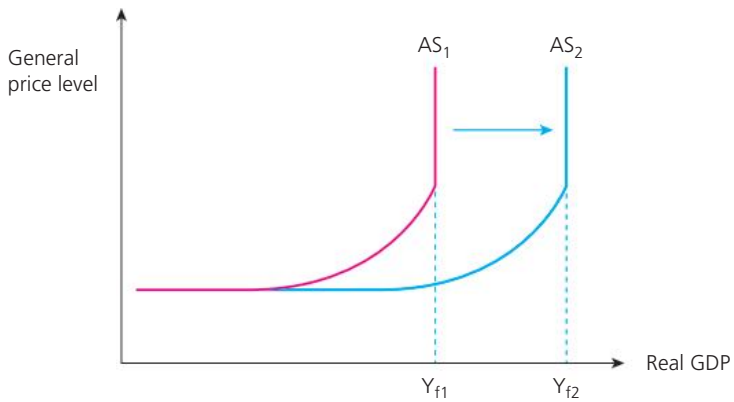


Figure 17.12 Shifts in the Keynesian AS curve

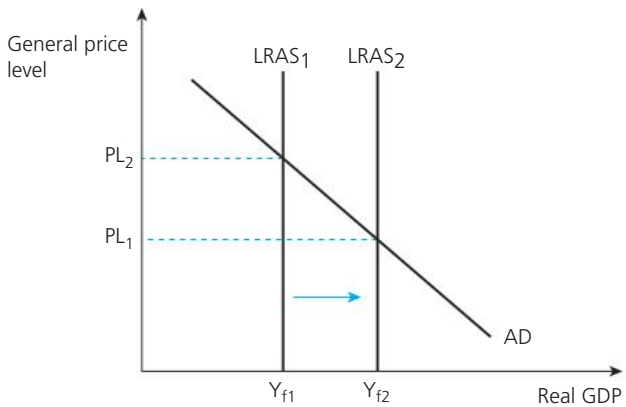


Figure 17.13 A shift in the LRAS curve (monetarist/new classical model)

Top tip!

This is a very important topic in macroeconomics as AD–AS analysis is used to explain the operations of the whole economy. Therefore, make sure you practise drawing AD–AS diagrams and are able to distinguish between the new-classical LRAS and Keynesian AS curves.

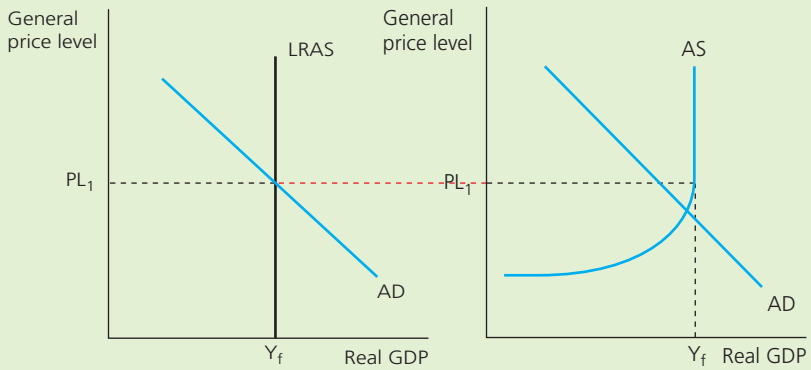


Figure 17.14 a) Classical view of LRAS; b) Keynesian view of AS

◆ **Macroeconomic equilibrium** exists when aggregate demand is equal to aggregate supply in the economy, that is, AD = AS.

Macroeconomic equilibrium (AO2, AO4)

Short run macroeconomic equilibrium

Short run **macroeconomic equilibrium** occurs when short run aggregate supply is equal to aggregate demand, that is, SRAS = AD measured, thus determining the general price level and real GDP level. Therefore, any change in AD and/or SRAS will change the macroeconomic equilibrium position.

For example, an increase in aggregate demand can be caused by an increase in net exports or lower income tax, ceteris paribus. In Figure 17.15, short run equilibrium is initially at point A, where AD_1 = SRAS. The higher aggregate demand increases both real GDP and the general price level (from Y_1 to Y_2 and PL_1 to PL_2 respectively), due to the rightwards shift from AD_1 to AD_2. This results in a new macroeconomic equilibrium at point B, where AD_2 = SRAS. The opposite is true if aggregate demand falls, perhaps due to a recession.

In Figure 17.16, an increase in factor costs of production and/or indirect taxes causes the SRAS to fall, shown by the leftwards shift of the SRAS curve from $SRAS_1$ to $SRAS_2$. This reduces real GDP from Y_1 to Y_2, creating a negative output gap, and increases the general price level from PL_1 to PL_2.

By contrast, an outwards shift of the SRAS curve will tend to increase real national output and reduce the general price level. This might be caused by factors such as reduced interest rates or improved training opportunities for employees, for example.

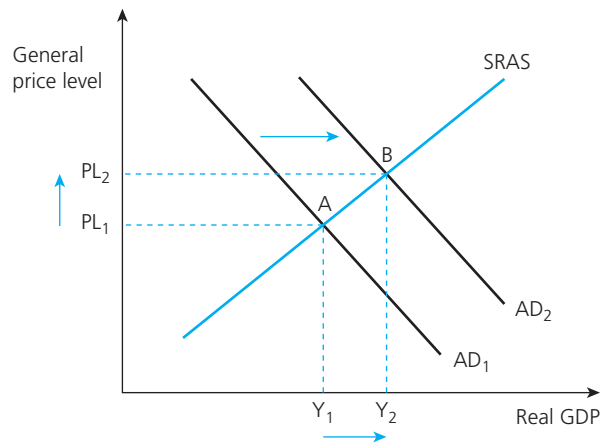


Figure 17.15 Short run equilibrium (a shift in the AD curve)

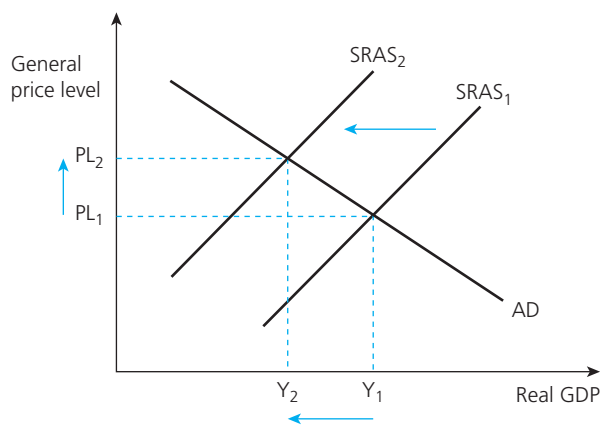


Figure 17.16 Short run equilibrium (a shift in the SRAS curve)

Top tip!

When analysing changes in aggregate demand and aggregate supply, consider the concept of price elasticity of demand (PED) and price elasticity of supply (PES). For example, an increase in AD will have a far greater impact on the general price level if AS is price inelastic.

PAPER 3 EXAM PRACTICE QUESTION 17.4 (HL ONLY)

Refer to the information in the table below and answer all questions that follow. All figures except for the price indices are expressed in billions of dollars ($bn).

Price index	Consumption	Investment	Government spending	Exports	Imports	Aggregate supply
128.0	20	4	5	6	10	53
124.0	22	6	7	7	9	47
120.0	24	8	9	8	8	41
116.0	26	10	11	9	7	35

a Calculate the value of net exports at a price index of 128. *[1 mark]*

b Calculate the equilibrium level of real national output in the economy. *[2 marks]*

c Calculate the price level (price index) when the value of aggregate demand is $49 billion. *[2 marks]*

d State the value of investment expenditure if the level of aggregate demand is $33 billion. *[1 mark]*

Equilibrium in the monetarist/new classical model

In the monetarist/new classical model of macroeconomic equilibrium, long run equilibrium occurs at the full employment level of output (potential output). Automatic adjustments through the forces of AD and AS restore the economy to full employment equilibrium in the long run. They assume wages and prices are flexible enough to maintain full employment in the long run.

◆ The **full employment level of output** exists at the point where unemployment is at its natural rate, so that everyone who is willing and able to work is in employment.

Monetarists/new classical economists assume the economy will reach full employment in the long run. The new classical model of AD–AS assumes that the LRAS curve is vertical at the full employment level of output (as shown by Y_f in Figure 17.17). **Full employment** occurs when those who are willing and able to work have or can find a job, so the economy is operating at full capacity, that is, it is not possible to increase real national output as all resources are fully utilized. Any further attempts to increase aggregate demand beyond the full employment level of national output, for example, from AD_1 to AD_2, will simply be inflationary, causing the general price level to increase from PL_1 to PL_2.

Note that full employment does not mean that there is zero unemployment in the economy as there will always be some unemployment that naturally exists because:

- a certain number of people are in-between jobs (frictional unemployment)
- redundancies are caused by cyclical factors in the year (seasonal unemployment), and
- there is skills mismatch in certain industries (structural unemployment).

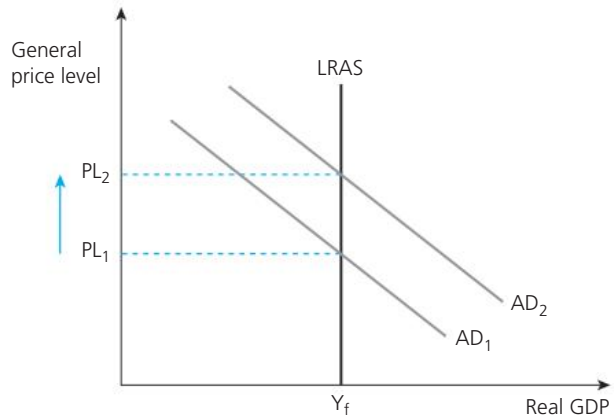


Figure 17.17 Long run equilibrium (Monetarist/new classical model)

Collectively, these three types of unemployment (see Chapter 19) are referred to as the **natural rate of unemployment**. Hence, unemployment at the full employment equilibrium level of national output is equal to the natural rate of unemployment. It represents the lowest level of unemployment that is compatible with price stability.

◆ The **natural rate of unemployment** (NRU) is the level of unemployment at the full employment equilibrium level of national output, comprising frictional, seasonal and structural unemployment.

Top tip!

'Full employment means that everyone who is able and willing to work has a job, so there is no unemployment in the economy.'

Despite the term, it is incorrect to assume that full employment occurs when there is zero unemployment in the economy. Instead, full employment occurs at the natural rate of unemployment. There will always be some degree of natural unemployment owing to frictional, seasonal and structural unemployment.

The monetarist model of AD–AS assumes automatic adjustment of the economy to full employment equilibrium. In the new classical model, any short-term fluctuations in national output will be only temporary as market forces will restore equilibrium to the full employment level of output in the long run. For example, in the case of an inflationary gap (see Figure 17.18), macroeconomic equilibrium is restored via the market mechanism in the following way:

- Long run equilibrium is at the full employment level of national output (Y_f) with the general price level at PL_1.
- An increase in aggregate demand from AD_1 to AD_2 increases the general price level from PL_1 to PL_2, causing AS to expand along the $SRAS_1$ curve, thereby creating an inflationary gap.
- This temporarily increases national output beyond its capacity, so raises production costs, such as higher wage costs.
- Hence, aggregate supply shifts from $SRAS_1$ to $SRAS_2$, causing the general price level in the economy to rise from PL_2 to PL_3 and the economy operating back at Y_f.
- Hence, macroeconomic equilibrium is restored at the full employment level of national output (Y_f).

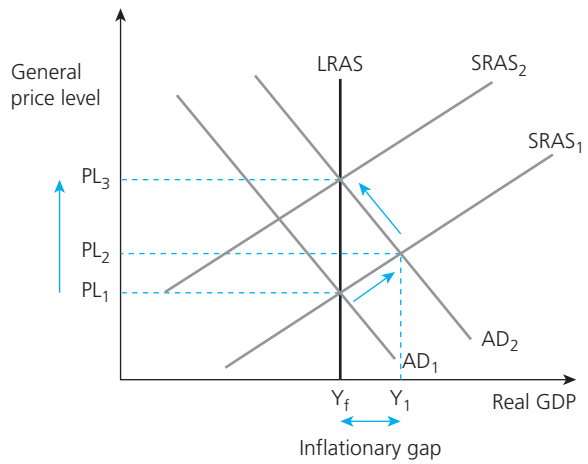


Figure 17.18 Restoring an inflationary gap (monetarist model)

Alternatively, in the case of a deflationary gap caused by a temporary fall in aggregate demand (see Figure 17.19), macroeconomic equilibrium is also restored by the market mechanism via the following process:

- Long run equilibrium is at Y_f, with the general price level initially at PL_1.
- A fall in aggregate demand, due to an economic downturn, causes a shift in aggregate demand from AD_1 to AD_2, thus reducing the general price level from PL_1 to PL_2 and resulting in a fall in real GDP from Y_f to Y_1, creating a temporary deflationary gap.

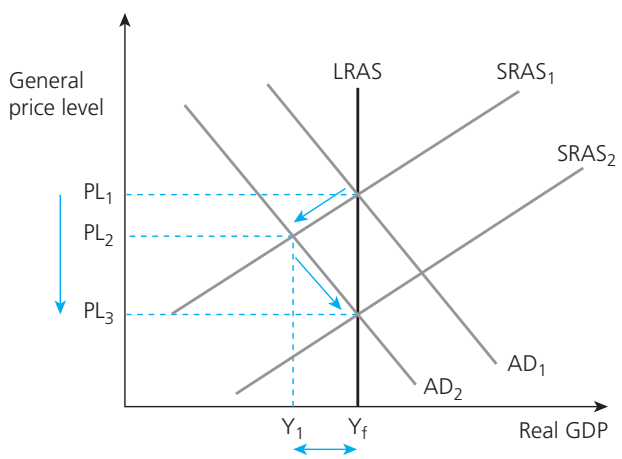


Figure 17.19 Restoring a deflationary gap (monetarist model)

- The lower price level reduces the production costs of firms in the economy, thereby helping to increase short run aggregate supply from $SRAS_1$ to $SRAS_2$, with the economy operating back at Y_f.
- Hence, macroeconomic equilibrium is restored at the full employment level of national output (Y_f), with the general price level falling from PL_2 to PL_3.

Hence, new classical economists believe that demand-side policies (see Chapters 24 and 25) are ineffective for achieving economic growth in the long run. Instead, they prefer to use supply-side policies (see Chapter 26) to shift the long run aggregate supply (LRAS) outwards, thereby achieving the macroeconomic objectives of economic growth, lower unemployment and a fall in the general price level in the long run.

Equilibrium in the Keynesian model

Keynesian economists do not believe that the economy can always maintain equilibrium at the full employment level of output but that there can be a persistence of recessionary gaps, that is, the macroeconomic equilibrium level of output might not equal the full employment level of output in the long-term. For example, high unemployment in the economy can create a negative multiplier effect, which reduces real GDP even further. This means the initial fall in household income would cause an overall decrease in consumption in the economy that is larger than the initial decrease in income (see Chapter 25). Hence, Keynes argued that government intervention is needed to prevent a prolonged negative output gap. Hence, the economy can be in macroeconomic equilibrium at any of the three sections of the Keynesian AS curve in both the short run and long run.

In Figure 17.20, there is spare capacity (unused resources) along the price elastic section of the Keynesian AS curve. Hence, an increase in aggregate demand from AD_1 to AD_2 increases real national output from Y_1 to Y_2 without putting any pressure on the general price level which remains at PL_e.

In Figure 17.21, the combination of increasing demand for scarce resources and labour shortages causes the general price level to rise from PL_1 to PL_2 as national output increases from Y_1 to Y_2. Note that the proportionate increase in real GDP is greater than the rise in the general price level. Contrary to the new classical model, the Keynesian model shows that increases in AD need not necessarily cause inflation in the economy, unless it is operating at or near its full employment level of output.

In Figure 17.22, at the full employment level of national output (Y_f), firms compete for highly limited resources as the economy is at full capacity. This forces up the general price level from PL_1 to PL_2, even though the economy cannot produce beyond its productive capacity.

Unlike monetarists, Keynesians argue that labour markets do not necessarily clear in order to restore macroeconomic equilibrium. Keynes argue that wages are 'sticky downwards' (that is, labour markets are inflexible) because:

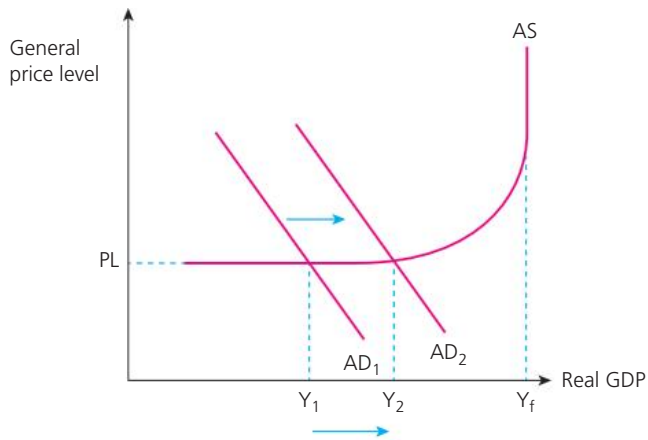


Figure 17.20 Equilibrium in the Keynesian model (phase 1)

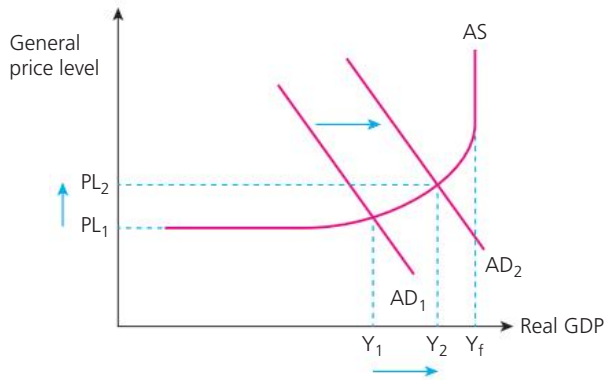


Figure 17.21 Equilibrium in the Keynesian model (phase 2)

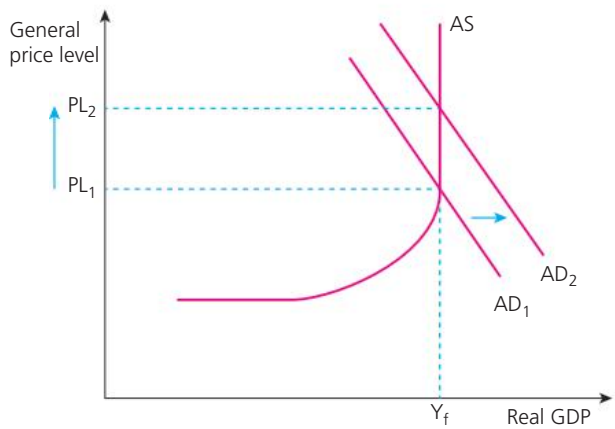


Figure 17.22 Equilibrium in the Keynesian model (phase 3)

- workers get used to a certain wage rate or income so are inflexible in accepting nominal pay cuts, especially when backed up by their trade union and industrial action to upload employee rights
- existing employment contracts can prevent wages and salaries from falling below the agreed (contracted) level
- some firms may prefer to cut employment rather than wages because pay cuts can reduce worker morale and labour productivity
- it is not legally possible to cut wages below the national minimum wage, even during a major economic downturn.

So, even if there is spare capacity in the economy up to Y_2 in Figure 17.22, due to an economic recession and high unemployment, wages can still remain sticky downwards thereby preventing the economy from restoring macroeconomic equilibrium.

Keynesians believe that the macroeconomy is not always able to self-correct, especially from deep recessions. Without government intervention, the economy can remain stuck in a deflationary gap because of mass unemployment in the economy. The decline in aggregate demand is often due to interconnected reasons beyond the country of households and firms in the private sector. For example, the global recession caused by the coronavirus outbreak led to a decline in:

- consumption expenditure (C) due to the decline in consumer confidence and falling wages
- export earnings (X) due to a major global recession and decline in world trade
- investment (I) due to the collapse in the banking sector and stock markets across the world.

The global pandemic caused economic growth in almost every country to be well below the average trend rate of growth, due to AD falling significantly below the productive capacity of the economy.

Top tip!

Make sure you can explain the difference between the equilibrium level of national income (Y_e) and the full employment level of national income (Y_f) and why this will result in either an inflationary gap ($Y_e > Y_f$) or a deflationary gap ($Y_e < Y_f$).

Figure 17.23 Governments around the world intervened to rescue businesses during the coronavirus outbreak

PAPER 2 EXAM PRACTICE QUESTION 17.5

Using an appropriate diagram, explain how a recession impacts the short run equilibrium position of an economy. *[4 marks]*

PAPER 2 EXAM PRACTICE QUESTION 17.6

Using an aggregate demand and aggregate supply diagram, explain what is likely to happen following a temporary reduction in both income tax and corporation tax. *[4 marks]*

Assumptions and implications of the monetarist/ new classical and Keynesian models (AO3)

The key assumption in the Keynesian model is wage inflexibility, which can cause persistent recessionary gaps so prevents the economy from reaching macroeconomic equilibrium. By contrast, the new classical model assumes this to be natural unemployment if it persists in the long run.

Monetarists (new classical economists) assume that long run equilibrium occurs at the potential or full employment level of real national output. In Figure 17.17 on page 257, this is shown by the intersection of AD_1 and the LRAS curve, where Y_f represents the full employment level of national output. Any fall in real national output below the macroeconomic equilibrium is assumed to be only temporary as the economy would return to equilibrium on its own accord. In particular, new classical economists assume that if the economy is below full employment, wages should fall and labour flexibility will clear the economy of any unemployment.

Therefore, monetarists argue that demand-side policies are ineffective in the long run and there is no need for government intervention. The implication is they prefer the use of supply-side policies (see Chapter 26) to shift the LRAS outwards to achieve economic growth in the long run. These policies include improved productivity of factors of production, technological progress and incentives to stimulate investments, research and development, and innovations.

In reality, it is rather difficult to know precisely the full employment level of output. It can be impractical for every firm in the economy to operate at 100 per cent capacity and it is not practical to clearly determine the total number of people in natural unemployment.

Keynesian economists assume that wages are 'sticky' downwards, that is, workers are not so willing to accept a pay cut even in a recession, so the market forces of AD and AS will not necessarily restore macroeconomic equilibrium at the full employment level of output. Hence, the implication of this is that there is a need for government intervention to restore long run macroeconomic equilibrium through the use of loose monetary policy and/or expansionary fiscal policy (see Chapters 24 and 25 respectively).

Keynes used the phrase 'sticky downwards' to describe a situation in which a variable (such as the wage rate) is resistant to change, thereby reducing the effectiveness of the market forces of demand and supply in the labour market. Therefore, wages being sticky downwards means that wages can move up with ease (who wouldn't accept a pay rise?!) while labour unions

Theory of Knowledge

Discuss how language affects perceptions about economic events or situations.

are assumed to resist any attempt to have the wages of their members being reduced, even if this is in the best interest of the workforce in the long run. The result is inflationary without any corresponding increase in output, so Keynesians assume this can leave the economy in a prolonged recessionary gap.

Theory of Knowledge

To what extent is it rational for workers to accept a pay cut?

Figure 17.24 Under what circumstances, if any, should people willingly accept a pay cut?

Therefore, this implies the need for government intervention if persistent recessionary gaps exist. This means there is a justification for government expenditure in the economy in order to close a prolonged deflationary gap (as the market economy will not close this gap). The government can choose to use expansionary fiscal and/or monetary policies to boost aggregate demand in order to match long run aggregate supply, raising the real national output to Y_f, thereby closing the negative output gap (see Figure 17.11 on page 253). Expansionary fiscal policy is the use of tax cuts and/or increased government spending to stimulate more economic activity (see Chapter 25). Expansionary monetary policy is the use of lower interest rates to encourage consumption and investments in the economy (see Chapter 24). This analysis provides justification for government intervention – increased government spending financed by borrowing to stimulate confidence levels and national expenditure.

Top tip!

If you are an HL student you need to make sure you review the contents of this topic carefully as it has major implications on policymaking (see Chapters 24–26). Hence, understanding the contrasting perspectives of macroeconomic equilibrium in the monetarist/new classical model and the Keynesian model will be extremely useful as a framework for answering the policy paper question in Paper 3.

WISE ChoICES

An economy can be productively efficient but allocatively inefficient at the same time, because the latter is concerned with the optimal distribution of scarce resources. For example, an economy might devote a significant amount of scarce resources to provide national defence in a productively efficient way, instead of allocating the funds for private education and healthcare (resulting in a socially undesirable allocation of resources). What other examples can you think of in economics where **efficiency** gains conflict with **scarcity** of resources in the economy?

Chapter summary

- As aggregate demand has an inverse relationship with the general (average) price level, the AD curve is downward sloping.
- Factors that cause a change in any component of aggregate demand (AD = C + I + G + X – M) will cause a shift in the AD curve, ceteris paribus.
- A favourable change in non-price factors that affect aggregate supply will shift the SRAS curve; for example, lower production costs or a cut in indirect taxes.
- Macroeconomic equilibrium in the economy exists when aggregate demand is equal to aggregate supply, that is, AD = AS.
- Keynesians and monetarists/new classical economists differ in their view about the shape of the LRAS curve.
- The monetarist/new classical view is that LRAS is vertical at the full employment level of national output because the economy is operating at its full capacity. Hence, it is not possible to increase production beyond the economy's potential output in response to growing aggregate demand in the economy.
- Monetarist/new classical economists argue potential output is based on the economy's quantity and quality factors of production, so the price level does not affect the LRAS.
- Changes in the quantity and/or quality of factors of production result in permanent shifts to the LRAS curve; for example, technological progress and innovations.
- In the monetarist/new classical model of macroeconomic equilibrium, long run equilibrium occurs at the full employment level of output (potential output). Automatic adjustments through the forces of AD and AS restore the economy to full employment equilibrium in the long run.
- Unemployment at full employment equilibrium is equal to the natural rate of unemployment.
- There will always be some degree of natural unemployment in the economy due to frictional unemployment (those between jobs) and structural unemployment (caused by skills mismatch). So, full employment occurs at the natural rate of unemployment.
- Monetarists (or new classical economists) suggest that any short run equilibrium (below the full employment level of real national output) is only temporary. This is because market forces are assumed to restore macroeconomic equilibrium in the long run, at the full employment level of national output.
- The key assumption in the Keynesian model is wage inflexibility, which can cause persistent recessionary gaps so prevents the economy from reaching macroeconomic equilibrium. By contrast, the new classical model assumes this to be natural unemployment if it persists in the long run.

18 Macroeconomic objectives – economic growth

Real-world issue

Why does economic activity vary over time and why does this matter?

Conceptual understandings (WISE ChoICES)

- **Change** in the conditions of the demand and supply sides of the economy cause economic activity to vary over time.
- Fluctuations in economic activity impact the **economic well-being** of individuals and societies.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- economic growth, with reference to short-term growth, actual growth in the PPC model, and the role of AD in the AD/AS model (AO2)
- economic growth, with reference to long-term growth, shifts of the PPC (growth in production possibilities), and the role of LRAS in the AD/AS model (AO2)
- the measurement of economic growth (AO2)
- a PPC model diagram showing actual growth and growth in production possibilities (AO4)
- diagrams that show a shift in AD, leading to an increase in real output (AO4)
- diagrams that show a shift in LRAS, leading to an increase in full employment output (AO4)
- the consequences of economic growth, with reference to the impact on living standards, the environment and income distribution (AO3).

Economic growth (AO2, AO4)

Economic growth refers to an increase in a country's real gross domestic product (GDP) over time. GDP measures the monetary value of goods and services produced within a country for a given period of time, usually one year. Real GDP is the value of an economy's national output that has been adjusted for inflation in order to reflect the true value of goods and services produced in a given year.

◆ **Economic growth** refers to a sustained increase in a country's real GDP over time. This is usually expressed as the annual percentage change in real national output.

The most common interpretation of economic growth is an increase in real GDP for two consecutive quarters. An alternative interpretation is a sustained increase in the potential productive capacity of the economy. Economic growth is usually expressed as the country's annual percentage change in the value of the country's real national output.

By contrast, **negative economic growth** occurs when the level of economic activity declines, that is, there is a fall in real GDP. This is associated with an economic downturn in the economy, such as a recession or slump in the business cycle (see Chapter 16).

◆ **Negative economic growth** occurs when the level of economic activity declines, that means, there is a fall in real GDP. This is associated with a recession in the business cycle.

In general terms, economic growth occurs when there is an increase in the quantity and/or quality of the economy's factors of production (see Chapter 1), such as an increase in labour productivity or improvements in the state of technology. Economists believe that sustained economic growth

is an important macroeconomic objective because it is the most practical way for the majority of people in a country to have a higher standard of living.

Different countries have different rates of economic growth for various reasons. For example:

- **Factor endowments** refers to the quantity and quality of a country's factors of production. The USA and Saudi Arabia are well endowed in the supply of oil, France has plenty of arable land for its agricultural output, and Australia has many natural resources for production such as coal, gold and iron ore. The more factor resources that a country has, the more likely it is to achieve economic growth.
- **The size and skills of the labour force** impacts on the country's economic growth. For example, the large labour force in China and India and Germany's highly skilled workforce have contributed to the economic growth of these countries.
- **Investment expenditure** in capital and human resources is vital for long-term competitiveness and economic growth as it boosts the country's productive capacity. Countries that invest a greater proportion of their GDP in capital investment and encourage foreign direct investment (FDI) can stimulate sustainable economic growth and development.
- **Discovery of raw materials** such as the discovery of oil, or any other tradable commodities, will increase the country's productive capacity. This can help to shift the economy's production possibility curve (PPC) outwards.
- **Labour productivity** refers to the output produced in a given time period. It is determined by several interrelated factors such as the qualifications, experience, training, skills and motivation of the labour force. Higher labour productivity tends to lead to greater economic growth. Countries that spend a greater proportion of their GDP on investments in education and training are more likely to achieve sustainable economic growth.
- **The mobility of labour** refers to the extent to which workers are willing and able to change jobs (*occupational mobility*) and move to different locations for employment (*geographical mobility*). The more occupationally and geographically mobile workers are, the greater economic growth tends to be.

◆ The **labour force** refers to people of working age who are in employment (including the self-employed) plus those who are seeking work, that is, those who are available for work at current real wage rates.

■ **Figure 18.1** Vineyard in Burgundy, France, where the land and climate are perfect for wine production

Short-term growth (AO2, AO4)

Recall from Chapter 1 (page 19) that the production possibility curve (PPC) shows the maximum possible combinations of output for two different products that the economy can achieve when all resources are employed efficiently.

Actual output refers to the current, rather than potential, level of real GDP in the economy. It is represented by any point in a PPC diagram. **Actual growth** occurs in the short-term when an economy operates below its full employment level of national income but moves towards its potential level of GDP (as illustrated by the PPC) by using more of its resources and/or more efficiently. It is measured by the annual percentage change in a country's real national output (GDP). In Figure 18.2, this is shown by the economy moving from point A to point B, as a result of using previously unemployed resources. Hence, the country is able to produce more goods *and* services.

◆ **Actual output** refers to the current, rather than potential, level of real GDP in the economy.

◆ **Actual growth** refers to short-term growth which occurs when an economy moves towards its potential level of GDP (as illustrated by the PPC) by using more of its resources in more efficient ways.

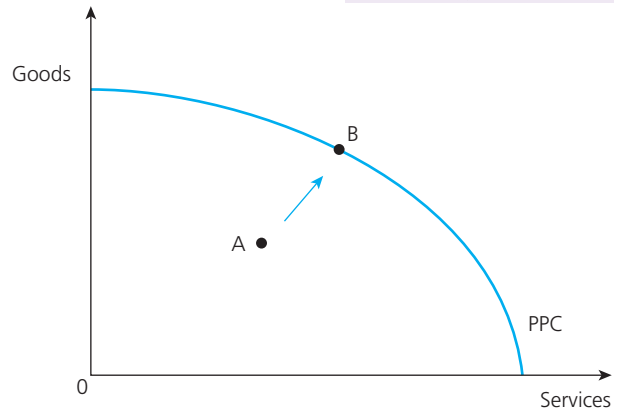


Figure 18.2 Actual growth in the PPC model

Alternatively, actual economic growth can be illustrated using an AD–AS diagram (see Figure 18.3). Real gross domestic product (real GDP), shown along the *x*-axis, is a monetary measure of the value of all output of goods and services in the economy, per time period, that is adjusted for inflation. It is commonly used to measure the real value of aggregate demand in the economy (see Chapter 17). The higher the level of aggregate demand in the economy, the larger the country's GDP tends to be. Recall that AD = C + I + G + (X – M), where:

- C = Consumption expenditure of individuals and households
- I = Investment expenditure of firms
- G = Government expenditure
- X = Export earnings of the economy
- M = Import expenditure of the economy.

In Figure 18.3, short run equilibrium is initially at point A, where AD_1 = SRAS (the short run aggregate supply), and real GDP at Y_1. A rightwards shift of aggregate demand from AD_1 to AD_2 results in real GDP rising from Y_1 to Y_2, although this also causes an increase in the general price level from PL_1 to PL_2. The new equilibrium is therefore at point B, where AD_2 = SRAS.

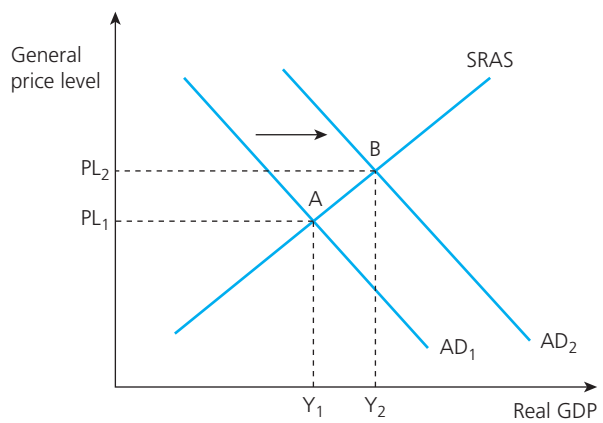


Figure 18.3 Actual growth in the short run AD–AS model

Long-term growth (AO2, AO4)

Potential output is associated with an increase in quantity and/or quality of factors of production in the long run. Potential growth also helps to achieve two macroeconomic objectives simultaneously: employment and economic growth. Using the PPC model, potential economic growth as a long run concept is demonstrated by an outwards shift of the economy's PPC from PPC_1 to PPC_2 or from point B to point C in Figure 18.4 due to an increase in the productive capacity of the economy. This can be caused by the discovery of new resources and/or improvements in the quality of factors of production.

◆ **Potential output** refers to the possible level of real GDP of an economy, as shown on its production possibility curve, if all resources are used efficiently. It is a long run phenomenon represented by an outwards shift of the PPC.

Alternatively, long-term growth can be illustrated by an AD–AS diagram. Potential growth is shown by an outwards shift of the long run aggregate supply (LRAS curve), as in Figure 18.4. Potential economic growth can be shown by a rightwards shift of the

LRAS curve, as shown in Figure 18.5. Sustained economic growth is achieved by an increase in the potential output of the economy, illustrated by the outwards shift of the LRAS curve from $LRAS_1$ to $LRAS_2$. This is caused by an increase in the quantity and/or quality of resources, thereby shifting the long run aggregate supply curve (LRAS) to the right. For example, an increase in the labour supply or improvements in the state of technology will shift the LRAS curve outwards, as shown in Figure 18.5.

Figure 18.5 shows that long-term economic growth occurs when there is an increase in the potential output of the economy, caused by the outwards shift of the long run aggregate supply (LRAS) curve from $LRAS_1$ to $LRAS_2$. This may be desirable as a macroeconomic objective as the real GDP increases from Y_{f1} (the previous level of real GDP at the full employment level of national output) to Y_{f2}. In addition, the general price level falls from PL_1 to PL_2, which helps to reduce the cost of living.

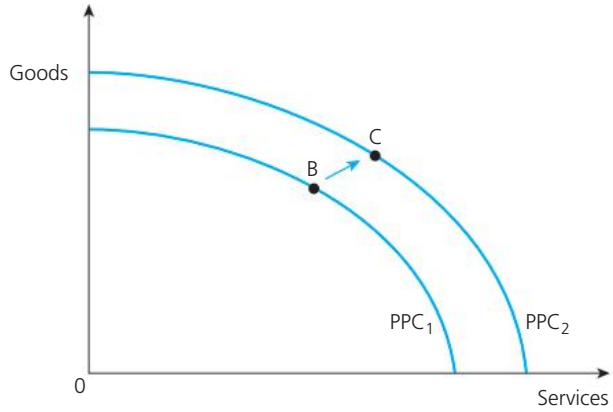


Figure 18.4 Potential growth in the PPC model

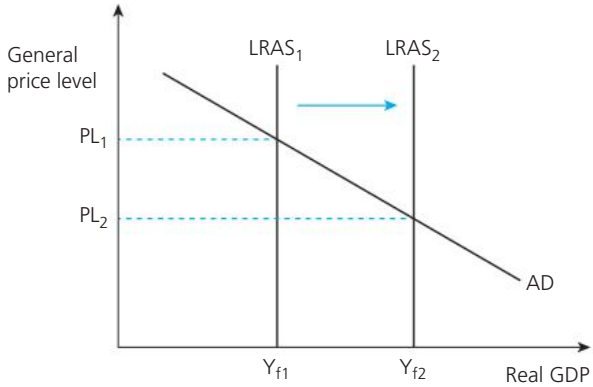


Figure 18.5 Potential growth in the AD–AS model

Top tip!

There are numerous ways to define the term economic growth, such as defining growth as a rise in real GDP over at least two consecutive quarters or by using the PPC model to define growth as a long run concept of a sustained rise in the economy's potential output. Therefore, be careful how you define this term, and substantiate your answer.

Top tip!

Although economic growth is generally regarded as the main single indicator of economic well-being or the general standard of living in a country, there are other measures that are used, such as the Human Development Index (HDI). This composite index (see Chapter 34) measures three dimensions of living standards: healthcare (life expectancy), education (years of schooling) and income levels (real GDP per capita).

◆ **Nominal economic growth** (or **nominal GDP**) is the annual rate of change in the monetary value of GDP. Changes in nominal GDP figures are expressed at **current prices**.

Measurement of economic growth (AO2)

Economic growth is measured by calculating changes in real gross domestic product (GDP) over a specified period of time. Economic growth can be measured in nominal and real terms. **Nominal economic growth** is the annual rate of change in the monetary value of GDP. This is expressed in terms of *current prices*, that is, the money value of GDP at the current point in time. **Real economic growth** adjusts nominal economic growth to take account of inflation (that is, changes in the general price level as measured by changes in consumer prices). It reflects the true value of goods and services produced in a given year because inflation artificially raises the value of a country's output.

◆ **Real economic growth** (or **real GDP**) is the annual rate of change in the value of nominal GDP after it has accounted for changes in the price level, that is, inflation. Changes in real GDP figures are expressed at **constant prices**.

Changes in real GDP figures are expressed at *constant prices*, that is, removing the impact of price increases to allow a continual or constant comparison of real changes to the value of GDP. For example, if nominal GDP rises by 5 per cent but part of this was due to the 3 per cent increase in average prices in the economy, then real GDP has increase by only approximately 2 per cent. So, if the percentage change in nominal GDP rises by a larger amount than the rate of inflation, real GDP will increase (that is, there is real economic growth). By contrast, if the rate of change in nominal GDP is less than for average price levels, this would be a sign of an economy falling into a recession.

A GDP deflator (see page 231) is needed to measure the economy's real GDP. It is an index number used to convert nominal GDP to real GDP by eliminating the impact of inflation on the calculation of GDP. The figure for the GDP deflator is a broad measure of inflation or changes in consumer prices. Real GDP is calculated by dividing nominal GDP by the GDP deflator and then multiplying the result by 100. As an example, consider the data in Table 18.1:

Table 18.1

Year	Nominal GDP ($bn)	GDP deflator
2019	110.0	100.0
2020	115.5	103.5
2021	121.0	105.0

To calculate the real GDP, it is necessary to deflate the nominal GDP by using the GDP deflator (to eliminate the effects of inflation). In Table 18.2 the base year is 2019 as the GDP deflator is 100:

Table 18.2

Year	Nominal GDP ($bn)	GDP deflator	Real GDP ($bn) (to 2 dp)
2019	110.0	100.0	(110.0 ÷ 100.0) × 100 = 110.00
2020	115.5	103.5	(115.5 ÷ 103.5) × 100 = 111.59
2021	121.0	105.0	(121.0 ÷ 105.0) × 100 = 115.24

In order to calculate the economic growth rate, it is necessary to calculate the percentage change in the year-on-year real GDP figures (Table 18.3):

Table 18.3

Year	Nominal GDP ($bn)	GDP deflator	Real GDP ($bn)	Growth rate (%)
2019	110.0	100.0	110.00	–
2020	115.5	103.5	111.59	(111.59 – 110.0) ÷ 110.0 = 1.45
2021	121.0	105.0	115.24	(115.24 – 111.59) ÷ 111.59 = 3.27

WORKED EXAMPLE

Year	Nominal GDP ($bn)	GDP deflator	Real GDP ($bn)	Nominal growth rate (%)	Real growth rate (%)
2018	120.0	100.0	120.00	–	–
2019	126.5	102.8	123.05	5.41%	2.54
2020	136.2	106.4	128.00	7.67%	4.02

- With 2018 as the base year, the nominal GDP is the same as the real GDP for the year, that is, $120bn.
- With inflation running at 2.8% in 2019, the nominal value of GDP includes higher prices due to inflation. A GDP deflator of 102.8 means that real GDP in 2019 is actually $126.5bn ÷ 102.8 = $123.05bn.
- The real growth rate between 2018 and 2019 is therefore (123.05 – 120.0) ÷ 120 = 2.54%. Note that growth in nominal GDP is higher: (126.5 – 120) ÷ 120 = 5.41%.
- Similarly, in 2020 the nominal GDP is deflated to derive at: ($136.2bn ÷ 106.4) = $128.0bn.
- The growth rate between 2019 and 2020 is (128.0 – 123.05) ÷ 123.05 = 4.02%. Again, note that growth in nominal GDP is higher: (136.2 – 126.5) ÷ 126.5 = 7.67%.

It can also be useful to measure real GDP in terms of per capita (per person). For example, countries such as China, India, Indonesia, Brazil and USA have a large nominal GDP partly due to their large population size. This might suggest that all countries with large populations have a high GDP and therefore high standards of living. However, when nominal GDP is divided by the population size, to give real GDP per capita (or real GDP per head) it reveals the mean average of the country's national income. It is a key measure of a country's economic growth rate and standards of living for the average person. Ceteris paribus, the larger the population size, the lower the GDP per head (see Table 18.4).

Table 18.4 GDP per capita, select countries (2020)

Country	Nominal GDP ($ billion)	GDP per capita (US$)
USA	20,544.34	54,541.70
China	13,608.15	7,755.00
India	2,718.73	2,104.20
Brazil	1,868.63	11,026.24
Indonesia	1,042.17	4,284.70
Norway	434.75	92,121.40
Kenya	87.91	1,201.10
Luxembourg	69.49	107,243.20

Source: Trading Economics **http://www.tradingeconomics.com/country-list/gdp-per-capita**

ATL ACTIVITY 18.1

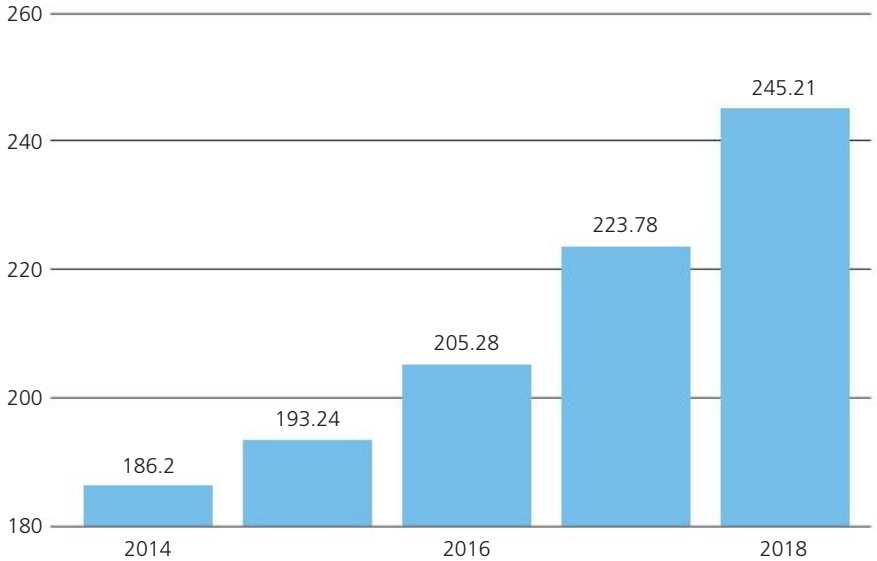
Investigate the economic growth rates for your country (or a country of your choice) over the past decade. A good starting point for economic data is to use the Trading Economics website: **https://tradingeconomics.com/countries**. Then using other online sources, investigate the main reasons for the trends in economic growth for your chosen country.

ATL ACTIVITY 18.2

A 'megacity' is a very large city, typically with a population of more than 10 million people. Data from the World Economic Forum and United Nations suggest that Tokyo will remain the world's largest city until 2030, despite its population growth being in decline. There were 34 megacities in 2020, and this is expected to increase to 43 megacities by 2030. Discuss the economic opportunities and threats of higher population density in very large cities.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 18.1

The chart below shows the nominal national output for Vietnam over a five-year period. Monetary values are expressed in US dollars (billions).



Source: Trading economics.com, World Bank
https://tradingeconomics.com/vietnam/gdp

Figure 18.6 Vietnam's GDP $bn (2014–20)

a Calculate the year in which Vietnam experienced its highest gain in nominal GDP over the period shown. *[2 marks]*

b Calculate the year in which economic growth was at its highest (round figures to 2 decimal places). *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 18.2

Study the data below and answer the questions that follow.

Country	Nominal GDP ($bn)	Population (million)
USA	20,544	327.0
China	13,608	1,400.0
Japan	4,971	126.0
UK	2,855	66.19

Source: Trading Economics **http://www.tradingeconomics.com/**

a Determine which country has the highest and which has the lowest nominal GDP per capita. *[3 marks]*

b Use the data to explain why it is important to account for the population size when comparing nominal GDP figures between different countries. *[4 marks]*

Common mistake

Be careful when interpreting percentage changes in your calculations. Too often, students comment that when the rate of economic growth drops, there is a recession. This is only true if there is negative economic growth. An economy that grew by 3 per cent last year and grows by 'only' 1 per cent this year, has still achieved positive economic growth.

Consequences of economic growth (AO3)

There are both positive and negative consequences of economic growth. These can be categorized into three areas: the impact of economic growth on living standards, the impact on the environment and the impact on income distribution.

Impact on living standards

In general, economic growth tends to lead to higher standards of living for the average person in the country. Higher real income per capita enables people to spend more money to meet their needs and wants. This has several interrelated impacts on the economy, including:

- The reduction or elimination of absolute poverty in the country owing to higher real incomes per person. As more people are able to purchase essential goods and services, absolute poverty may become less of an issue in the country.
- The creation of new jobs and lower unemployment in the economy also help to improve standards of living in the country. This helps to raise consumption and encourages further investment in capital, helping to sustain growth in the economy.
- Increased tax revenues (from taxes imposed on expenditure and income) enable the government to fund more merit goods and services. For example, increased levels of economic activity generate more tax revenues from sales taxes (on consumption or expenditure), corporate taxes (on the profits of firms) and import taxes. In the long run, this helps to further sustain the growth of the economy and to facilitate further improvements in standards of living.
- Increased consumer spending will lead to higher sales revenues for firms and greater profits. This might then lead to a rise in planned capital investment spending as firms look to increase the size of their capital stock to expand their productive capacity. For example, the increased use and reliance on mobile and wireless technologies prompts a rise in investment expenditure on broadband and internet networks.
- As income levels increase, the amount spent on demerit goods (such as tobacco, alcohol and gambling) also tends to increase. This will tend to cause a social welfare loss in the long run.
- However, the risk of inflation occurs if the economy continually grows owing to excessive aggregate demand in the country. The dangers of demand-pull inflation (see Chapter 20) can lead to prices of goods and services rising to unsustainable levels, with negative consequences on the economy, such as a decline in the country's international competitiveness. Another example is the high property prices in areas where economic prosperity has been high (see Table 18.5). According to the World Economic Forum (WEF), the average worker in Hong Kong would have to save 22 years' worth of their income in order to purchase a 650 square feet apartment in the city.

Table 18.5 The most expensive cities in the world for real estate*

Rank	City	Payback period*
1	Hong Kong	22
2	London	15
3	Paris	14
4	Singapore	12
5	Tokyo	11
6	New York	11
7	Amsterdam	10
8	Vancouver	9
9	Sydney	8
10	Munich	7

Source: The World Economic Forum (www.weforum.org/agenda/2019/03/the-worlds-most-expensive-places-to-own-a-home/)

* Number of years required to work to afford a 650 square feet apartment

Impact on the environment

Economic growth tends to come with environmental consequences. High rates of economic growth can create negative externalities (see Chapter 11) that cause problems for the environment, such as air pollution, road congestion, climate change, land erosion and the loss of biodiversity. Economic growth can therefore create market failures caused by resource depletion (such as deforestation and overfishing) and damage to ecosystems. Such environmental impacts can damage the social and economic well-being of people and their quality of life in the long run. For example, household and industrial waste have had devastating impacts on the natural environment. There is growing concern regarding the volume of plastic waste in the oceans and the exponential growth in waste disposal of unwanted clothing and consumer electronic products (e-waste).

Resource depletion by the current generation is not always sustainable, thereby threatening intergenerational equities. Hence, economic growth does not necessarily resolve a country's socioeconomic problems, such as resource depletion, market failures (see Chapter 11) and income inequality (see Chapter 19).

Figure 18.7 Environmental pollution is a drawback of rapid economic growth

Nevertheless, most economists believe that economic growth achieved through environmentally sustainable methods (such as the increased use of green technologies and renewable energy sources) is an important macroeconomic objective because it is the most practical measure of standards of living in a country. More countries are now reporting **green GDP** as a measure of gross domestic product, which adjusts the country's GDP to take account of the value of environmental degradation, damage and destruction associated with economic growth.

◆ **Green GDP** adjusts the country's GDP to take account of the value of environmental degradation, damage and destruction associated with economic growth.

Top tip!

While economic growth is generally regarded as a desirable macroeconomic objective, remember that individuals do not benefit equally. There are also opportunity costs associated with economic growth, such as negative externalities of production and consumption, and unintended consequences such as climate change and environmental damage.

Theory of Knowledge

To what extent is it possible to even know if economic growth is actually beneficial to others in society? Who are the winners and losers of economic growth?

Impact on income distribution

Economic growth often creates greater disparities in the distribution of income and wealth, widening the gap between rich and poor, that is, the rich get richer and the poor get relatively poorer. Critics argue that the pursuit of economic prosperity does not resolve socioeconomic problems such as income inequalities (see Chapter 15). Although the economy as a whole might experience economic growth, not everyone will benefit nor will they do so in the same way. *Forbes* magazine reported that in 2017, the top 10 billionaires in the USA (who all happened to be male) had a combined wealth of more than $427 billion – the equivalent wealth of the poorest 3.6 billion people on the planet! Put another way, their wealth was worth the same as the poorest half of the world's population or more than two times larger than Vietnam's GDP or 24 times greater than the GDP of Cambodia!

However, economic growth also leads to greater tax revenues (from taxes imposed on expenditure, income, profits and the sale of imports). Hence, economic growth enables the government to use these tax revenues to redistribute income and wealth in the economy.

Top tip!

Policies to promote economic growth include loose monetary policy, expansionary fiscal policy and supply-side policies. These are covered in Chapters 24, 25 and 26 respectively.

PAPER 2 EXAM PRACTICE QUESTION 18.3

According to *The Economist's* Economist Intelligence Unit, Macao's economy grew by 14.3% back in 2013 – the highest economic growth rate for any country back then. The island nation had enjoyed 9.8% growth in 2012 with gambling revenue increasing by approximately 14% to about $38 billion, making it the world's biggest gambling market ahead of Las Vegas. In 2011, Macao enjoyed a stunning 20.7% growth rate. The country is also investing huge amounts of money to attract a wider range of tourists with casino giants such as Sands and MGM Resort also investing large sums of money into the economy. However, the Covid-19 coronavirus outbreak had devastating impacts on economic activity throughout the Asia Pacific region. This had a large impact on tourism and gambling expenditure, lowering Macao's economic growth to 5.2% in 2019 and 3.3% in 2020. The government had announced plans for economic diversification, but progress on these policies was delayed until 2021.

Figure 18.8 Macao is the gambling capital of the world, in terms of dollars spent

a Define the term *economic growth*. *[2 marks]*

b Using an AD/AS diagram and information from the text above, explain how investment in Macao could boost its economic growth. *[4 marks]*

WISE ChoICES

Discuss the extent to which economic growth improves **economic well-being** without compromising **sustainability**.

Chapter summary

- Economic growth is the increase in the level of a country's real national output over time, usually a year. It is commonly measured by calculating the annual percentage change in the country's real gross domestic product (real GDP).
- Economic growth is indicative of a country's expansion of its potential productive capacity. Hence, it can be represented in a diagram by an outwards shift of the PPC.
- Economic growth can be shown diagrammatically by a rightwards shift of the LRAS curve. Alternatively, it can be shown by an outwards shift of an economy's production possibility curve.
- Causes of economic growth depend on the quality and quantity of a country's factors of production, such as its factor endowments, the size and skills of the labour force, labour productivity rates, and the amount of investment expenditure (including foreign direct investments in the economy).
- Economic growth can be measured using real gross domestic product (GDP) and expressed as real GDP per head (or real GDP per capita). Real GDP refers to the value of national income (GDP) that is adjusted for inflation.
- Nominal GDP figures are expressed as *current prices* (the actual monetary value of GDP) whereas real GDP figures are expressed as *constant prices* (the value of GDP after the effects of inflation have been removed).
- Real GDP is calculated by dividing nominal GDP by the GDP deflator (for that year, in order to account for the rate of inflation), and then multiplying the result by 100.
- Nominal gross domestic product (GDP) measures the monetary value of goods and services produced within a country, per year, without removing the impact of inflation.
- GDP per capita (or GDP per head) measures the GDP of a country divided by its population size. It is generally regarded as a more accurate representation of a country's standard of living for the average person.
- Economic growth tends to lead to higher standards of living for most people in the country. Higher real income levels enable people to spend more money to meet their needs and wants. This helps to reduce or eliminate absolute poverty in the country.
- By contrast, negative economic growth results in a recession followed by unemployment and a drop in standards of living.
- Economic growth often involves using up the world's scarce resources at rates that are not necessarily sustainable, such as deforestation and overfishing. This leads to market failures and negative externalities, causing climate change and problems to the natural environment and damage to ecosystems.

19 Macroeconomic objectives – low unemployment

Real-world issue

Why does economic activity vary over time and why does this matter?

Conceptual understandings (WISE ChoICES)

- **Change** in the conditions of the demand and supply sides of the economy cause economic activity to vary over time.
- Fluctuations in economic activity impact the **economic well-being** of individuals and societies.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- low unemployment as a macroeconomic objective (AO2)
- the measurement of unemployment and the unemployment rate (AO2)
- the unemployment rate from a set of data (AO4)
- the difficulties of measuring unemployment (AO2)
- the causes of unemployment, with reference to cyclical, structural, seasonal and frictional unemployment (AO2)
- a minimum wage diagram to show unemployment (AO4)
- diagrams that show a fall in the demand for labour for a particular market or geographical area (AO4)
- a deflationary gap diagram to show cyclical unemployment (AO4)
- the natural rate of unemployment, in terms of the sum of the structural, seasonal and frictional unemployment (AO2)
- the personal, social and economic costs of unemployment (AO2).

Low unemployment (AO2, AO4)

Employment refers to the use of factors of production in the production process. The term is usually applied to the use of labour resources. Governments strive to ensure that everyone who is available and willing to work is able to find employment. In Figure 19.1, labour market equilibrium exists when the aggregate demand for labour (AD_L) is equal to the aggregate supply of the labour (AS_L). The labour supply refers to the number of people able and willing to work at the prevailing market equilibrium wage rate (W_e). This means that everyone of working age who is willing and able to work at the prevailing market equilibrium wage rate (W_e) is in employment. **Formal sector employment** refers to officially recorded employment, with workers paying income taxes and contributing to the country's official gross domestic product (GDP).

Unemployment exists when the forces of demand and supply are in disequilibrium. Unemployment refers to the situation when people are willing and able to work and actively seeking employment but are unable to find work. It represents inefficiency in the economy due to the non-use of scarce resources in the production process.

◆ **Formal sector employment** refers to officially recorded employment, with workers paying income taxes and contributing to the country's official GDP

◆ **Unemployment** refers to the situation when people are willing and able to work and actively seeking employment but are unable to find work.

Low unemployment is a key macroeconomic objective because it:

- Complements economic growth (another macroeconomic objective) because higher employment tends to lead to greater national expenditure. Hence, low unemployment tends to increase economic well-being and the standards of living in an economy.
- Increases tax revenues from a range of sources, such as income tax (from employment), sales taxes (from increased expenditure on goods and services) and stamp duty (from the sale and purchase of private and commercial land and property). This enables the government to use the funds to further develop the economy, such as government expenditure on education, healthcare and infrastructure.

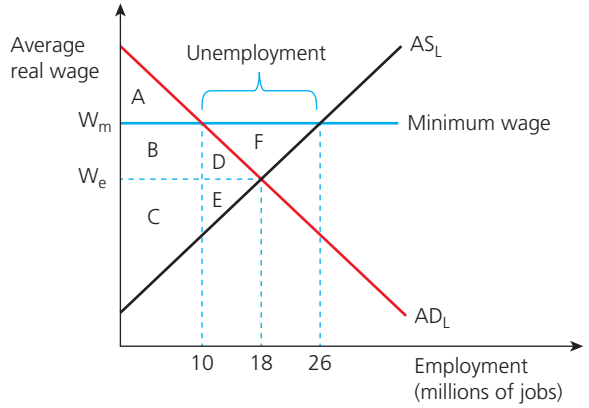


Figure 19.1 Impact of a minimum wage on the labour market

- Reduces the burden on the government because there is less need for taxpayers to fund welfare benefits as more people are in employment.
- Prevents a 'brain drain' from the economy, whereby skilled workers pursue better employment opportunities in other countries.

Figure 19.1 shows real wage unemployment, which exists when real wages are above their market clearing level due to the imposition of a minimum wage by the government.

The following can be seen in Figure 19.1:

- Equilibrium exists at a wage rate of W_e, with 18 million people employed in the labour market where the aggregate demand for labour (AD_L) is equal to the aggregate supply of labour (AS_L) in the economy.
- The higher wage rate (W_m) raises labour costs for firms, so reduces the AD_L to 10 million but there is a greater incentive for workers to offer their labour services thus increasing the AS_L to 26 million. Hence, the minimum wage risks causing unemployment owing to the excess supply of labour (26 million minus 10 million). In fact, any wage rate above the equilibrium causes excess supply (shown by area F).
- Area A shows the employer surplus as some firms are willing and able to pay more than the minimum wage (W_m).
- Area B+C shows the employee surplus as there are 10 million people willing and able to work for less than W_m but receive this as a higher wage rate.
- The minimum wage causes a market disequilibrium. There is market inefficiency, represented by a welfare loss to society, shown by area D+E.

Measurement of unemployment and the unemployment rate

The generally accepted definition of unemployment refers to people who are both able (available) and willing to work at the prevailing real wage rate but cannot find a job, despite actively seeking employment. The **unemployment rate** measures the number of people who are officially unemployed as a percentage of the country's workforce, per time period. This is calculated by using the formula:

$$\text{Unemployment rate} = \frac{\text{Number of unemployed people}}{\text{Labour force}} \times 100$$

◆ The **unemployment rate** measures the percentage of the labour force that is unemployed. It is calculated by the formula: (Total unemployed ÷ Total labour force) x 100.

The **labour force** consists of the employed, the self-employed *and* the unemployed, that is, all those people in employment plus all those actively seeking employment. For example, suppose a country has a workforce of 50 million people, of which 5 million people are actively seeking employment but unable to find jobs. This means the country's unemployment rate is 10 per cent (that is, 5m ÷ 50m).

◆ The **labour force** refers to people of working age who are in employment (including the self-employed) plus those who are seeking work, that is, those who are available for work at current real wage rates.

The United Nation's **International Labour Organization** (ILO) states 15 as the minimum age to enter the labour force. There is no official upper limit, but many countries use a range between 65 and 70. For example, the official retirement age for men and women is 67 years in Norway, Italy, Iceland and Greece. The retirement age is 60 years in Japan, South Korea, Malaysia, Peru and Saudi Arabia. Anyone aged between 15 and the official retirement age and is both able and willing to work is considered part of the workforce. This does mean that full-time students aged 15+ are not considered, as they are not available to work.

The ILO measures unemployment based on the number of people who are:

- willing to work, but unable to find employment
- actively looking for work, that is, they have looked for a job in the last four weeks, and
- able to start work within the next two weeks (that is, they are readily available for work), or
- waiting to start a new job within the next two weeks.

An alternative to the ILO measure of unemployment is the **claimant count**. This measure includes people who are eligible to claim unemployment benefits. There are specific requirements to qualify for unemployment benefits, which vary from country to country but typically include:

◆ The **claimant count** is a measure of the number of people who are out of work, actively seeking employment but claiming unemployment benefits.

- being 18 years old or above, and not in full-time education or government training schemes
- being available (able) to work and actively seeking (willing to) work
- have household savings below a certain amount, which is determined by the government and reviewed on an annual basis to reflect changes in the cost of living
- part-time workers who may also qualify for unemployment benefit in some cases, if they do not work more than a stipulated number of hours per week.

However, one criticism of using the claimant count to measure unemployment is that it does not necessarily reflect the true level of unemployment in the economy because some people may be deterred from claiming as they cannot prove they are actively looking for work, such as part-time workers. In addition, some individuals may make fraudulent claims in order to qualify for welfare benefits. Hence, it is generally recognized that the claimant count of unemployment underestimates the actual figures.

Irrespective of the measure or definition of unemployment, it represents an inefficient use of any economy's scarce resources, which hinders its potential national output.

WORKED EXAMPLE

Using the data below, calculate the total number of people unemployed. *[2 marks]*

Labour force	30 million
Population of working age	35 million
Unemployment rate	7.9%

Answer

- The total number of people classified as unemployed = 7.9% of 30 million (not everyone of working age is willing and able to work)
- 30m × 0.079 = 2.37 million people

Common mistake

Students often define or comment on unemployment being a situation when people do not have a job. This is not entirely accurate as there are people who are not willing and/or able to work, even if they are of working age; for example, full-time university students or those who have chosen to retire early. Economists consider unemployment to exist if people are both able (available) and willing to work but cannot find employment despite actively trying to do so.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 19.1

Use the data below to calculate the unemployment rate. *[2 marks]*

Total population	100 million
Percentage of population employed	74.8%
Population of unemployed	15 million
Retired population	20%

PAPER 3 EXAM PRACTICE QUESTION 19.2 (HL ONLY)

The diagram below shows the aggregate supply of labour (AS_L) and the aggregate demand for labour (AD_L) in the economy. Explain why a real wage rate of W_2 will cause real-wage unemployment in the economy. *[4 marks]*

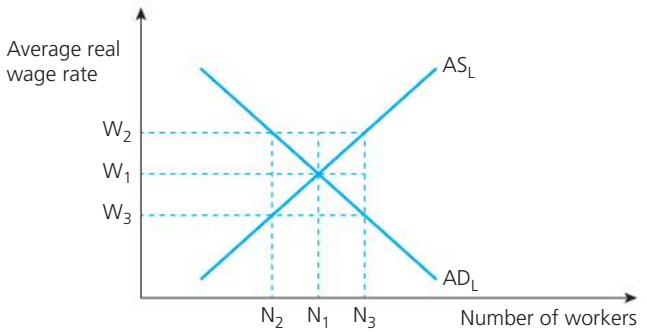


Figure 19.2 Impact of a minimum wage

Difficulties of measuring unemployment (AO2)

Owing to the various different interpretations and measures of unemployment, it can be difficult to accurately measure unemployment. Reasons for this include: hidden unemployment, voluntary unemployment, underemployment and disparities in the measurement and reporting of unemployment in the economy.

Hidden unemployment

The concept of **hidden unemployment** is important to economists. There are plenty of people who are not considered in the official calculation of unemployment statistics (purely owing to the choice of measurement) rather than the core meaning of unemployment (the redundancy or non-use of factors of production). As these people escape the official measure of unemployment, there is an underestimation of the true or actual rate of unemployment. Hence, hidden unemployment is often referred to as **disguised unemployment**.

One major reason for hidden unemployment is the existence of discouraged workers. **Discouraged workers** are people who have stopped searching for jobs because they have become disheartened having been rejected by prospective employers too many times. Hence, they are not included in the labour force as they are no longer actively looking for a job. Both the ILO and claimant count measures of unemployment consider only those who are actively looking for a job. Hence, discouraged workers result in an underestimation of the economy's unemployment rate.

◆ **Hidden unemployment** (or **disguised unemployment**) refers to those who are technically unemployed – as they do not have full-time jobs – but are not included in the official measurement of unemployment; for example, discouraged workers and underemployed workers.

◆ **Discouraged workers** are people who have stopped looking for jobs as they have become disheartened having tried but been rejected by prospective employers too many times.

Voluntary unemployment

There are also economically inactive people who are **voluntarily unemployed** – people of working age who choose not to actively look for work. Some of the reasons for being voluntarily economically inactive include:

- people who look after elderly relatives or those with long-term sickness
- parents who choose not to work full-time in order to care for their young children
- individuals who have decided to take early retirement.

◆ **Voluntarily unemployed** people are economically inactive individuals who choose not to pursue full-time employment.

Figure 19.3 Should unpaid volunteers be considered as employed or unemployed?

These people are classed as being economically inactive rather than being considered as part of the unemployed population.

Underemployment

Underemployment exists when people are inadequately employed, reflecting the underutilization of the labour force. Although they are technically employed, the underemployed are not able to fully use their skills or abilities. For example:

- Involuntary part-time workers who cannot find full-time jobs. Although part-time workers can be officially classed as employed (depending on how many hours they work per week), they may be working fewer hours than they would like.
- Skills mismatch which results in highly skilled, qualified and experienced people working in relatively low-paying occupations owing to limited job opportunities.

The underutilization of the economy's human resources is inefficient and limits the economy's productivity. This has detrimental impacts on economic growth and the country's international competitiveness.

◆ **Underemployment** refers to the underutilization of the labour resources in the economy, which limits the economy's productivity and efficiency.

Disparities

The measurement of unemployment is an average measure for the whole economy at one point in time. It ignores disparities, such as:

- **Regional disparities** – Different regions of the country face different rates of unemployment. For example, unemployment in 2019 for the USA averaged 3.6 per cent, but Alaska recorded 6.1 per cent unemployment while South Carolina and Utah saw only 2.3 per cent unemployment.
- **Ethnic disparities** – Ethnic minority groups tend to suffer from higher than average rates of unemployment, and for longer periods of time. For example, the unemployment rate for African-Americans in New York is three times higher than the average official unemployment rate.
- **Age disparities** – Unemployment rates among the young and older people in the workforce are higher than those officially reported for the nation. Following the global financial crisis of 2008, Greece and Spain experienced unemployment of around 28 per cent in 2013, although youth unemployment reached 58.7 per cent and 56.1 per cent respectively.
- **Gender disparities** – Females tend to experience higher average rates of unemployment than males. On average, men also re-enter the labour market quicker following a period of unemployment. According to the ILO, gender inequalities in unemployment rates are exceptionally high in the Middle East and North Africa. An alternative or additional measure is therefore to calculate the **female participation rate** – the proportion of women who are economically active in the labour force.

◆ The **female participation rate** measures the proportion of women who are economically active in the labour force.

Top tip!

Remember that not all people of working age are *willing* and/or *able* to participate in full-time employment. This will partly depend on the country's welfare benefits scheme and social attitudes towards women in the workforce. For example, women in Saudi Arabia face huge barriers to entering the labour force. In 2019, the Middle Eastern nation ended a 61-year ban on females being allowed to drive. This created huge employment opportunities for 15 million women, who were finally able to have the freedom of movement enjoyed by their male counterparts.

Theory of Knowledge

Discuss the following phrases:

a voluntary unemployment
b voluntary employment (voluntary workers)
c unemployment benefits
d constant change.

Are they oxymoronic? What do these phrases tell us about the relationship between knowledge and language, or between knowledge and politics?

PAPER 2 AND 3 EXAM PRACTICE QUESTION 19.3

According to the International Monetary Fund (IMF), Pakistan's annual unemployment rate between 2010 and 2020 was kept steady at around 6%. This, according to the Central Intelligence Agency (CIA) meant Pakistan's gross domestic product grew by 4.24% in 2018 and another 5.3% in 2019. These changes have helped to reduce some of the poverty in the country with a population of more than 212 million people.

a Explain **two** reasons why it might be difficult at times to determine the exact rate of unemployment in a country. *[4 marks]*

b Explain **two** likely consequences of low unemployment for the Pakistani economy. *[4 marks]*

Causes of unemployment (AO2, AO4)

The main causes of unemployment can be categorized in terms of the four main types of unemployment: cyclical (demand deficient), structural, seasonal and frictional unemployment.

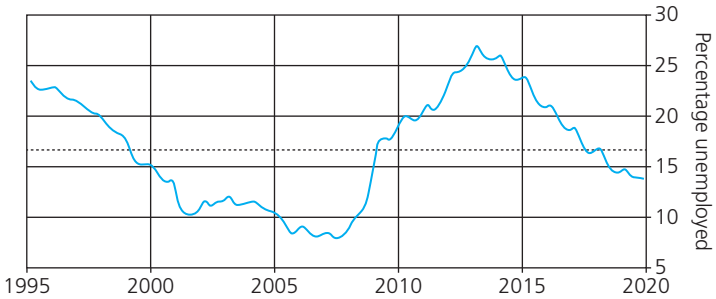
Cyclical unemployment

Cyclical unemployment is unemployment that results from a downturn in the business cycle (see Chapter 16). It is caused by a lack of aggregate demand in the economy, which causes a fall in real national output (see Figure 19.5) and therefore job losses. Hence, cyclical unemployment is also referred to as **demand deficient unemployment**. It is the most severe type or cause of unemployment as it can affect every industry in the economy. For example, car manufacturers will make some workers redundant during a recession as fewer people buy new cars owing to a fall in their real national income. When the demand for cars and related products fall (such as petrol, insurance and servicing), the derived demand for workers in the car manufacturing and related industries also declines as the automakers no longer need to hire as many employees.

◆ **Cyclical unemployment** (or **demand-deficient unemployment**) is unemployment caused by a lack of demand for goods and services in the economy.

Cyclical unemployment usually results in mass job losses as firms try to control costs, protect their profitability and prevent business failure. For example, the global financial crisis of 2008 caused huge unemployment in Spain, rising from 8 per cent before the crisis to 26.94 per cent in 2013 (see Figure 19.4).

Demand deficient unemployment is associated with a decline in the business cycle (during recessions and slumps). In Figure 19.5, the decline in aggregate demand from AD_1 to AD_2 causes national output (or real GDP) to fall from Y_1 to Y_2. This creates widespread unemployment in the economy, and therefore a fall in the general price level from PL_1 to PL_2.



Source: Trading Economics https://tradingeconomics.com/spain/unemployment-rate

Figure 19.4 Spain's unemployment rate (%) (1995-2020)

Alternatively, cyclical unemployment can be represented using an AD–AS diagram that shows a **deflationary gap** (see Figure 19.6). A deflationary gap (also called a **recessionary gap** or **negative output gap**) occurs when the economy operates below the full employment equilibrium, that is, the level of aggregate demand is insufficient (deficient) to create full employment.

◆ A **deflationary gap** is the negative output gap in the economy, shown by the difference between the level of aggregate demand and that needed to achieve full employment.

The deflationary gap is shown by the difference between the full employment level of national output (Y_f) and the actual national output (Y_e) in the short run. (P_e and Y_e) is below the potential output at full employment (Y_f). Therefore, the government could choose to use loose monetary policy and/or expansionary fiscal policy (see Chapters 24 and 25) to raise aggregate demand from AD_1 to AD_2, the level that matches the long run aggregate supply (LRAS). This helps to restore full employment equilibrium at Y_f. Hence, closing the deflationary gap helps to reduce cyclical unemployment in the economy.

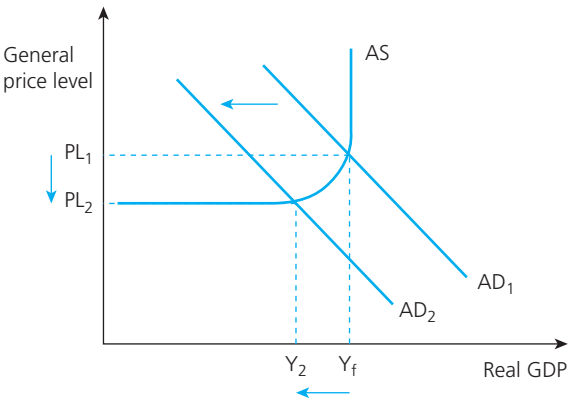


Figure 19.5 Demand deficient unemployment

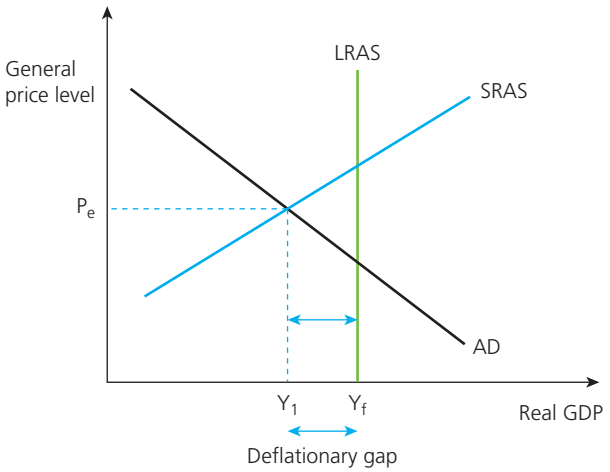


Figure 19.6 Unemployment and the deflationary gap

Theory of Knowledge

Recession-proof occupations might include the following: government officials (civil servants), doctors in the public sector, private sector teachers and jobs at fast food retailers such as McDonald's. Why do you think this is? What does it say about human behaviour's relationship with knowledge and evidence?

Structural unemployment

Structural unemployment occurs when the demand for goods and services in a particular industry continually falls, thereby reducing the demand for particular labour skills, that is, when there is a long-term (permanent) decline in the demand for labour in a specific industry. It arises from the skills mismatch between the jobs available in the economy and the skills of the available workers in the labour market.

◆ **Structural unemployment** exists when the demand for labour is less than the supply of labour in a particular industry.

Structural unemployment can also be caused by changes in geographical locations of industries (for cost advantages) and labour market rigidities (such as the unwillingness of workers to accept lower wage rates). It can also arise owing to physical capital, such as robots and automation, replacing human capital in the production process. For example, the UK has suffered from structural unemployment in the motor manufacturing, shipbuilding, coal mining, steel manufacturing and textiles industries. Typically, these industries locate in certain areas of regions of the country, so structural unemployment is often associated with regional unemployment.

CASE STUDY

The Full Monty (movie)

The Full Monty (1997) is one of the UK's most successful movies, with a production budget of just $3.5 million but grossing over $258 million at the box office. The movie is set in Sheffield, UK, telling the story of six fictitious unemployed men who struggle to find alternative jobs after structural unemployment hits the British steel industry in the 1970s. Despite being classed as a comedy, the movie also tells of the dangers and seriousness of mass-scale structural unemployment, including poverty, depression, family issues and mental health issues.

Structural unemployment is often permanent because it is difficult for many of the unemployed to learn new skills (occupational immobility) or to move to a different area where their skills are still in demand (geographical immobility). Jobs lost owing to changing market trends means that new jobs are needed but these usually demand different skills and qualifications. Hence, government-funded education and retraining programmes are common solutions used to tackle structural unemployment in order to help these people to find alternative employment.

Figure 19.7 shows the impact of a fall in the demand for labour in a particular market or in a specific geographical area due to structural unemployment. This will bring about a decline in the derived aggregate demand for labour in the industry from AD_{L1} to AD_{L2}, which causes the employment level to drop from N_1 to N_2. The lower demand for labour in the industry also causes the wage rate to fall from W_1 to W_2. As a consequence of such unemployment, the industry and the associated geographical area suffer from structural and long-term changes in demand. It is usually difficult for those in structural unemployment to find a new job without some assistance and retraining.

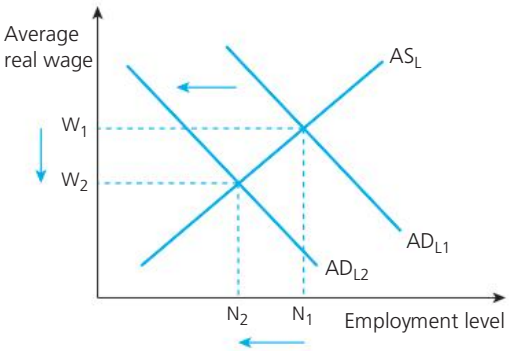


Figure 19.7 The impact of structural change on the demand for labour

WISE ChoICES

One of the most common causes of structural unemployment in many countries is the growing integration and interdependence of the world's economies (often referred to as *globalization*). The growing **interdependence** through international trade deals, for example, has caused many multinational companies to relocate their operations to more cost-effective countries. However, this can unintentionally cause structural unemployment in the domestic economy. It is usually difficult for the structurally unemployed to find other jobs as their skills and specialized background are not needed in the current labour market.

To what extent does **interdependence** on a global scale offer less **choice** to individuals and societies?

Seasonal unemployment

Seasonal unemployment is unemployment caused by regular and periodical changes in demand for certain goods and services at different times of the year. There are seasonal peaks and troughs in the production and sales for many products. For example, fruit pickers are in high demand during the summer months while retailers in many parts of the world hire more temporary workers during the Christmas trading season. Other examples of seasonal unemployment include:

- ice-cream retailers during the winter
- construction workers during the winter season
- ski instructors during the summer
- hand car wash firms during the monsoon (rainy) season
- school bus drivers and teaching assistants during school holidays.

Seasonal unemployment is particularly harmful to those on low incomes trying to manage their income and expenditure throughout the year. Areas of the country that suffer from high seasonal unemployment become relatively unattractive, which eventually leads to local workers relocating to other areas to seek better employment opportunities. Owing to the nature of seasonal unemployment, official unemployment statistics are often seasonally adjusted. This means the measure of unemployment is modified to account for the usual seasonal fluctuations in demand for certain goods and services.

Possible solutions to dealing with seasonal unemployment include:

- diversifying the economy to cater for goods and services irrespective of the seasons, such as developing opportunities to provide tourism services during the winter months
- government retraining programmes to improve the occupational and/or geographical mobility of those who are seasonally unemployed.

◆ **Seasonal unemployment** is unemployment caused by regular and periodical changes in the derived demand for labour at different times of the year.

Figure 19.8 Strawberry and other fruit pickers – a seasonal occupation

Frictional unemployment

Frictional unemployment (also called **transitional unemployment**) occurs when people are in transition between jobs due to the time delay between leaving a job and finding or starting a new one. There is a delay because the labour market is unable to immediately match demand (people looking for suitable jobs based on their interest and skills set) with the supply of appropriate jobs. It takes time for people to find and apply for the right jobs and for employers to recruit the right people. The greater the labour market imperfections, the longer this delay will tend to be.

◆ **Frictional unemployment** is a cause or type of unemployment that exists when people are temporarily unemployed while they actively search for a new job or are waiting to start a new job.

Examples of frictional unemployment include:

- people quitting their jobs in search of better-paid and/or more fulfilling jobs
- those who leave their previous job because they were unhappy working there
- some people may take time off to look after their family, such as elderly parents or young children
- people who take a sabbatical (like a gap year but for full-time employees) in order to pursue further studies and/or to travel
- those who are of working age but retire early only to have the desire at a later date to re-enter the world of world.

Nevertheless, frictional unemployment is unavoidable and can be a healthy sign of the labour market as workers switch to more meaningful and rewarding jobs. The constant creation and destruction of jobs also creates frictional unemployment. Frictional unemployment is always present because it takes time for the labour market to match available jobs with suitable people looking for jobs. Hence, frictional unemployment can occur even when there is full employment (the situation when anyone who wants a job at the prevailing market wage rate is able to have one). However, in a recession, frictional unemployment is likely to fall as workers feel less confident in being able to find alternative employment.

Natural rate of unemployment (AO2)

The natural rate of unemployment (NRU) refers to the equilibrium rate of unemployment, determined by calculating the level of unemployment when the labour market is in equilibrium. This means the NRU occurs where the aggregate supply of labour (AS_L) is in balance with the aggregate demand for labour (AD_L). Hence, at the NRU, everyone who wants to work at the prevailing real wage rate has a job. In other words, there is no involuntary unemployment.

Some voluntary unemployment exists at the NRU because various people remain out of a job while searching for employment opportunities that offer better real wages or improved terms and conditions of employment. Hence, the natural rate of unemployment is the sum of structural, seasonal and frictional unemployment in the economy.

Figure 19.9 shows the aggregate demand for labour (AD_L) and aggregate supply of labour (AS_L) curves. The labour market is in equilibrium at a real wage rate of W_e, with N_L workers employed. However, at this wage rate, the labour force (N_F) is greater than the number of people willing and able to work at W_e, so voluntary unemployment exists. The horizontal distance between the supply of labour curve and the labour force (LF) at each real wage rate

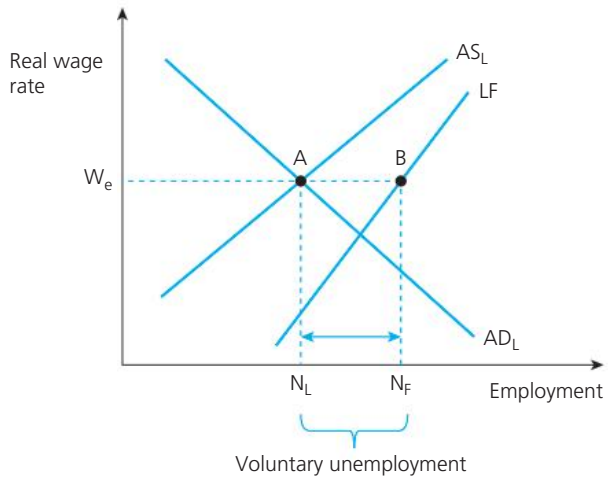


Figure 19.9 The natural rate of unemployment

represents natural unemployment. For example, at W_e, voluntary unemployment is the distance between A and B, or N_L to N_F, which represents the natural unemployment. This consists of frictional, seasonal and structural unemployment.

Top tip!

The higher the real wage rate, the lower the amount of voluntary or natural unemployment as there are greater incentives for people to work. In theory, at a high-enough real wage rate, everyone who is able and willing to work can find employment. At this point, the AS_L and LF curves converge.

Policies to reduce the natural rate of unemployment usually focus on correcting labour market imperfections by making labour markets more competitive and flexible. These include:

- Reforms to the welfare benefits system to create incentives for people to work. For example, market economists argue that the existence of unemployment benefits incentivizes some people to be voluntarily unemployed.
- Reforms to trade unions by reducing their collective bargaining power. Labour unions often cause labour market rigidities by pressurizing governments to increase the minimum wage and terms and conditions of employment. These matters tend to raise costs for employers, thereby preventing labour markets from reaching an equilibrium.
- Reducing barriers to labour mobility so that people can find jobs in different industries, thereby reducing frictional, seasonal and structural unemployment in the economy. For example, the government could provide funding for job centres and employment agencies to help the unemployed find jobs.
- Personal income tax reductions can help improve incentives to work. Lower marginal rates of tax (see Chapter 23) can encourage people to find and accept paid employment.

Costs of unemployment (AO2)

The costs of unemployment can be categorized as personal costs, social costs and economic costs, although in reality these are all interrelated.

The **personal costs** of unemployment include the following:

- **Stress** – The unemployed are likely to suffer from stress and depression. In extreme cases, unemployment can even lead to suicides.
- **Low self-esteem** – The long-term unemployed may lack self-confidence, self-worth and self-respect. This can be damaging for both their physical and mental well-being.
- **Poverty** – The lack of income means the unemployed may experience poverty, including hunger, malnutrition and poor health. Prolonged periods of unemployment can lead to personal bankruptcy and homelessness.
- **Family breakdowns** – Unemployment often causes family arguments, including separation and divorce.

The **social costs** of unemployment include the following:

- **Crime and anti-social behaviour** – As unemployment causes personal costs such as deprivation and desperation, this can lead to increased crime and antisocial behaviour, such as theft, alcoholism and vandalism.

Figure 19.10 Social deprivation is a major cost of unemployment

- **Indebtedness** – Lower income, caused by unemployment, leads to increased indebtedness for individuals, firms and the government. Persistent indebtedness, especially unaffordable debt, can cause bankruptcies, which can cause impoverishment, hunger, disease, drug addiction, homelessness, crime and other social problems.
- **Social deprivation** – The local community can suffer if there is mass unemployment, such as structural unemployment in a particular region of the country. These social costs include absolute poverty, falling house prices (and other asset values), and increased crime rates.

The **economic costs** of unemployment include the following:

- **A loss of GDP** – Lower gross domestic product (negative economic growth) has detrimental consequences on the economy including a fall in its international competitiveness, that is, its ability to compete with foreign rivals, both domestically and in overseas markets. Whichever measure or definition of unemployment is used, it represents an inefficient use of the economy's scarce resources. Therefore, unemployment hinders the economy's potential output.
- **Loss of tax revenues** – Unemployment results in lower income and expenditure, thus resulting in lower tax revenues for the government. This limits the ability of the government to take corrective measures to stabilize the economy.
- **Increased cost of unemployment benefits** – Unemployment creates an increased opportunity cost of government expenditure on unemployment and welfare benefits. Prolonged periods of high unemployment can therefore lead to increased government debts, especially as tax revenues are likely to fall during periods of prolonged unemployment.
- **Loss of income for individuals** – Unemployment results in lower household income with negative consequences for individuals and their families, and therefore the economy as a whole. This is due to the multiplier effect (see Chapter 25). If prolonged, unemployment can cause (or increase) mass-scale poverty in the economy and further negative economic growth.
- **Greater disparities in the distribution of income and wealth** – As women, young adults, ethnic minority groups and those living in rural areas tend to suffer more from prolonged periods of unemployment, the result will be greater discrepancies in the distribution of income and wealth in the economy.

ATL ACTIVITY 19.1

An inquiry into the costs of unemployment – investigate the economic costs of unemployment for different stakeholders in a country of your choice.

Chapter summary

- Governments strive to ensure that everyone who is able and willing to work is able to find employment.
- Lowering unemployment helps the economy to make the most of its labour resources.
- The *unemployment rate* measures the number of unemployed as a percentage of the country's workforce, per time period.
- The *labour force* consists of all those who are in employment (the employed plus the self-employed) *and* those actively seeking employment (the unemployed).
- The *claimant count* refers to the number of people receiving unemployment benefits, expressed as a percentage of the labour force.
- The ILO measure of unemployment considers the number of people who are willing to work but unable to find employment, actively looking for work, able to start work and/or waiting to start a new job.
- Difficulties in measuring unemployment include hidden unemployment (employed people who are excluded from the measure of unemployment owing to the chosen definition of unemployment), underemployment (people who want full-time jobs but are only able to get part-time jobs), voluntary unemployment and disparities (such as regional, gender, age and ethnic disparities).
- Cyclical unemployment is caused by a downturn in the business cycle, resulting in mass job losses in the economy, as firms try to control costs, protect their profitability and prevent business failure.
- Structural unemployment occurs when the demand for labour is less than the supply of labour in a particular industry, that is, there is a decline in the industry, which typically causes regional unemployment.
- Seasonal unemployment is caused by regular and periodical changes in the derived demand for labour during different times of the year. It is particularly harmful to individuals and households on low incomes.
- Frictional unemployment is the result of the delay between workers leaving a job and finding or starting a new one. It is always present and naturally occurs, even in a growing and stable economy.
- Unemployment causes personal, social and economic costs, such as stress, social deprivation and lower economic growth.

20 Macroeconomic objectives – low and stable rate of inflation

Real-world issue

Why does economic activity vary over time and why does this matter?

Conceptual understandings (WISE ChoICES)

- **Change** in the conditions of the demand and supply sides of the economy cause economic activity to vary over time.
- Fluctuations in economic activity impact the **economic well-being** of individuals and societies.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- why a low and stable rate of inflation is an a macroeconomic objective (AO2)
- the inflation rate, using consumer price index data (AO2, AO4)
- a weighted price index, using a set of data provided (AO4) (HL only)
- the inflation rate from a set of data using quantities purchased as weights in the CPI (AO4)
- the limitations of the CPI in measuring inflation (AO2)
- the difference between demand-pull and cost-push inflation (AO2)
- diagrams that show demand-pull inflation (AO4)
- diagrams that show cost-push inflation (AO4)
- the costs of a high inflation rate (AO2)
- the causes of deflation, with reference to changes in AD and/or SRAS (AO2)
- the difference between disinflation and deflation (AO2)
- the costs of deflation (AO2)
- a diagram that shows deflation (AO4)
- the relative costs of unemployment versus inflation (AO3).

Low and stable rate of inflation (AO2, AO4)

Inflation is the sustained (continual) rise in the general price level in an economy over a given period of time. This does not mean the price of every good and service increases, but that prices are rising on average. Therefore, people need to spend more money to buy the same amount of goods and services as before. Governments set a target inflation rate as a key macroeconomic objective as inflation reduces the purchasing power of money and a country's international competitiveness. **Price stability** exists when the general level of prices remains broadly constant. Price stability does not necessarily mean there is zero inflation in the economy but a low and stable rate of inflation.

◆ **Inflation** is the sustained rise in the general price level in an economy over time, that is, prices in the economy are, on average, rising over time.

◆ **Price stability** occurs when the general level of prices remains largely constant due to a low and stable rate of inflation in the economy.

CASE STUDY

A case of hyperinflation

Hyperinflation (also known as **runaway inflation**) refers to extortionately high and uncontrollable rates of inflation that causes serious macroeconomic problems. For example, in 2019, Venezuela's economic crisis caused inflation to exceed 1.3 million per cent! However, this is still relatively low compared to Zimbabwe's hyperinflation rate of 231 million per cent in July 2008 due to the country's gross mismanagement of the economy between 2003 and 2009. This was not helped by the country printing banknotes with a value of 100 trillion Zimbabwean dollars (ZWD100,000,000,000,000)! This meant the currency became literally worthless, and was eventually abandoned. Today, the southern African country still does not have its own official currency, with many of the nation's 17.3 million people preferring to use the US dollar, South African rand or UK pound sterling. According to the World Bank, Zimbabwe's GDP per capita is only $1,322 (around $3.62 per day). This is equivalent to just 10 per cent of the world average.

Figure 20.1 The 100 trillion dollars bill was introduced in 2009

Top tip!

Inflation is not necessarily detrimental to the economy as a low and stable rate of inflation is a sign of economic growth. Low rates of inflation, of 1 or 2 per cent, are not usually harmful to the economy as higher prices can encourage firms to supply more output. It is when inflation rises too quickly that it can disrupt decision-making for individuals, firms and governments.

◆ **Hyperinflation** (or **runaway inflation**) refers to extortionately high and uncontrollable rates of inflation that causes serious macroeconomic problems.

Measuring the inflation rate (AO2, AO4)

The **consumer price index** (CPI) is a weighted index of average consumer prices of goods and services over time. It is the most common method used to measure inflation (and therefore changes in the cost of living) for the typical or average household in the economy. Calculating changes in the CPI will give the rate of inflation.

◆ The **consumer price index** (CPI) is a weighted index of average consumer prices of goods and services over time, used to measure inflation (or changes in the cost of living) for the average household in the economy.

A **base year**, with an assigned index number of 100, is used as the starting period when calculating a price index such as the CPI. So, a price index of 115.2 means that prices have, in general, increased by 15.2 per cent since the base year. Changes in the index number are used to show inflation in subsequent years. Hence, if prices were to rise by another 5 per cent in the following year, the price index number would become 120.96 (that is, 115.2 × 1.05), that is, on average, prices are 20.96 per cent higher since the base year. An increase in the consumer price index (that is, inflation) will cause a fall in the purchasing power of money in the economy.

In practice, price changes in the CPI are measured on a monthly basis but reported for a twelve-month period. There are two steps to do this:

1. Collection of the price data – a sample of prices are collected from a range of retail locations (such as supermarkets and the internet) for a selection of goods and services that are representative of the average household's expenditure. The price data are collected on a monthly basis.
2. Assigning statistical weights – each item of expenditure in the representative list of goods and services is assigned a different statistical weight to represent different patterns of spending

over time. The statistical weights are used to reflect the relative importance of household spending on these items. There are two ways to apply these weights to the CPI:

- Volume of quantities purchased – It is assumed that the more times and the greater the quantities purchased, the more important the item is to the average household. The weights are therefore based on quantities purchased per month. So, for example, if the typical family buys more food items than clothes or candles per month, then food products will have a larger weighting in the CPI.
- Value of quantities purchased – Similarly, it can be assumed that the more money that an average family spends on a product as a proportion of its overall spending, the more important that product is to them. So, if the typical household spends more money on housing expenses (such as rents or mortgages) than on food or clothing, housing costs would have a larger statistical weight applied in the CPI than food or clothing.

Statistical weights are usually determined by the percentage of household income spent on the goods and services in the representative basket for the average family. Therefore, the statistical weights for goods and services that take up a greater proportion of the typical household's income are assigned a greater value. For instance, in any given month, the typical household in the economy will spend more money on petrol (gas) than on strawberries. Hence, the statistical weight assigned to spending on petrol is greater than that for strawberries, because changes in the price of petrol will have a more significant impact on the average household than changes in the price of strawberries. Essentially, a 10 per cent increase in the price of petrol will affect most people far more than a 50 per cent increase in the price of items such as strawberries, light bulbs, batteries or socks.

Typically, the prices of items such as staple food products, clothing, petrol and transportation are included in the CPI. However, the statistical weights in the CPI are revised periodically in order to reflect changes and trends in the consumption expenditure of a typical household in the economy. For example, greater household expenditure on smartphones, online apps, tablet computers and internet services (broadband service providers and online streaming services such as Netflix or Amazon Prime Video) require a review or update on what is included in the CPI and their relative weights.

WORKED EXAMPLE

The simplified example below, with three products in the representative basket of goods and services, shows how an unweighted CPI is calculated. Assume 2017 is the base year, when the total price of the basket of the three products was $20.

Product	Price in 2019	Price in 2020
Pizza	$9	$10
Cinema ticket	$10	$11
Petrol	$3	$3.5
Total basket price	$22.0	$24.5

To calculate the inflation rate between 2019 and 2020, first calculate the price indices for the two years:

- 2019: $22 ÷ $20 × 100 = 110.0 (that is, prices in 2019 were 10% higher on average than in 2017).
- 2020: $24.5 ÷ $20 × 100 = 122.5 (that is, prices in 2020 were 22.5% higher on average than in 2017).
- The inflation rate between 2019 and 2020 is the percentage change in the price indices during these two time periods, that is, [(122.5 – 110) ÷ 110] × 100 = **11.36%**.

WISE ChoICES

Investigate how the goods and services in the CPI have been subject to change in a country of your choice. To what extent do the items in the CPI improve economic well-being in the country?

WORKED EXAMPLE

The example below is an extension of the above, showing how a weighted CPI is calculated. This method uses the volume of quantities purchased per time period as the statistical weight. Assume that the same quantities were purchased in 2019 and 2020.

Item	Price per unit in 2019	Price in 2020	Quantity in basket	Value of basket in 2019	Value of basket in 2020
Pizza	$8	$10	6	$48	$60
Cinema ticket	$9	$10	3	$27	$30
Public transport	$4	$5	15	$60	$75
Total basket price	$23	$26		$135	$165

To calculate the inflation rate between 2019 and 2020, first calculate the value of the basket of goods and services for both years:

- 2019: ($8 × 6) + ($9 × 3) + ($4 × 15) = $135
- 2020: ($10 × 6) + ($10 × 3) + ($5 × 15) = $135
- The inflation rate between 2019 and 2020 is the percentage change in the weighted price of the basket of goods and services, that is, [(165 – 135) ÷ 135] × 100 = **22.22%**.

WORKED EXAMPLE

In reality, the products in the CPI are of different importance to the average household, so statistical weights are applied to reflect this. As another example, suppose that food consumption takes up 40 per cent of average household spending, entertainment represents 20 per cent, transport equals 25 per cent and all other items account for the remaining 15 per cent of household consumption. To create a weighted CPI, economists multiply the price index of each item by the statistical weight for the item. Applying the weights gives the following results:

Item	Price index	Weight	Weighted CPI
Food	110.0	0.40	110 × 0.4 = 44.0
Entertainment	115.0	0.20	115 × 0.2 = 23.0
Transport	116.4	0.25	116.4 × 0.25 = 29.1
Others	123.3	0.15	123.3 × 0.15 = 18.5
Weighted index			114.6

While the price of food has increased the least (only 10 per cent), the spending on food accounts for 40 per cent of the typical household, so higher food prices have a much larger impact on the cost of living. Without using weights, the general price index would be 116.18, that is, (110 + 115 + 116.4 + 123.3) ÷ 4. However, the weighted index reduces the CPI to 114.6 because the relatively higher prices of the non-food items account for a smaller proportion of spending by the typical household. Therefore, the use of a weighted CPI is more accurate in measuring changes in inflation and, hence, the cost of living for the average household.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 20.1

Study the following data and answer the question that follows.

Year	Inflation rate (%)	Wage increase (%)
1	2.5	3.0
2	3.1	3.5
3	2.9	3.1

a Identify the year in which there was the largest increase in real wages. Explain your answer. *[3 marks]*

b Explain why average wages were higher in Year 3 than in Year 2. *[3 marks]*

ATL ACTIVITY 20.1

Investigate the components of the CPI for a country of your choice. How have the weights for different products changed over time, and what might this suggest about the society?

PAPER 2 AND 3 EXAM PRACTICE QUESTION 20.2

The data below show the inflation rates for a country over three years.

Year	1st	2nd	3rd
Inflation rate (%)	2.5	1.7	2.3

a Define the *inflation rate*. *[2 marks]*

b Explain why inflation was at its highest level in the third year. *[2 marks]*

PAPER 3 EXAM PRACTICE QUESTION 20.3 (HL ONLY)

Calculate the weighted price index from the information below. *[3 marks]*

Item	Price index	Statistical weight
Food and drink	120	10
Transportation	130	20
Leisure and entertainment	140	30
Housing	150	40

PAPER 2 AND 3 EXAM PRACTICE QUESTION 20.4

Calculate the inflation rate if the consumer price index changes from 123.0 in Year 1 to 129.15 in Year 2. *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 20.5

Calculate the consumer price index if there is 3.0% inflation during the year if the price index was previously at 130. *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 20.6

Calculate how much a basket of goods and services currently priced at $1,200 would be if the CPI increased from 125.0 to 135.0. *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 20.7 (HL ONLY)

The data below are for a hypothetical country, Jukeland.

Item	Consumer price index	Statistical weight
Clothing	110	10
Food	120	20
Housing	130	30
Others	140	40

a Define the term *consumer price index (CPI)*. *[2 marks]*

b 'The typical household in Jukeland spends more money on housing than on food or clothing.' Explain this statement. *[2 marks]*

c Use the data above to calculate the weighted consumer price index for Jukeland. *[4 marks]*

Common mistake

Candidates often define the CPI with reference to changes in the average price of a representative basket of *goods*, without acknowledging that *services* are also included in the calculation.

Top tip!

The CPI, as the key measure of inflation for most countries, is also important as a benchmark when central banks set interest rates (see Chapter 24).

The limitations of the CPI in measuring inflation (AO2)

When evaluating the measurement of inflation, it is worth noting there are limitations to using the CPI to measure inflation.

Atypical households

Although the CPI is the most widely used price index for measuring inflation, it takes only an average measure, that is, the CPI only considers the expenditure of the 'average' household, whatever this might actually mean in multicultural and diverse societies in the real world. The average pensioner or full-time university student will have different spending habits from a 'typical' family household. Therefore, the CPI could have minimal if any relevance for atypical households, such as expatriates working in the country.

Regional and international disparities

Prices can vary enormously between countries making international comparisons difficult. For example, housing costs are extremely high in Hong Kong and Singapore but far more affordable in Hungary or Scotland. Similarly, the CPI does not reflect regional differences and disparities in prices and the cost of living. For example, prices tend to rise faster in central business districts than in rural areas of the country. Hence, the CPI does not necessarily measure changes in average price levels (and, hence, the cost of living) for all stakeholders in the economy.

Theory of Knowledge

What role does culture have in determining how inflation affects individuals within a society, and why does this matter?

Different income earners

Different people, on different levels of income, can experience a different rate of inflation because their pattern of expenditure is not necessarily or accurately reflected by the CPI. High-income earners and wealthy households are less sensitive to changes in prices, where low-income individuals and households are far more affected by inflation. Furthermore, different income earners experience their own rates of inflation owing to the differences in their spending habits. Individuals who cannot afford a car or their own home are largely unaffected by higher petrol prices or by house price inflation. Hence, as an average measure, the CPI does not account for different consumption patterns between high-, middle- and low-income earners.

Changes in product quality

As a price index, the CPI ignores changes in the quality of goods and services over time. For example, producers of televisions, computers, cars and smartphones might charge higher prices for products, but the higher specifications and build quality are not reflected in the calculation of the CPI.

Different patterns of consumption over time

Consumption patterns change over time, thereby affecting price levels. Each product included in the CPI is weighted (such as by multiplying the price of the product by the number of units purchased by the typical household during the year). However, as the popularity of most goods and services change over time, these need to be reflected in the CPI calculations but are not easy to determine when making historical comparisons of inflation. Therefore, the CPI can give a misleading or inaccurate representation of inflation over time.

Time lags

Inflation figures and calculations may not accurately reflect changes in consumption patterns due to time lags in collecting data to compile the CPI. Given the huge amount of data collection and reporting needed to construct the CPI, the data may well be out of date or inaccurate by the time of publication.

Volume or value of quantities purchased

Finally, the calculation of the inflation rate by using *quantities purchased* as weights in the CPI (rather than the percentage of income spent on the items) is also a limitation. For example, the typical household might spend $80 on refuelling their family car every two weeks (so, makes only two purchases per month), whereas the same family might buy eight loaves of bread in the same time period. Hence, while applying a weight in the CPI based on the quantities purchased is simple, it is also somewhat meaningless.

Theory of Knowledge

The CPI is compiled using data for the 'average' or typical household in the country, based on their choice of which goods and services are bought on a regular basis. Therefore, to what extent can we rely on the CPI to represent or know about a change in the cost of living for everyone in the economy?

Theory of Knowledge

Given that different countries may differ in the way they calculate the CPI – especially as different types of goods and services and statistical weights are included in the representative basket – what knowledge, if any, can be determined by international comparisons of inflation rates?

Causes of inflation (AO2, AO4)

Economists differ in their view on the root cause of inflation. Keynesians believe that inflation is caused by an increase in aggregate demand in the economy, so AD must be controlled. For example, deflationary fiscal policy (see Chapter 25) can be used to prevent a rise in level of consumption and investment expenditure. By contrast, monetarists argue that there is a need to control the economy's **money supply** in order to control a rise in the general price level. In general, there are two main causes of inflation: **demand-pull inflation** and **cost-push inflation**.

◆ **Money supply** is the amount of money in circulation within the economy at a particular point in time, as determined by the central bank. The money supply comprises legal tender (banknotes and coins), loans, credit, bank deposits and central bank reserves.

Demand-pull inflation is inflation triggered by higher levels of aggregate demand (AD) in the economy, which drives up the general price level (see Figure 20.2). It is caused by excessive aggregate demand (AD increasing at a faster rate than aggregate supply in the economy), thereby forcing up the general price level. Ceteris paribus, an increase in any component of aggregate demand (consumption, investment, government spending and net exports) will cause demand-pull inflation. Diagrammatically, this can be shown by a rightwards shift of the aggregate demand curve. For example, this might be due to higher real GDP per head, lower levels of unemployment, an increase in net export earnings, lower interest rates or cuts in income tax.

◆ **Demand-pull inflation** is inflation caused by higher levels of aggregate demand (AD) in the economy, which drives up the general price level.

Figure 20.2 shows that during an economic boom, the consumption of goods and services increases due to higher levels of employment and higher real GDP per capita. This is represented by a rightwards shift of the aggregate demand curve from AD_1 to AD_2, raising national income from Y_1 to Y_2 and increasing the general price level from PL_1 to PL_2 (demand-pull inflation). Hence, this causes demand-pull inflation, as the general price level has increased.

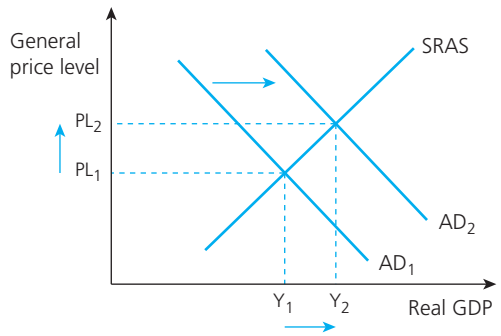


Figure 20.2 Demand-pull inflation

Top tip!

An increase in aggregate demand has a minimal impact on inflation if there is spare capacity in the economy, that is, if aggregate supply is relatively price elastic.

◆ **Cost-push inflation** refers to inflation caused by higher costs of production, such as higher labour costs or skyrocketing rents. It is shown by a leftwards shift of the aggregate supply (AS) curve, thereby forcing up average prices.

Cost-push inflation refers to inflation caused by higher costs of production, thus shifting the aggregate supply (AS) curve to the left and forcing up general prices (see Figure 20.3). This might be caused by higher raw material prices or soaring rents. The rising production costs force firms to raise their prices in order to maintain their profit margins. Diagrammatically, this is shown by a leftwards shift of the short run aggregate supply (SRAS) curve.

Causes of cost-push inflation include higher imported prices of raw materials, components (semi-finished goods) and finished goods for sale, higher labour costs in the economy, increased corporation taxes and escalating rents for commercial properties.

In Figure 20.3, higher costs of production shift the aggregate supply curve from $SRAS_1$ to $SRAS_2$. This forces up the general price level from PL_1 to PL_2 (cost-push inflation) but reduces real national output from Y_1 to Y_2.

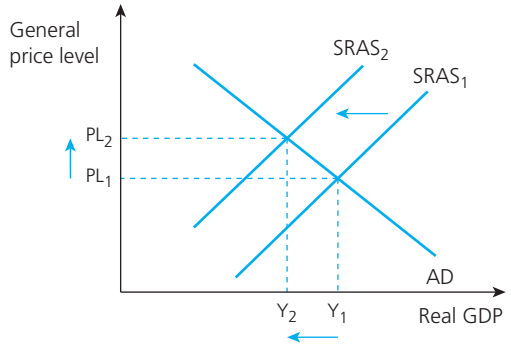


Figure 20.3 Cost-push inflation

PAPER 1 EXAM PRACTICE QUESTION 20.8

Explain the differences between cost-push and demand-pull inflation. *[10 marks]*

Costs of a high inflation rate (AO2)

Inflation diminishes the ability of money to function as a medium of exchange. This means the real purchasing power of households, firms and governments declines when there is inflation. This is because there is a fall in real national income as money is worth less than before. Governments aim to influence the inflation rate because inflation (or the lack of price stability) can cause negative consequences on different stakeholder groups. In general, inflation is also undesirable owing to the costs on the economy and society. These interrelated costs include: uncertainty, redistributive effects, the effects on saving, damage to export competitiveness, the impact on economic growth and inefficient resource allocation.

Uncertainty

Inflation reduces the real value of money, so lowers the spending power of households and firms in the economy. This creates uncertainty, especially when the inflation rate is very high, thereby reducing both consumer and business confidence levels. This means households and firms are more uncertain and confused about what to spend their money on as there are higher risks associated with such expenditure. For example, firms are usually less willing to invest because of the uncertainty about future costs, prices and profits. The uncertainty and confusion brought about by inflation can lead to lower economic growth over the long-term. By contrast, countries with low and stable inflation rates tend to have better economic prospects than countries experiencing higher rates of inflation.

Redistributive effects

The effects of inflation are not distributed evenly among different stakeholders in the economy, so a consequence of this is that there are greater costs to some people who suffer more from the fall in the purchasing power of their money. High inflation in the economy has a regressive effect on low-income households, including the elderly population in the country. This happens in particular when the prices of essential goods and services, such as food products and utilities (such as water, gas and electricity), rise at a rapid rate. Although all consumer groups tend to lose from inflation, low-income households and small-sized firms are most vulnerable as the cost of living and costs of production increase.

Those on fixed incomes (such as workers who earn a fixed wage rate or salaries) and those who receive fixed welfare payments (such as those on state pensions or unemployment benefits) also suffer from high inflation. This is because the real purchasing power of their fixed income declines with higher price levels in the economy. The payments received by fixed income earners are not necessarily adjusted in line with inflation, despite higher prices in the economy.

By contrast, land and property owners can benefit from higher prices of their physical assets. However, higher property and asset prices also make it even more challenging for those without their own home or land to do so.

Common mistake

Beware of stating that demand-pull and cost-push are the two *types* of inflation. This is incorrect as these are the *causes,* not the types of inflation.

Figure 20.4 Property owners tend to gain from higher prices

WISE ChoICES

Inflation harms the poorest members of society far more than those with higher incomes. Low-income earners tend to be far more sensitive to price changes for most goods and services. By contrast, those on high incomes have accumulated wealth so are less affected by higher prices in the economy. To what extent should there be greater **intervention** to improve greater **equity** in income distribution?

The effects on saving

Savers, be they individuals, firms or governments, will lose out from high inflation, assuming there is no change in interest rates for savings. This is because inflation reduces the real rate of return for savers. For example, if banks pay 1 per cent interest on people's savings accounts but the inflation rate is 2.5 per cent, then savers are losing out as the real interest rate is –1.5 per cent, that is, this is a negative real interest rate. Hence, inflation discourages savings as money becomes less effective as a store of value. This is particularly harmful to the elderly, who often have to rely on the interest on their savings as a source of income.

By contrast, borrowers can gain from inflation as the money they need to repay is worth less than when they initially borrowed it. For example, a mortgage at 5 per cent interest with inflation at 3.5 per cent means the real interest rate is only 1.5 per cent. Hence, the real value of the debt declines for the borrower.

However, this means that creditors (lenders) will lose from high inflation. This is because the money lent to borrowers becomes worth less than before as a result of inflation. Again, the real value of the money returned to lenders in the future is lower because of inflation in the economy.

Top tip!

High inflation has many opportunity costs on societies. For example, high inflation puts pressure on the government to increase the value of state pensions, the national minimum wage, unemployment benefits and other welfare payments because the cost of living climbs higher due to higher prices in the economy.

Damage to export competitiveness

If a country has persistently higher inflation rates than its trading partners, then its exports will become less price competitive, leading to a fall in exports and a deterioration in the country's current account on the balance of payments (see Chapter 32). Exports become more expensive, which generally lowers the international competitiveness of domestic firms. In turn, this reduces the sales and profitability of domestic businesses, which can slow or even reduce economic growth and cause higher unemployment in the economy. Hence, exporters lose from high and sustained inflation in the domestic economy.

Imports become relatively cheaper for domestic consumers and firms, so they may choose to substitute domestic products for imported ones. In addition, inflation can cause problems for countries that are heavily reliant on essential imports such as petroleum, steel, motor vehicles, electrical equipment and pharmaceuticals, for example, due to the fall in the purchasing power of money in the domestic economy. Hence, these importers can gain from the domestic economy facing inflationary pressures – especially if the demand for imports is highly price inelastic. Therefore, this will benefit foreign firms if domestic consumers spend more on imports.

Overall, as high inflation tends to damage a country's export competitiveness, this will also reduce the economy's international competitiveness in the long-term. This can cause the balance of payments to go into deficit (when the country imports more goods, services, and capital than they export), thereby causing lower economic activity and higher unemployment if the problem of high inflation is not dealt with.

The impact on economic growth

The combination of uncertainty and the lower expected real rates of return on capital investments (due to higher costs of production) tends to lower the amount of planned investment in the economy. In the long run, a lack of investment expenditure due to lower business confidence levels might damage the country's productive potential and productivity. Hence, high inflation is harmful to the country's economic growth.

Furthermore, workers are more likely to demand a pay rise during times of inflation in order to maintain their level of real income and spending power. As a result, labour costs rise and profits margins decline, ceteris paribus. In addition, as consumption is also likely to fall due to higher price levels, employers tend to lose out from inflation. This also has negative impacts on the country's level of economic growth.

Inefficient resource allocation

High inflation can cause inefficiencies in the economy as higher prices can distort resource allocation. For example, in the labour market, trade unions might negotiate higher wages for their members, although this causes higher costs of firms. However, without receiving higher wages to compensate for the higher cost of living, workers may become disgruntled and demotivated, thereby causing labour unrest and conflict within the economy. Therefore, the resulting **wage-price spiral** simply causes more problems for the economy such as higher costs of production, lower profitability and more unemployment fuelled by allocative inefficiencies.

◆ A **wage-price spiral** is a circular cause of higher inflation due to trade unions negotiating higher wages owing to higher costs of living, which causes higher costs of firms.

In consumer markets, individuals and households incur **menu costs** because inflation impacts on the prices charged by firms. Menu costs require firms to continually update their prices (even for online businesses) and to communicate these to their customers. Catalogues, price lists and menus have to be changed regularly, which is costly and inefficient for businesses.

◆ **Menu costs** are a cost of inflation because with prices continually rising, firms incur ongoing costs of changing and updating price lists and price tags.

Figure 20.5 Updating prices can be a time-consuming and costly task for firms

High inflation can also cause **shoe leather costs**, meaning that customers spend more time searching for the best deals, be it physically or online. They might also have to make more regular visits to the bank for cash withdrawals or devote more time for online banking. Shoe leather costs therefore represent a greater opportunity cost for customers in terms of time and effort needed to counteract the effects of inflation. Again, this causes greater inefficiencies in the economy.

◆ **Shoe leather costs** are a cost of inflation, as customers spend more time searching for the best deals; that is, it represents a greater opportunity cost for customers in terms of time and effort needed.

In summary, the interrelated economic costs of inflation can be remembered by the acronym **REUSER**: **R**edistributive effects, **E**xport competitiveness, **U**ncertainty, **S**avings, **E**conomic growth and **R**esource allocation.

Inflation that is monitored and controlled makes conditions far more predictable for economic stability, that is, there is greater certainty for consumers, producers and the government. By contrast, erratic and uncontrolled rates of inflation can be extremely harmful for consumer and producer confidence levels. Foreign direct investment (FDI) is also likely to fall as a result. For example, the inflation rate in Syria reached a historical high of 121.29 per cent in August 2013, meaning that the general price level in Syria increased by an average of over 121 per cent in a year. Hence, a product priced at 100 Syrian pounds would increase to more than 221 Syrian pounds by the end of the year. However, Syria's lowest rate of inflation was –31.05 per cent in September 2014. The huge fluctuations in its inflation rate makes conditions far less predictable for economic stability in Syria than in other countries with low and stable rates of inflation such as Canada, the UK and the USA (see Table 20.1). Hence, uncontrolled inflation can be detrimental to both economic growth and employment.

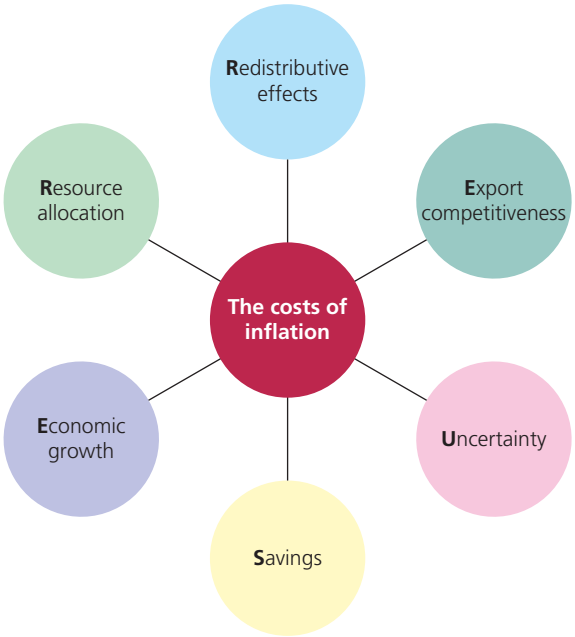


Figure 20.6 The costs of inflation (REUSER)

Table 20.1 Inflation rates around the world, selected countries, March 2020

Venezuela	9,585.50%
Zimbabwe	473.10%
South Sudan	69.00%
Sudan	64.28%
Argentina	51.50%
Liberia	30.55%
Iran	25.00%
Turkey	12.37%
India	7.59%
China	5.40%
USA	2.50%
Canada	2.40%
United Kingdom	1.80%
France	1.40%
Saudi Arabia	0.40%

Source: Trading Economics **https://tradingeconomics.com/country-list/inflation-rate**

Causes of deflation (AO2, AO4)

Deflation is the persistent fall in the general price level in an economy over time, that is, the inflation rate is negative. It is caused by a continual decline in aggregate demand (AD) and/or an increase in the short run aggregate supply (SRAS).

While it might seem that deflation is beneficial (because the general price level of goods and services has fallen, thereby increasing people's purchasing power), it depends on the cause of deflation. An adverse change in aggregate demand, causing a recession, is not beneficial to the economy even if average prices are falling. Ultimately, deflation can be significantly worse for the economy than inflation.

The consequences of deflation depend on the cause. **Benign deflation** is generally positive as the economy is able to produce more (an outwards shift of the SRAS curve), thus boosting national output and employment, without an increase in the general price level (see Figure 20.8).

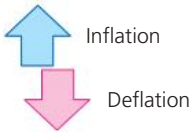


Figure 20.7 Deflation is a fall in the general price level

◆ **Deflation** refers to the persistent fall in the average price level in an economy over time, that is, lower prices in general.

Table 20.2 Deflation rates around the world, selected countries, March 2020

Libya	–6.0%
Mali	–3.3%
Fiji	–3.0%
Chad	–2.8%
Niger	–2.3%
UAE	–1.4%
Burkina Faso	–0.9%
Armenia	–0.5%
Ecuador	–0.3%
Switzerland	–0.1%

Source: Trading Economics https://tradingeconomics.com/country-list/inflation-rate

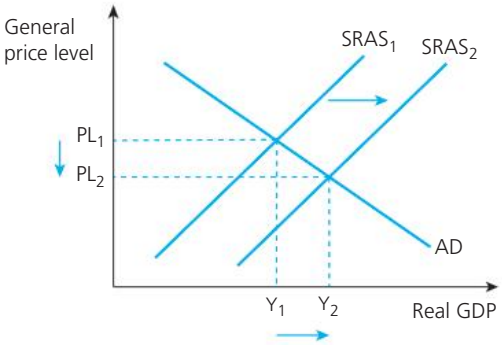


Figure 20.8 Deflation caused by supply factors

◆ **Benign deflation** is generally unharmful deflation as the economy is able to produce more without an increase in the general price level (shown by an outwards shift of the SRAS curve).

Top tip!

Although not explicitly stated in the IB Economics syllabus, benign deflation can come about by an outwards shift of the economy's LRAS (see Figure 20.9) or its PPC. For example, policies to increase the quantity and quality of factor resources, such as investments in human resource and 5G telecommunications technologies, can boost the economy's production capacity in the long-term, thereby increasing the country's national output without prices having to increase.

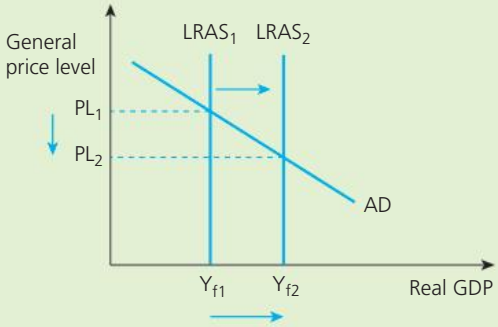


Figure 20.9 Deflation caused by a shift in the LRAS curve

- Deflation can also be caused by a rightwards shift of the LRAS curve from $LRAS_1$ to $LRAS_2$.
- This increases real national output from Y_{f1} to Y_{f2}, and forces down the general price level from PL_1 to PL_2.
- This causes benign deflation (deflation that is beneficial to the economy).

Deflation can be caused by an increase in the economy's short run aggregate supply, such as lower production costs, higher productivity or efficiency gains from technological progress. The higher level of aggregate supply drives down the general price level of goods and services from PL_1 to PL_2, as shown in Figure 20.9, while increasing real national income from Y_1 to Y_2. Hence, benign deflation (non-threatening deflation) is not necessarily disadvantageous to an economy if lower prices enable households to have cheaper access to a greater number and variety of goods and services, and for exporters to be more competitive on price.

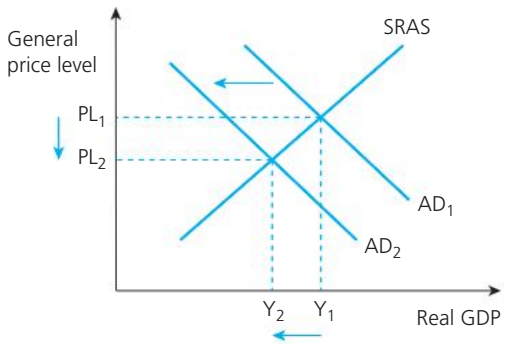


Figure 20.10 Deflation caused by demand factors

However, economists are concerned with **malign deflation**, which is generally harmful to the economy due to a decline in aggregate demand for goods and services in the economy. This is usually associated with an economic recession and rising levels of unemployment (see Figure 20.10).

In addition to a rightwards shift of the SRAS curve (benign deflation), deflation can also be caused by a leftwards shift of the aggregate demand curve from AD_1 to AD_2. This reduces real national income from Y_1 to Y_2, and forces down the general price level from PL_1 to PL_2. This causes malign deflation (deflation that is harmful to the economy due to the fall in real GDP).

◆ **Malign deflation** is generally harmful deflation due to a decline in aggregate demand for goods and services in the economy (shown by an inwards shift of the AD curve, and, hence, a fall in real GDP).

Common mistake

Some students argue that inflation is simply bad for the economy. This is not necessarily true. Low rates of inflation, of 1 or 2 per cent, are not usually harmful to the economy because higher prices can encourage firms to supply more output. It is when inflation rises too quickly that it can disrupt decision-making for households, firms and governments. In fact, deflation is far more of an economic problem than inflation tends to be.

Top tip!

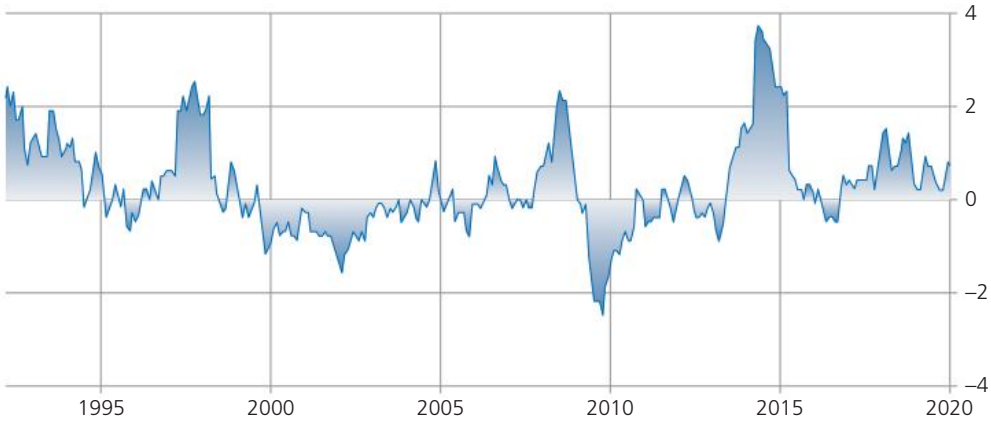
The extent to which an economy is affected by malign deflation will depend on the severity of deflation. Portugal's experience of zero rate inflation in 2013 would have been very different from Somalia's –15.35 per cent inflation rate in the same year.

PAPER 1 EXAM PRACTICE QUESTION 20.9

Explain the difference between a deflationary gap and deflation. *[10 marks]*

PAPER 2 EXAM PRACTICE QUESTION 20.10

For much of the past 30 years, Japan has suffered from deflation (see Figure 20.11).



Source: Trading Economics **www.tradingeconomics.com/japan/inflation-cpi**

Figure 20.11 Japan's inflation rate (%), 1990–2020

a Define the term *deflation*. *[2 marks]*

b Explain what evidence there is in the chart to suggest that Japan has suffered deflation for most of the past 30 years. *[2 marks]*

c Explain the impacts of prolonged deflation for the Japanese economy. *[4 marks]*

Disinflation and deflation (AO2)

Disinflation occurs when there is a fall in the rate of inflation, that is, prices are still rising, but at a slower pace. This is different from an actual fall in the general price level, which as stated above is called deflation. This occurs when the inflation rate is negative. Nevertheless, disinflation can lead to deflation if not controlled, with negative consequences for the economy and a drop in the standards of living in the country.

◆ **Disinflation** occurs when there is a fall in the rate of inflation, that is, prices are still rising, but at a slower pace.

Diagrammatically, deflation results in *lower* average prices (see Figure 20.12), whereas disinflation is shown by a smaller proportional *increase* in average prices (see Figure 20.13).

The fall in aggregate demand from AD_1 to AD_2 causes real national output to drop from Y_1 to Y_2 with the general price level falling from PL_1 to PL_2.

In Figure 20.13, an increase in aggregate demand from AD_1 to AD_2 causes national output to rise from Y_1 to Y_2 but with prices rising from PL_1 to PL_2, that is, there is demand-pull inflation. Disinflation occurs when the rate of increase in aggregate demand slows down (from AD_2 to AD_3, for example), which reduces the rate of increase in the general price level from PL_2 to PL_3. In other words, prices in general still increase but at a slower rate.

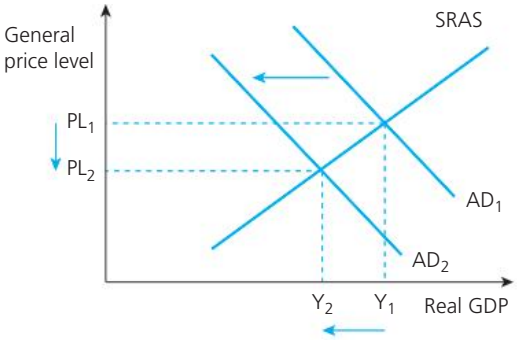


Figure 20.12 Deflation

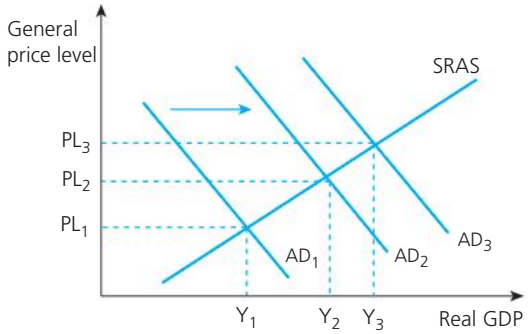


Figure 20.13 Disinflation

PAPER 2 AND 3 EXAM PRACTICE QUESTION 20.11

If the consumer price index (CPI) changes from 151 to 147, explain whether there has been disinflation or deflation. *[2 marks]*

PAPER 1 EXAM PRACTICE QUESTION 20.12

Explain why governments aim for price stability (low and sustainable rates of inflation) as a macroeconomic objective. *[10 marks]*

Top tip!

Make sure you can distinguish between *disinflation* and *deflation*. A fall in the rate of inflation (disinflation) means that prices are still rising on average, but at a slower rate. Be clear about the meaning of deflation – an actual fall in the general price level.

Costs of deflation (AO2)

There are seven specific interrelated costs of malign deflation mentioned in the IB Economics syllabus: uncertainty, redistributive effects, deferred consumption, association with high levels of cyclical unemployment and bankruptcies, increase in the real value of debt, inefficient resource allocation and policy ineffectiveness. These costs can be remembered by the acronym **R U CUPID:**

Uncertainty

As with inflation, deflation also causes uncertainty in the economy. With deflation, there is a subsequent increase in the real value of debts, so this reduces both consumer and business confidence levels. Such uncertainties add further economic problems in the country. For instance, consumers may limit their consumption expenditure owing to uncertainties in the economy. In addition, firms have less of an incentive to invest (capital expenditure) because during periods of deflation they receive lower prices and, hence, profitability. If the uncertainty continues, this can have a detrimental impact on employment and economic growth.

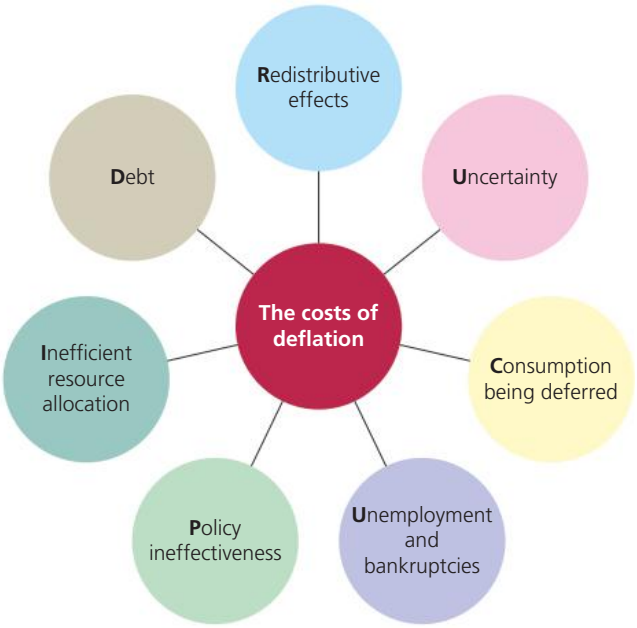


Figure 20.14 The costs of malign deflation

Figure 20.15 Negative equity exists when the value of mortgage debt exceeds the value of the property

Redistributive effects

Deflation causes a fall in the value of assets and household wealth. Owing to declining profitability, share prices fall during times of deflation. This means that dividends and the capital returns on holding shares and other fixed assets also fall, thus reducing the wealth of shareholders and investors.

Deflation also leads to a redistribution of income and wealth from debtors (borrowers) to creditors (lenders). During times of deflation, the real value of debts (such as household car loans or mortgages) will increase. For example, many homeowners suffered from *negative equity* during the financial crisis of 2008 as the value of their homes fell to a level below the value of their debt (the market price of their homes was worth less than the value of their outstanding mortgages). Furthermore, owing to the uncertainties and instability caused by deflation, many debtors may default on their loans, that is, fail to repay their escalating debts.

Deferred consumption

Deflation can cause consumption to be postponed as consumers wait in anticipation of further falls in the prices of goods and services. In particular, it can discourage consumers from purchasing luxury goods and non-essential products, because consumers anticipate that they will save money by waiting for these items to be cheaper in the future. Hence, deferred consumption slows down the level of economic activity. Therefore, deflation can cause a fall in business profits and investments, which can easily lead to a prolonged economic recession and unemployment.

Therefore, prolonged periods of deflation often lead to lower consumer spending, postponed investment expenditure and lower economic growth. Subsequently, this creates even more deflationary pressures on the economy due to the **deflationary spiral** (see Figure 20.16).

ATL ACTIVITY 20.2

Many countries have experienced negative equity, including the USA, the UK and Hong Kong. Investigate the impact of negative equity on homeowners in one of these economies or any other of your choice. What were the wider social and economic impacts of negative equity in the economy?

◆ A **deflationary spiral** is a circular process of downward pressure on the general price level caused by a weak economy, which leads to lower wages and confidence levels that in turn causes prices to fall further.

Association with high levels of cyclical unemployment and bankruptcies

As malign deflation occurs because of a fall in aggregate demand in the economy, this causes a fall in the derived demand for labour, which means, deflation can cause cyclical unemployment, hence, mass job losses throughout the economy.

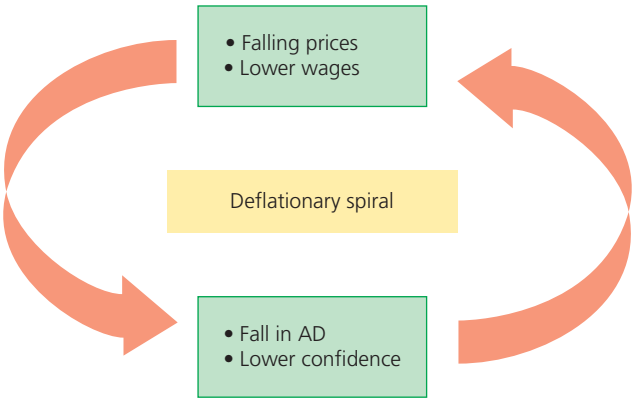


Figure 20.16 The deflationary spiral

It is rational and reasonable for firms that face lower prices for their goods and services to cut wages too, in order to counteract the reduction in their revenues and profit margins. Employers could also argue that during times of deflation, real wages rise. However, wages might not adjust instantaneously especially if workers are on fixed-term employment contracts or owing to resistance from trade unions to accept pay cuts for their members. John Maynard Keynes referred to this as 'sticky wages', by which he meant the reluctance of people to accept pay cuts even though this could have major macroeconomic benefits for everyone in the society. This, in turn, will lead to further job losses, because labour is a major factor input in the production process.

As cyclical unemployment results in consumers paying less during periods of deflation, firms suffer from lower sales revenues and profits. This makes it more difficult for businesses to pay their costs and liabilities (debts such as mortgages), thus causing a large number of bankruptcies if there is prolonged deflation in the economy.

Increase in the real value of debt

The real cost of debts (borrowing) increases when there is deflation because real interest rates rise when the price level falls. For example, if interest rates average 1.0 per cent but the inflation rate is –1.5 per cent, then the real interest rate is approximately 2.5 per cent. This means that deflation makes it more difficult for debtors (borrowers) to pay off their debts because households and firms have to spend a larger percentage of their disposable income on meeting debt repayments. Note that during periods of deflation, it is likely that firms will also be getting lower revenue, and consumers will likely be getting lower wages. Essentially, the increase in the real value of debt reduces the spending power of consumers and firms in the economy.

With more bankruptcies, unemployment and lower levels of economic activity, tax revenues are likely to fall while the amount of government spending rises (due to the economic decline associated with malign inflation). Hence, deflation can also create a budget deficit for the government, thereby further increasing the real value of public sector debt. Therefore, persistent deflation reduces both consumption and investment expenditure, with detrimental impacts on real national output and employment in the economy.

Inefficient resource allocation

Deflation is usually caused by a weak economy, that is, a significant fall in the level of aggregate demand and real GDP that leads producers to cut prices on an ongoing basis in order to incentivize consumers to buy their products. The damaging effects of deflation also cause market distortions, such as greater uncertainty, deferred consumption and investment, bankruptcies, cyclical unemployment and an increase in the real value of debt; all of which lead to inefficient resource allocation in the economy.

Policy ineffectiveness

Policy measures adopted by the government to achieve macroeconomic goals such as economic growth and low unemployment become less effective when the economy faces deflation. In particular, the use of loose monetary policies (see Chapter 24) by the Central Bank or monetary authority becomes ineffective as interest rates are already very low during times of deflation. Given the weak economy, governments would ordinarily lower interest rates to boost aggregate demand, but nominal interest rates cannot fall below zero so monetary policy becomes ineffective in a situation with deflation and an economic recession. As it becomes more challenging to alleviate deflation, policy ineffectiveness can also lead to lower economic growth and higher unemployment in the economy.

Common mistake

'Deflation makes people better off because things are cheaper in the economy.'

Deflation is usually associated with a weakening economy, so incomes are likely to be falling or at least stagnant. It is also associated with higher unemployment due to the lack of aggregate demand in the economy. So, although prices are generally falling, so is the amount of household income and business revenue, so consumption and investment expenditure also fall. Debt is also likely to rise as a result of this. Deflation is only 'good' if prices are falling while personal disposable income and business profits are rising (as in the case of benign deflation). Therefore, the wider macroeconomic consequences of recession, indebtedness and unemployment mean that the economy is not likely to be better off in times of deflation.

Top tip!

It can be difficult for an economy to break out of a downward deflationary spiral (Japan suffered deflation for a 20-year period from 1991 to 2010, a period that became known as *The Lost Score* or *Lost 20 Years*). Business and consumer confidence levels need to increase significantly in order to boost aggregate demand. It might be possible, for example, to cut interest rates to encourage consumer spending and increase investment in the economy.

Relative costs of unemployment versus inflation (AO3)

Both low unemployment and low inflation are key macroeconomic objectives. This is because policymakers know that both high unemployment and high inflation have major costs on the macroeconomy. For example, there are personal, social and economic costs of high unemployment (see Chapter 19). The costs of high inflation can be remembered by the **REUSER** acronym: Redistributive effects, Export competitiveness, Uncertainty, Savings, Economic growth and Resource allocation (see page 300). Examples of the costs of inflation and unemployment are outlined in Table 20.3.

Table 20.3 The costs of unemployment and inflation

Costs of unemployment	Costs of inflation
• Psychological costs • Social problems • Loss of household incomes (earnings) • Higher government borrowing • Slower economic growth • Rising income and wealth inequalities	• Increases the cost of living • Deferred consumption and investment • Reduces international competitiveness • Erodes savings • Uncertainty • Menu costs and shoe leather costs

Although these costs are not mutually exclusive, as unemployment and inflation are interlinked, governments often struggle to decide which macroeconomic objective to pursue as a priority. Priorities may change over time, and context is also important to consider. For example, during the global financial crisis of 2008 which caused many countries to go into a deep recession, countries such as Portugal, Italy, Greece and Spain pre-initialized lowering unemployment even if this would lead to demand-pull inflation in their economies. By contrast, countries such as Venezuela and Zimbabwe (see Table 20.1) will want to prioritize lowering their very high rates of inflation in order to alleviate the major macroeconomic problems caused by unsustainable rises in prices and the cost of living.

ATL ACTIVITY 20.3

In August 2019, Venezuela's inflation surpassed 10 million per cent. Investigate the causes and impacts of Venezuela's hyperinflation for four different stakeholder groups of your choice. Evaluate the degree of success of the government in tackling the problems of hyperinflation in the country.

CASE STUDY

The worst case of runaway inflation?

Hungary faced runaway inflation (hyperinflation) in 1946. After the Second World War, Hungary witnessed its prices doubling every 15 hours! The country introduced the 100 quintillion pengo bank note (100 quintillion = 100,000,000,000,000,000,000!). This is the largest ever denomination banknote to have become legal tender. At the time, 100 quintillion pengo was equal to around $0.20. Today, Hungary's highest denomination banknote is a modest 20,000 forint (the currency that replaced the pengo). As a member of the European Union (EU), the long-term goal of the government is to replace the forint with the euro. For Hungary to achieve this aim, it will need to ensure there is price stability that is sustained over the long-term.

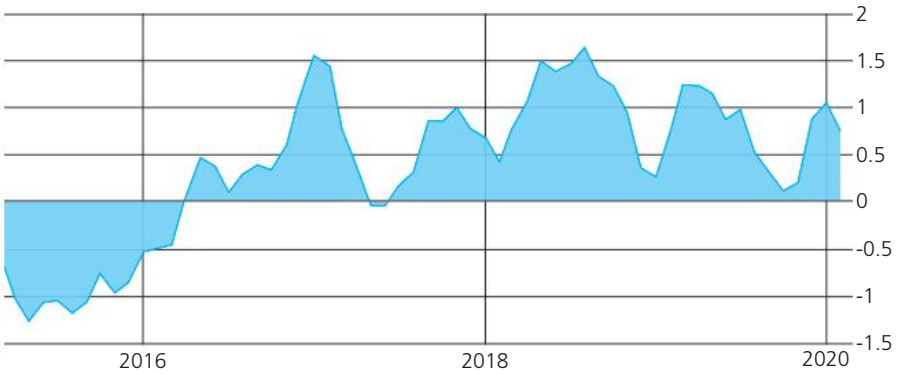
Figure 20.17 The 1 pengo coin

American economist Arthur Melvin Okun (1928–80) developed the 'misery index' – the sum of the unemployment rate and the inflation rate for a country. Essentially, the higher the values of unemployment and inflation, the greater the misery index is for a country, and the lower its economic well-being tends to be. It is not a highly accurate measure of economic well-being but does attempt to quantify the extent to which inflation and unemployment directly contribute to this. For example, in a country where jobs are scarce, if the cost of living increases rapidly this will cause major social and economic problems. The Cato Institute, an American think tank organization, publishes an annual misery index report for 95 countries, based on their inflation and unemployment data (see Table 20.4). Unsurprisingly, Venezuela's long-term hyperinflation and high unemployment problems since 2014 mean that the country has the highest misery score. By contrast, Thailand is the least miserable country owing to its remarkable track record of low unemployment (see Table 20.4).

Table 20.4 The misery index, selected countries (2019)

Rank	Country
1	Venezuela
2	Argentina
3	Iran
4	Brazil
5	Turkey
…	
91	China
92	Austria
93	Japan
94	Hungary
95	Thailand

Source: https://www.visualcapitalist.com/the-most-miserable-countries-in-the-world/



Source: Trading Economics **www.tradingeconomics.com/thailand/inflation-cpi**

Figure 20.18 Thailand's relatively stable inflation rates (2015–20)

In deciding which macroeconomic objective to pursue as a priority, policymakers consider the breadth and severity of the relative costs of unemployment and inflation (see Table 20.3). They also need to consider the magnitude of inflation and unemployment rates in the economy; the former will take priority if, for example, the economy is experiencing hyperinflation.

Chapter summary

- Inflation is the persistent rise in the general price level in an economy, over time.
- Inflation, deflation and disinflation are typically measured by using a consumer price index (CPI). This is a statistically weighted index measuring the change in prices of a representative basket of goods and services consumed by the average household.
- Governments set a target inflation rate as a key macroeconomic objective owing to the potentially harmful impacts of inflation on the economy.
- Inflation results in higher prices of goods and services, whereas deflation causes prices of goods and services to fall, on average.
- Statistical weights are assigned to different goods and services to reflect the relative importance of household spending on these items.
- The loss of purchasing power owing to inflation can cause greater levels of uncertainty in the economy, with detrimental impacts on economic growth and employment in the economy.
- The international competitiveness of a country tends to decline when there is domestic inflation as exports become less price competitive. This causes a decline in the profits of domestic firms, leading to a fall in export earnings for the country.
- As inflation dampens business confidence in the economy, firms are less likely to spend money on capital investments during this time.
- Deflation is the persistent fall in the general price level in an economy over time, that is, the inflation rate is negative. It can be caused by a continual decline in aggregate demand and/or an increase in aggregate supply.
- Disinflation occurs when there is a fall in the rate of inflation, that is, prices are still rising, but at a slower rate. Disinflation can lead to deflation if not controlled.
- Benign deflation is not necessarily a concern, as it can be caused by an increase in the economy's productive capacity. Malign deflation is harmful to the economy, as it is caused by lower aggregate demand in the economy, causing a decline in real GDP.
- Economists and policymakers are challenged by prioritizing between unemployment and inflation; while both are harmful to the economy, it is not always clear which one causes more damage to the economy.

Theory of Knowledge

It has been said that unemployment is a price worth paying. On what grounds can this statement be justified? Discuss whether you agree with this statement in the context of a country of your choice.

21 Macroeconomic objectives – sustainable level of government (national) debt (HL only)

Real-world issue

Why does economic activity vary over time and why does this matter?

Conceptual understandings (WISE ChoICES)

- **Change** in the conditions of the demand and supply sides of the economy cause economic activity to vary over time.
- Fluctuations in economic activity impact the **economic well-being** of individuals and societies.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- the macroeconomic objectives of a sustainable level of government debt (national debt) (AO2)
- sustainability of national debt in terms of measuring government debt as a percentage of GDP (AO2)
- the relationship between a budget deficit and government debt (AO2)
- the costs of a high government debt (AO2).

Sustainable level of government (national) debt (HL only) (AO2)

A **budget deficit** exists when the value of government spending exceeds its revenue per time period. For example, during the coronavirus pandemic, governments across the world increased their level of spending in order to support households and businesses owing to the major economic slowdown. So, if government spending exceeds tax revenues (G > T), the public sector runs a budget deficit. **Government debt** (also known as **national debt**) is the sum of all accumulated government budget deficits from previous years. It represents the total amount of money owed by the government to its domestic and foreign creditors, such as commercial banks, the IMF and the World Bank. A sustainable level of government debt is important for effective operations of the macroeconomy over the long-term.

While it is acceptable and perhaps affordable for governments to borrow money to finance their spending (and therefore to incur budget deficits), it is not sustainable. Just as individuals cannot continually spend more than they earn, the same applies to the government. However, a sustainable level of debt means debt is affordable – individuals may borrow money (mortgage) to finance the purchase of their homes as long as their earnings over the long-term mean they can repay the mortgage and interest charges incurred on the debt. In the same way, the extent to which national debt is affordable is determined by measuring the debt in terms of the country's gross domestic product (GDP).

◆ A **budget deficit** exists when the value of government spending exceeds government revenue per time period, that is, G > T.

◆ **Government debt** (or **national debt**) refers to the sum of all accumulated government budget deficits from previous years, that is, it is the total amount of money owed by the government to its domestic and foreign creditors.

PAPER 1 EXAM PRACTICE QUESTION 21.1

Explain the difference between a *budget deficit* and *government debt*. *[10 marks]*

Measurement of government debt (AO2)

The main method used to measure and report government debt is to express the national debt as a percentage of GDP, rather than to convey this in absolute terms. This shows the percentage of annual national income that the government owes to its creditors. The **debt to GDP ratio** shows a country's national debt as a percentage of its GDP (national income). The higher the debt to GDP ratio, the more unaffordable the debt.

◆ The **debt to GDP ratio** shows a country's national debt as a percentage of its GDP (national income).

Table 21.1 Government debt, selected countries, March 2020

Rank	Country	Government debt ($)
1	USA	23,409,959,000,000
2	Italy	2,652,196,000,000
3	Jamaica	1,998,668,000,000
4	Spain	1,329,687,275,000
5	Brazil	1,109,207,000,000
6	Sri Lanka	69,596,000,000
7	Ukraine	30,421,426,000
8	Serbia	26,642,000,000
9	Belarus	18,844,403,000
10	Kazakhstan	18,079,072,000

Source: Trading Economics **https://tradingeconomics.com/country-list/government-debt**

Note that the data in Table 21.1 shows the nominal or absolute amount of government debt. However, this may have little meaning when trying to make international or historical comparisons. For example, Spain has a larger amount of debt ($1.329 trillion) than Brazil ($1.109 trillion) but this does not really mean much until we look at the proportion of their debt as a proportion of national income.

Table 21.2 Government debt to GDP ratio

Country	Government debt ($)	Nominal GDP ($)	Debt to GDP ratio (%)
Spain	1,329,687,275,000	1,460,000,000,000	91.07
Brazil	1,109,207,000,000	2,020,000,000,000	54.91

In Table 21.2, it can be seen that Spain has a smaller nominal GDP value than Brazil ($1.46 trillion and $2.02 trillion respectively). So, when national debt is expressed as a percentage of nominal GDP, Spain's debt to GDP ratio is over 91 per cent whereas for Brazil it is just below 55 per cent. Hence, Brazil has far more affordable debt than Spain. Table 21.3 shows the countries with the largest debt to GDP ratios.

Table 21.3 Countries with the largest debt to GDP ratios, March 2020

Rank	Country	Debt to GDP ratio (%)
1	Japan	238
2	Greece	181
3	Lebanon	151
4	Italy	135
5	Cape Verde	124
6	Portugal	122
7	Mozambique	113
8	Singapore	112
9	Bhutan	110
10	USA	107

Source: Trading Economics **https://tradingeconomics.com/country-list/government-debt-to-gdp**

Budget deficit and government debt (AO2)

There is a direct relationship between a budget deficit and government (national) debt. This is because each time a government runs a budget deficit, it adds to the national debt. Government debt is the accumulation of all government budget deficits in previous years. It is important that this debt is managed carefully as debt interest means that national debt will simply grow exponentially if repayments are not made.

Recall that a budget deficit exists when government spending (G) exceeds government tax revenue (T) per time period, that is, G > T. This assumes that tax revenue is the only source of government revenue, and excludes other sources such as one-off privatization proceeds. This can be beneficial in the short run as the government spending represents an injection into the circular flow of income, thereby stimulating economic growth and maintaining and/or creating jobs. However, budget deficits are not sustainable in the long run. For example, Table 21.3 shows that Japan has a debt to GDP ratio of 238 per cent. This means that for every $1m of GDP, the nation owes $2.38m. Hence, over the long-term, the government needs to pay off this debt by reducing its spending and/or raising tax revenues. Any period of time when the government can run a budget surplus (T > G), this surplus can be used to pay off some of the national debt. Again, this assumes tax revenues are the only source of government revenue.

Figure 21.1 shows the value of Spain's debt to GDP ratio (%). Owing to Spain's inability to cope with the global financial crisis of 2007–08, national debt soared from 39.5 per cent to more than 100 per cent in 2014 (when the unemployment rate peaked at over 26 per cent). The country has since been able to reduce its debt to GDP ratio, although the economic crisis of the coronavirus outbreak in 2019–20 caused this to rise yet again.

ATL ACTIVITY 21.1

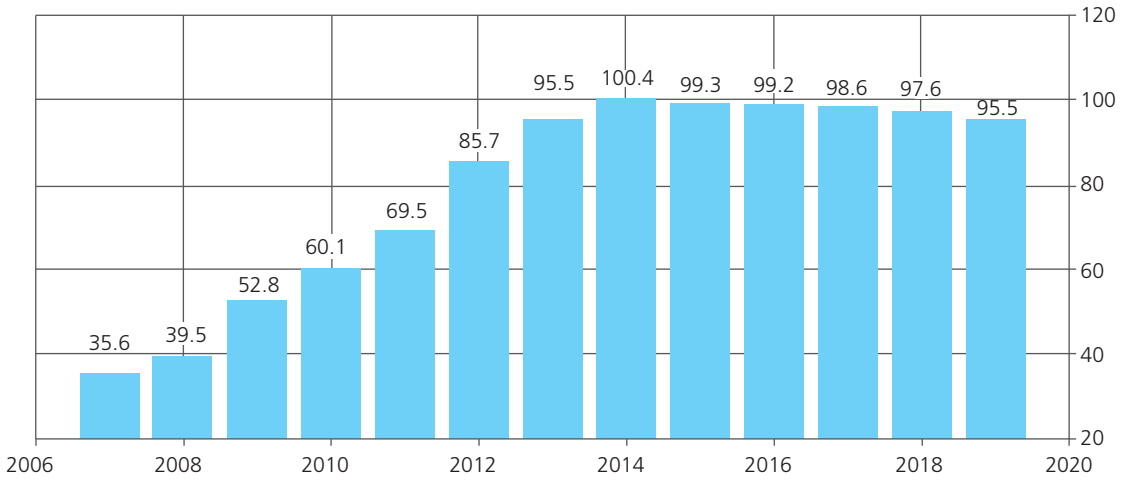
Visit **https://usdebtclock.org/** to see the current (live) value of US debt – you may be surprised to see how quickly the debt is rising. This is not because the government is borrowing more money, but because of the debt interest being accumulated on the national debt. Investigate the national debt for another country of your choice and determine the extent to which you think this debt is affordable debt for the country.

Top tip!

HL students – make sure you can explain that the **multiplier effect** exacerbates the total change in national income that results from a change in taxation or government spending (see Chapter 25).

Theory of Knowledge

What kind of knowledge should policymakers consider when making decisions in favour of pursuing budget deficits or reducing the national debt?



Source: Trading Economics https://tradingeconomics.com/spain/government-debt-to-gdp

Figure 21.1 Spain's debt to GDP ratio (%)

Costs of government debt (AO2)

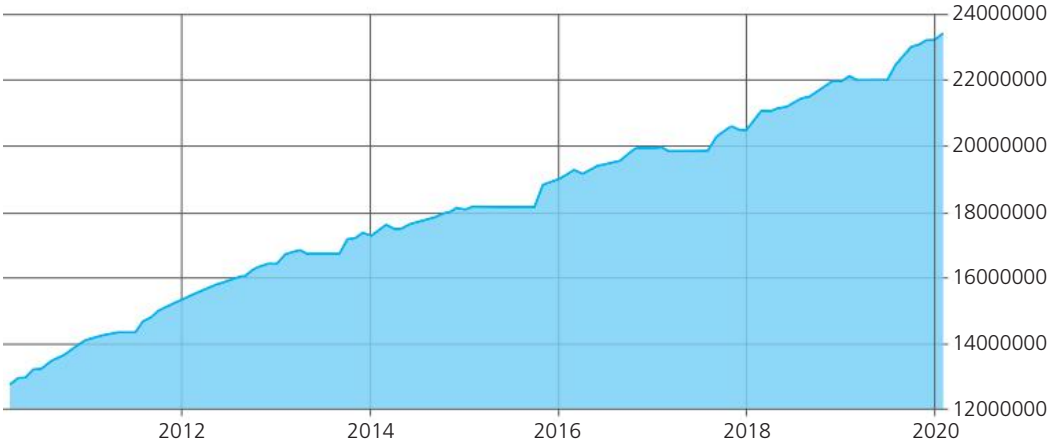
Costs of a high government (national) debt can be classified in three areas: debt servicing costs, credit ratings, and the impacts on future taxation and government spending.

Debt servicing costs

Debt servicing costs refer to the costs of financing the national debt, that is, the loan repayment plus interest charges incurred on the loan. The larger the amount of government debt and the higher the interest rate, the greater the debt servicing costs will be. The value of the US government's debt has continued to soar (see Figure 21.2) largely due to the compound interest that simply adds to the growing government debt.

◆ **Debt servicing costs** are the expenses of financing the national debt, that is, repaying the principal amount borrowed plus interest costs.

In the example of the USA, a national debt of $24 trillion in 2020 even at a low interest rate of 0.5 per cent per year would equate to government debt of around $24.12 trillion by the end of the year. This means debt interest is $120 billion or $328,767 per day! The US bank lending rate for the first quarter of 2020 was 4.75 per cent – at this rate, the annual US debt repayment would be $1.14 trillion or more than $3.1 billion per day! There is clearly a very large opportunity cost to debt servicing.



Source: Trading Economics https://tradingeconomics.com/united-states/government-debt

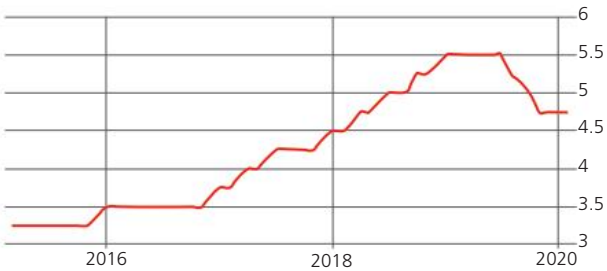
Figure 21.2 US government debt, ($m), 2010–20

Credit ratings

Credit rating refers to the measure of a borrower's ability to repay a loan. A person, organization or country that has a high credit rating is more likely to be able to secure loan capital from a financial institution such as commercial banks because they are estimated to be more able to fulfil their financial commitments. Credit ratings are based on a combination of criteria related to the borrower's credit history and perceived level of risk. This will depend on various factors such as:

◆ **Credit rating** refers to the measure of a borrower's ability to repay a loan. A high credit rating means the borrower has a good credit history and a perceived low level of risk.

- The borrower's previous credit record – those able to repay previous loans will get a higher credit score or rating, whereas those who have defaulted on loans (unable to repay these) will tend to struggle to secure more credit from financial lenders.
- The amount the borrower needs – the higher the amount of the loan, the more risk this represents for the lender.



Source: Trading Economics https://tradingeconomics.com/united-states/bank-lending-rate

Figure 21.3 US bank lending rate (%), 2015–20

Figure 21.4 Countries with large amounts of existing debt will tend to have a poor credit rating

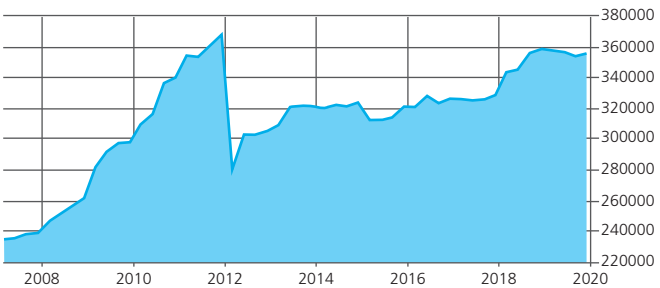
- The amount of existing loans (if any) – individuals, firms, and nations with a large amount of existing debts will tend to find it more difficult to get further loan capital, as this carries a high degree of risk.
- The income of the borrower (or the real GDP for a country) – the higher the national income of the borrowing country, the more likely it is to be able to borrow money to finance its national debts. However, banks and other financiers will look at the country's debt to GDP ratio to determine whether the country can realistically afford to service its debt repayments.

A country struggling to repay its national debt but one that also has a low credit rating will find it increasingly challenging to pay off its liabilities. Hence, having a high credit rating makes the country more attractive to financial lenders such as commercial banks.

The impacts on future taxation and government spending

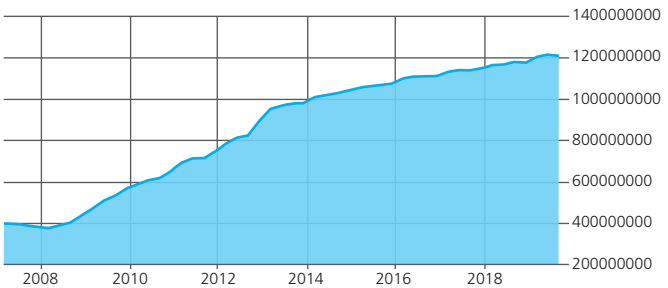
As shown in Figure 21.2 and Figures 21.5 and 21.6, national debt that is not managed and repaid simply continues to increase, making it even more challenging to pay off. The likely impact of this is **austerity measures**, that is, cutbacks in fiscal spending in order to repay national debt caused by budget deficits in the past. Essentially, this entails the combination of increases in taxation and

◆ **Austerity measures** refer to cutbacks in fiscal spending in order to repay national debt. This is achieved by reducing government spending and raising taxation.



Source: Trading Economics https://tradingeconomics.com/greece/government-debt

Figure 21.5 Greece's national debt, (€m), 2007–20



Source: Trading Economics https://tradingeconomics.com/spain/government-debt

Figure 21.6 Spain's national debt, (€m), 2007–20

reduced levels of government spending. Austerity measures in Greece and Spain were inevitable following the huge debts incurred by these countries during the global financial crisis of 2008 (see Figures 21.5 and 21.6).

For countries such as the USA, Greece and Spain, there are opportunity costs to the tax revenues and government spending in the future. Raising taxes to pay off the national debt can cause disincentives in the economy, as people and firms will face higher marginal tax rates (see Chapter 23). Reducing government spending can cause welfare losses in the economy, especially for low-income households. It also puts the government in a more vulnerable position if it has to deal with emergency crises such as the outbreak of infectious diseases or natural disasters such as earthquakes, tsunamis and tornadoes.

ATL ACTIVITY 21.2

Investigate the long-term economic costs and benefits of austerity measures. A good starting point is here: **https://bbc.in/390zG0d**

Common mistake

'In general, budget deficits are harmful to the economy whereas budget surpluses add more money into the economy.'

While a budget deficit causes an increase in the national debt, this may be deemed to be beneficial to the economy, especially during an economic recession. Running a budget surplus is not necessarily a good thing – it could mean that the surplus can be used to repay some of the national debt, but it could also mean taxes were higher than perhaps necessary, which could have hindered incentives to work.

PAPER 1 EXAM PRACTICE QUESTION 21.2

Explain the relationship between a budget deficit and government debt. *[10 marks]*

WISE ChoICES

To what extent does government **intervention** in the macroeconomy to combat the problems of inflation cause more harm than good in terms of long-term **economic well-being**?

Chapter summary

- A sustainable level of government debt is important for effective operations of the macroeconomy over the long-term.
- A budget deficit exists when the value of government spending exceeds government revenue, thereby contributing to the national debt.
- National debt is the accumulated budget deficits of a country and can have positive impacts in the short run (as it represents an injection into the circular flow of income), whereas it incurs costs in the long run.
- The main method of measuring and reporting government debt is to express this as a percentage of GDP. This is known as the *debt to GDP ratio*.
- There is a direct relationship between budget deficits and government debt, that is, each budget deficit adds to the value of national debt.
- The costs of high government debt include: debt servicing costs, low credit ratings, and negative impacts on future taxation and government spending policies.

22 Macroeconomic objectives – potential conflict between macroeconomic objectives

Real-world issue

Why does economic activity vary over time and why does this matter?

Conceptual understandings (WISE ChoICES)

- **Change** in the conditions of the demand and supply sides of the economy cause economic activity to vary over time.
- Fluctuations in economic activity impact the **economic well-being** of individuals and societies.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- the potential conflict between low unemployment and low inflation (AO3)
- an AD–AS diagram to show the trade-off between unemployment and inflation (AO3, AO4) (HL only)
- the trade-off between unemployment and inflation, with reference to the short run and long run Phillips curve (HL only)
- a Phillips curve diagram to show the short run and long run relationship between inflation and unemployment (HL only)
- the potential conflict between high economic growth and low inflation (AO3)
- the potential conflict between high economic growth and environmental sustainability (AO3)
- the potential conflict between high economic growth and equity in income distribution (AO3).

Potential conflict between macroeconomic objectives (AO3, AO4)

It is not always possible for a government to simultaneously achieve all of its macroeconomic objectives. This is because there are a number of potential conflicts between these macroeconomic goals, which leaves the government with several trade-offs or **conflicting macroeconomic objectives**. Ultimately, the government has to decide which macroeconomic aim is the most important to the economy at a particular point in time. This section of the IB Economics syllabus examines four broad areas where there is potential conflict between the main macroeconomic objectives:

- low unemployment and low inflation
- high economic growth and low inflation
- high economic growth and environmental sustainability
- high economic growth and equity in income distribution.

◆ **Conflicting macroeconomic objectives** means that it is not always possible for a government to simultaneously achieve all of its macroeconomic objectives owing to potential trade-offs.

Low unemployment and low inflation (AO3, AO4)

When the economy grows rapidly and more people are in employment, it is likely that inflationary pressures will occur. Low unemployment means more people will have extra money to spend as the economy reaches full employment. However, this is likely to increase aggregate demand and therefore cause demand-pull inflation (see Chapter 20) because aggregate demand in the economy increases faster than aggregate supply. In addition, low unemployment may cause cost-push inflation because full employment makes it harder for firms to attract skilled labour, so this can lead to wage inflation (and, hence, higher prices in general). Therefore, the pursuit of low unemployment (or full employment) can conflict with trying to achieve low inflation.

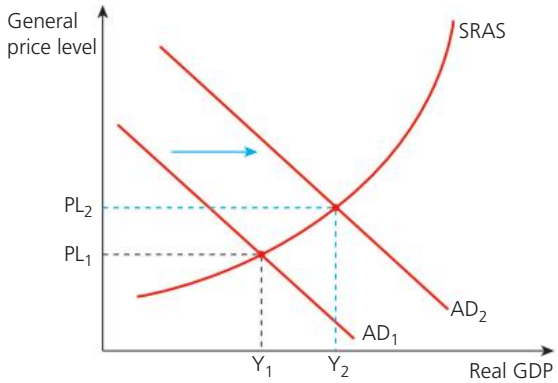


Figure 22.1 The conflict between low unemployment and low inflation

In Figure 22.1, an increase in aggregate demand shifts AD_1 to AD_2, thereby reducing unemployment in the economy. This is because the derived demand for workers increases during an upturn in the business cycle (see Chapter 16), raising national income from Y_1 to Y_2. However, the opportunity cost of pursuing low unemployment is the increase in the general price level from PL_1 to PL_2.

Trade-off between unemployment and inflation (HL only)

The trade-off between unemployment and inflation can be demonstrated by the Phillips curve (see Figure 22.2). The model is named after New Zealand economist William Phillips (1958) who initially used data from the UK from 1861–1913 to show the trade-off between the unemployment rate and the rate of change in nominal wages (which correlates to price stability).

The **short run Phillips curve** (SRPC) shows a potential trade-off between low unemployment and low inflation as macroeconomic objectives. A fall in unemployment, due to an increase in aggregate demand and economic growth, creates more consumption expenditure thereby fuelling higher prices in the economy. This is represented in the diagram by a movement up (to the left) along the SRPC because as the rate of unemployment falls, the rate of wage inflation increases, thus creating a trade-off. As such, the Philips curve is an essential tool used to analyse macroeconomic policy.

◆ The **short run Phillips curve** (SRPC) shows a potential trade-off between pursuing low unemployment and low inflation as macroeconomic objectives.

The short run Phillips curve also shows there is a **natural rate of unemployment** (NRU) – the sum of structural, seasonal, and frictional unemployment in the economy, irrespective of the general price level. The NRU is the unemployment rate that exists at full employment,

◆ The **natural rate of unemployment** (NRU) is the level of unemployment at the full employment equilibrium level of national output, comprising frictional, seasonal and structural unemployment.

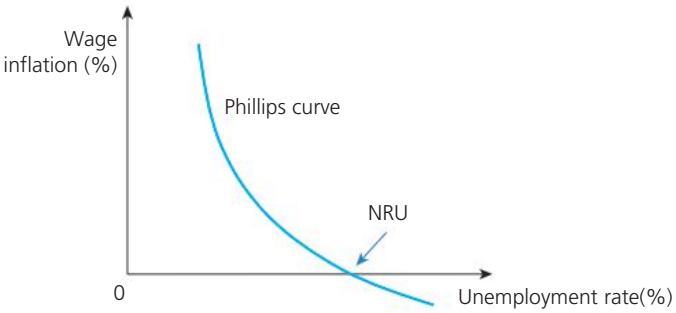
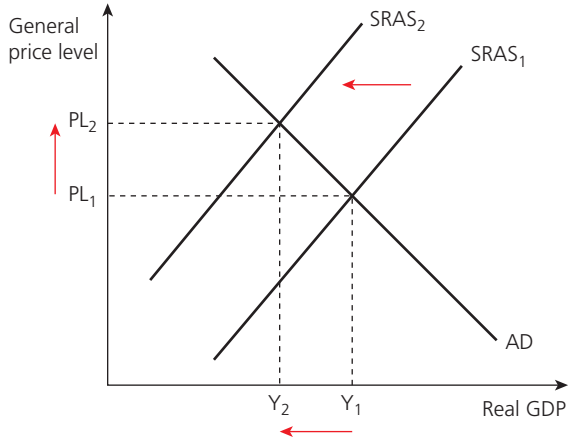


Figure 22.2 The short run Phillips curve (SRPC)

that is, where the demand for labour equals the supply of labour. When unemployment is above its natural rate, deflation occurs because the inflation rate becomes negative. Essentially, the rate of unemployment is so high that there is a recession or depression in the economy, thus forcing prices in general to fall (deflation or negative inflation).

The Phillips curve model suggest that a government might not be able to achieve low unemployment and low inflation at the same time, that is, there is an opportunity cost in striving to achieve either of these macroeconomic objectives. For any given target rate of inflation, the Phillips curve indicates how much unemployment to expect. However, later studies showed that this trade-off seemed to exist only in the short run. The Phillips curve lost some credibility in the 1970s owing to the existence of **stagflation** – a situation that occurs when real national GDP, and, hence, the level of employment, falls at the same time as there being inflation in the economy. Stagflation is often caused by supply shocks, such as an oil crisis, which causes world oil prices to soar. This occurred in many parts of the world during the 1970s following the tripling in the price of oil.



■ **Figure 22.3** Stagflation caused by an oil crisis

◆ **Stagflation** occurs when there is rising inflation but falling national output and rising unemployment.

In Figure 22.3, the higher oil prices increase production costs so shifts the SRAS curve to the left from $SRAS_1$ to $SRAS_2$. This causes real GDP to decline from Y_1 to Y_2 at the same time as an increase in the price level from PL_1 to PL_2, that is, a situation of stagflation in the economy. Other possible causes of such supply shocks include financial crises, natural disasters (such as major earthquakes or severe flooding or droughts) and the widespread outbreak of contagious (infectious) diseases.

■ **Figure 22.4** Food shortages during the coronavirus outbreak (2020) caused supply-side shocks around the world

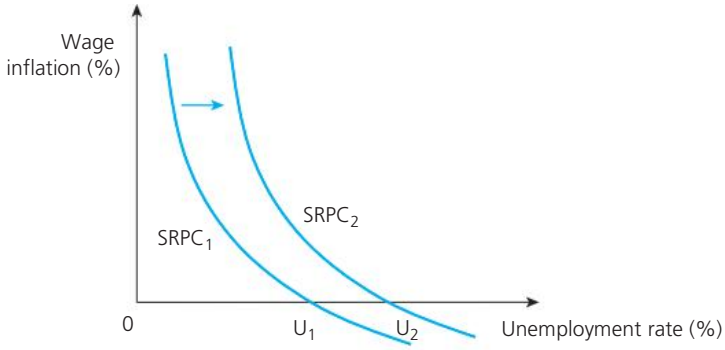


Figure 22.5 Stagflation and the short run Phillips curve

A period of stagflation in the economy will shift the Phillips curve outwards to the right, from $SRPC_1$ to $SRPC_2$ as shown in Figure 22.5. This increases the natural rate of unemployment from U_1 to U_2, thus stagflation creates a worse trade-off between low unemployment and low inflation.

The **long run Phillips curve** (LRPC) was shown to be vertical at the natural rate of unemployment (NRU), and so there is no trade-off as inflation would be stable. The NRU is the level of unemployment that is unavoidable (see Chapter 19). It can be found by adding up the number of people in frictional, seasonal and structural unemployment. Later studies showed that the Phillips curve can shift over time. For example, a reduction in structural unemployment caused by greater occupational and geographical mobility in the labour market will shift the Phillips curve to the left. Hence, the NRU is likely to fall if the government encourages more supply-side policies such as (re)training programmes for the unemployed. In this case, there is no trade-off between inflation and unemployment in the long run (see Figure 22.7).

◆ The **long run Phillips curve** (LRPC) shows that in the long run, there is only a single rate of unemployment (the natural rate) that is consistent with a stable inflation rate. Hence, the LRPC is vertical, that is, there is no trade-off between low inflation and low unemployment.

Figure 22.6 Reducing the natural rate of unemployment is a macroeconomic objective

In Figure 22.7, the SRPC shifts outwards owing to a supply-side shock caused by a prolonged global financial crisis, which leads to stagflation. In response, the government may use expansionary fiscal and/or monetary policies in an attempt to reduce the level of unemployment, moving the economy from point A to point B. However, this causes an increase in the general price level, so real wage rates fall because inflation reduces the real value of nominal wages. Therefore, workers push for higher nominal wages from W_1 to W_2, which further increases the costs of production for firms, which shifts the SRPC from $SRPC_1$ to $SRPC_2$, with unemployment reverting to its natural rate. Hence, attempts to reduce unemployment below its natural rate will be inflationary in the long run, as shown by the movement from point B to point C.

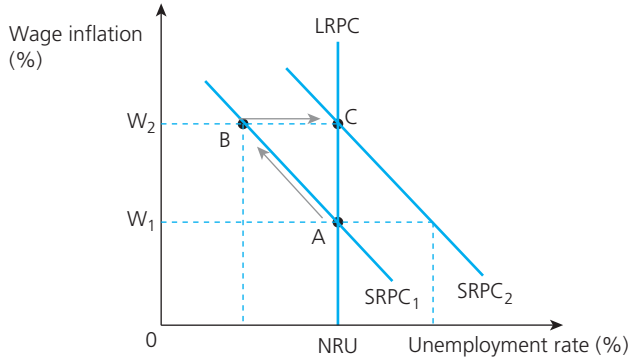


Figure 22.7 The long run Phillips curve (LRPC)

Hence, the LRPC model shows that any short-term attempts to reduce unemployment beyond its natural rate (NRU) will be ineffective and unsustainable in the long run as the resulting inflation leads to workers expecting and demanding higher nominal wages. The conclusion drawn was that any attempt to push unemployment below the NRU would cause accelerating inflation, with no long-term gain in the number of people in employment.

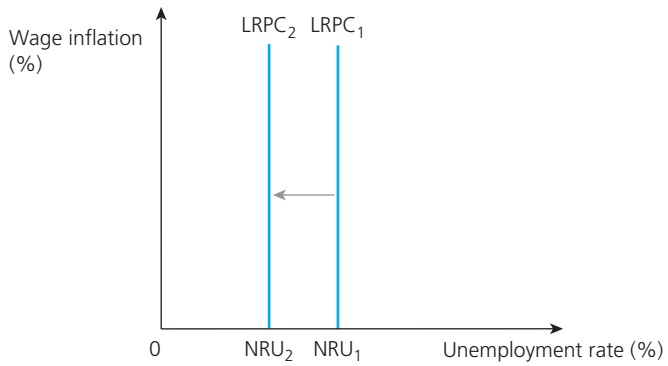


Figure 22.8 Shifts in the long run Phillips curve (LRPC)

Most governments strive to reduce the NRU (shifting the LRPC to the left) by creating incentives to work and encouraging more (re)training schemes for the unemployed to improve their occupational mobility. Between 1993 and 2008 (before the global financial meltdown), unemployment fell to record lows in many parts of the world, but inflation did not rise, as predicted by the Phillips curve. Many economists explain this by pointing to the successful supply-side policies that were pursued over the last 20 years, thereby shifting the LRPC to the left and reducing the NRU. This can be illustrated in Figure 22.8, with the LRPC shifting to the left and reducing the natural rate of unemployment in the economy from NRU_1 to NRU_2.

High economic growth and low inflation (AO3)

Economic growth is usually associated with an increase in the level of aggregate demand in the country. It occurs when there is an increase in any of its components, that is, consumption, investment, government spending and net exports. However, as the economy continually grows, the risks associated with inflation becomes higher especially as the economy approaches its full employment level of national output (see Figure 17.21 – Equilibrium in the Keynesian model (phase 2), page 259). If aggregate demand rises faster than aggregate supply in the economy, *demand-pull inflation* occurs (see Chapter 20), causing the general price level of goods

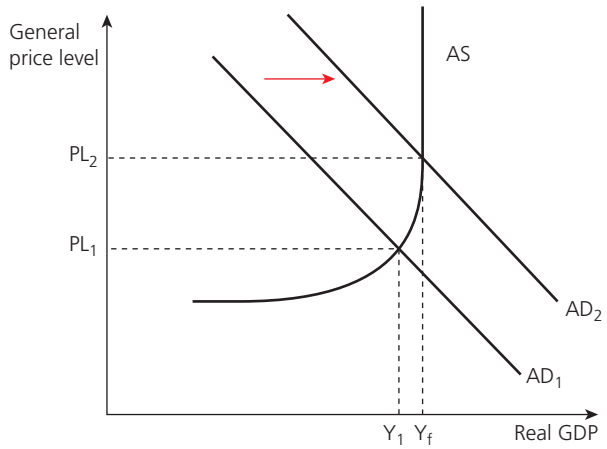


Figure 22.9 The trade-off between low unemployment and low inflation

and services to increase. This can also be shown in Figure 22.9 by the increase in the general price level from PL_1 to PL_2, caused by the increase in aggregate demand from AD_1 to AD_2, which increases real GDP from Y_1 to Y_f (the full employment level of national income).

Cost-push inflation can also occur because high economic growth and, hence, full employment make it more difficult for firms to attract skilled labour, so this can lead to wage inflation (and, hence, higher prices in general). Therefore, the pursuit of high economic growth may conflict with the macroeconomic objective of low inflation.

Monetary policy can generally be used to reduce inflation (by raising interest rates) or increase economic growth (by cutting interest rates). Hence, monetary policy (see Chapter 24) cannot solve both inflation and recession at the same time, again signifying the potential conflict between low inflation and low unemployment as macroeconomic objectives.

Nevertheless, growth does not necessarily mean uncontrollable inflation. Economic growth that is sustainable means the economy grows at a pace that enables a low and stable rate of inflation to take place, as aggregate supply increases in response to the higher level of aggregate demand in the economy. Low and manageable rates of inflation of up to 2 per cent are not usually damaging to the economy as higher prices can encourage firms to increase aggregate supply. It is only when inflation rises too quickly that it becomes harmful as it disrupts consumption for individuals and investment decisions for firms.

Top tip!

While the pursuit of economic growth can conflict with low inflation, they do not have to be mutually exclusive. Inflation that is monitored and controlled makes conditions far more predictable for consumers, producers and the government. Hence, controlled inflation can be positive for economic growth. Uncontrolled inflation will reduce the international competitiveness of the country, which can lead to negative economic growth and job losses.

Common mistake

'It is not possible for an economy to simultaneously achieve economic growth and price stability, as these macroeconomic objectives are in conflict.'

This can occur when both AD and AS increase, thereby creating more economic growth (real GDP growth) and employment, while even allowing prices to fall.

High economic growth and environmental sustainability (AO3)

Economic growth is generally associated with issues regarding environmental sustainability. As an economy grows, increased levels of production and consumption can create negative externalities (see Chapter 11) that cause problems for the environment, such as air pollution caused by over-congested roads, and threatening environmental sustainability through climate change, the loss of ecosystems, land erosion and destruction of water systems.

Negative externalities of production occur when producing a good or service causes a harmful effect to a third party, such as the excessive use of fossil fuels, which pollute the air with carbon emissions. As a country grows, many more people purchase private vehicles and households go on more overseas holidays, thereby adding to their carbon footprint.

Such environmental impacts can damage the well-being of individuals and societies and the standard of living in the long-term. For example, the amount of plastic waste in rivers and oceans has had devastating impacts on the natural environment. Hence, the pursuit of high economic growth can conflict with the goal of environmental sustainability.

Figure 22.10 The Taj Mahal and nearby banks of Yamuna River, polluted with garbage

Theory of Knowledge

If we 'know' that the world is being destroyed by the pursuit of economic growth and prosperity for the current generation, why are individuals and governments across the world not doing more to rectify this matter? Therefore, to what extent are humans truly rational beings?

Nevertheless, economic growth does not always lead to resource depletion or the destruction of the environment, and can be achieved through environmentally sustainable methods, such as the use of green technologies and renewable energy sources. When calculating economic growth (through changes in real GDP), an increasing number of countries are accounting for the value of environmental degradation, damage and destruction associated with increased levels of economic activity. In addition, governments and educational establishments are increasingly educating society about the importance of environmental protection, and promoting the use of energy conservation, waste minimization and the 3 Rs of environmental protection (reducing, reusing and recycling of the planet's scarce resources).

WISE ChoICES

Due to the potential conflicting nature of macroeconomic objectives, how does a government decide which form of **change** is best for the **economic well-being** of individuals and societies?

PAPER 1 EXAM PRACTICE QUESTION 22.1

Explain why there may be a trade-off between high economic growth and low inflation in the economy. *[10 marks]*

PAPER 1 EXAM PRACTICE QUESTION 22.2 (HL ONLY)

Explain why there may be potential conflict between low inflation and low unemployment in the short run. *[10 marks]*

WISE ChoICES

Does economic activity always cause **scarcity** of resources and therefore result in challenges for environmental **sustainability**?

WISE ChoICES

To what extent does the **interdependence** of macroeconomic objectives hinder a government's ability to achieve **economic well-being** for individuals and societies?

High economic growth and equity in income distribution (AO3)

Rapid economic growth often creates greater disparities in the distribution of income and wealth within the economy, that is, there is a widening gap between the rich and poor. For example, the world's richest people (see Table 22.1) had a combined wealth of $743.8bn in 2019, which represents an increase of over 74 per cent in just two years! By contrast, between 2017 and 2019, the USA economy grew by about 7.5 per cent and the European Union grew by around 6.4 per cent. In 2020, the combined wealth of these 10 people exceeded the nominal GDP of South Africa ($350bn), Singapore ($382bn), Argentina ($450bn), Switzerland ($715bn) and even Turkey ($740bn) – all of which are G20 countries. The Group of Twenty (G20) consists of the world's major advanced and emerging economies, working together for international co-operation on important aspects of international economic priorities.

Table 22.1

Rank	Name	Wealth ($bn)
1	Jeff Bezos	131.0
2	Bill Gates	96.5
3	Warren Buffett	82.5
4	Bernard Arnault and family	76.0
5	Carlo Slim Helu and family	64.0
6	Amancio Ortega	62.7
7	Larry Ellison	62.5
8	Mark Zuckerberg	62.3
9	Michael Bloomberg	55.5
10	Larry Page	50.8

Source: Forbes https://www.forbes.com/billionaires/#496c1863251c

ATL ACTIVITY 22.1

It has been suggested that if the world's ten richest men gave up 99% of their wealth, world poverty could be resolved overnight. By reallocating this wealth to the poorest 1% on the planet, each person would receive almost $10,000.

In small groups, discuss the social and economic arguments for and against the redistribution of wealth in such a way.

You should consider at least two of the WISE ChoICES concepts in your discussions.

The pursuit of rapid economic growth does not necessarily solve the socioeconomic problems of income inequalities (see Chapter 23). Although the country as a whole might experience economic growth, not everyone will benefit in the same way. Those earning minimum wage or on zero-contract hours are unlikely to earn enough income to build wealth in the same way

as high-income earners who have a higher marginal propensity to save (see Chapter 25). Zero-contract hours are a form of employment contract where the employer is not obliged to provide workers with any minimum working hours, while employees are not obliged to accept any work offered, which creates huge flexibility in the labour market.

However, economic growth also leads to greater tax revenues (from taxes imposed on income, profits, expenditure and the sale of imported products). Hence, economic growth enables the government to use these tax revenues to redistribute income and wealth in the economy. Therefore, so long as the tax system is progressive and equitable, there is not necessarily a conflict between economic growth and the equitable distribution of income and wealth. Progressive taxes mean an increasing percentage of income is paid in taxes as the income level increases (see Chapter 23).

ATL ACTIVITY 22.2

Investigate the extent to which people are on the national minimum wage and/or zero-hours contracts in your country or a country of your choice.

Chapter summary

- The pursuit of economic growth can pose a threat to environmental sustainability owing to the overuse of the planet's scarce and non-renewable resources.
- Economic growth is not necessarily equitable and often causes greater income inequalities as the rich get richer and the poor get relatively poorer.
- The macroeconomic objectives of low unemployment and low inflation can conflict because as the economy grows rapidly and more people are in employment, it is likely that inflationary pressures will occur.
- The macroeconomic objectives of high economic growth and low inflation can conflict because growth is associated with an increase in the level of aggregate demand in the country, which can cause demand-pull inflation especially as the economy approaches full employment.
- The macroeconomic objectives of high economic growth and environmental sustainability can conflict because greater production and consumption can create problems for the environment, such as air pollution caused by over-congested roads.
- The macroeconomic objectives of high economic growth and equity in income distribution can conflict because not everyone in society will benefit from growth in the same way.
- The short run Phillips curve (SRPC) shows a trade-off between unemployment and inflation. Lower unemployment, caused by a rise in aggregate demand, creates more spending in the economy thereby causing higher inflation. (HL only)
- The natural rate of unemployment (NRU) occurs the sum of frictional, structural and seasonal unemployment in the economy, irrespective of the general price level, that is, there is no trade-off between low unemployment and low inflation. (HL only)
- An increase in aggregate demand will tend to cause a movement up (to the left) along the short run Phillips curve because unemployment falls while the general price level begins to rise. (HL only)
- Any short-term attempts to reduce unemployment beyond its natural rate will be ineffective and unsustainable in the long run as the resulting inflation leads to workers expecting and demanding higher nominal wages. (HL only)
- The long run Phillips curve (LRPC) shows that in the long run, unemployment will revert to its natural rate that is consistent with a stable inflation rate so there is no trade-off between low inflation and low unemployment. (HL only)

WISE ChoICES

Does the **choice** to pursue macroeconomic objectives ultimately threaten the **sustainability** of the world's **scarce** resources?

23 Economics of inequality and poverty

Real-world issue

Why does economic activity vary over time and why does this matter?

Conceptual understandings (WISE ChoICES)

- **Change** in the conditions of the demand and supply sides of the economy cause economic activity to vary over time.
- Fluctuations in economic activity impact the **economic well-being** of individuals and societies.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- the relationship between equality and equity (AO2)
- the meaning of economic inequality in terms of unequal distribution of income and unequal distribution of wealth (AO2)
- the Lorenz curve and Gini coefficient (index) as a measure of economic inequality (AO2)
- a Lorenz curve diagram to show the distribution of income and possible changes in the distribution of income (AO4)
- a Lorenz curve from income quintile data (HL only)
- the meaning of poverty, including the difference between absolute and relative poverty (AO2)
- single indicators as measures of poverty including international poverty lines and minimum income standards (AO2)
- composite indicators as measures of poverty indicators including the Multidimensional Poverty Index (MPI) (AO2)
- the difficulties in measuring poverty (AO2)
- the main causes of economic inequality and poverty (AO2)
- the impact of income and wealth inequality on economic growth, standards of living and social stability (AO3)
- the role of taxation in reducing poverty, income and wealth inequalities, with reference to progressive, regressive and proportional taxes and with reference to direct and indirect taxes (AO2)
- the role of average and marginal tax rates (AO3) (HL only)
- the amount of indirect tax paid from a given level/amount of expenditure (AO4) (HL only)
- the total tax and average tax rates from a set of data (AO4) (HL only)
- further policies to reduce poverty, income and wealth inequality (AO3).

Relationship between equality and equity (AO2)

In the study of economics, the term equity differs from equality in reference to the distribution of income and wealth. **Equity** is based on the argument that income inequalities (such as wage differentials) are needed to create economic incentives for people to study and work harder; it is about economic fairness or justified inequalities, such as those with higher levels of qualifications, skills and experience being paid more than those without these attributes. **Equality** means the equal distribution of income in the economy, thus minimizing or eliminating the income gap between the rich and poor. It means there is parity in income (earnings) between individuals, where everyone is paid or receives the same, so that no inequalities exist. Economies face an

◆ **Equity** means economic fairness, such as those with higher levels of qualifications, skills and experience being paid more, that is, the existence of justified inequalities.

◆ **Equality** means there is parity in income (earnings) between individuals, that is, everyone is paid or receives the same, so no inequalities exist.

unequitable and/or unequal distribution of income and wealth owing to the natural unequal ownership of factors of production in a free market economy. For example, consider the wage differentials between professional footballers, doctors or pilots with those earning the national minimum wage.

Economists can evaluate government policies to promote equity by assessing their potential positive or negative effects on efficiency in the allocation of scarce resources. The three main ways of government intervention to promote equity are: taxation policy, government expenditure (through direct provision or through subsidies) and the use of transfer payments.

WISE ChoICES

To what extent does a person with a high disposable income enjoy improved **economic well-being**? What does a government need to **change** to improve the quality of life for people in society?

The meaning of economic inequality (AO2)

Owing to the inherent differences in investment in healthcare and education between countries and regions, the level of skills development of people is different in different parts of the world. The level and the quality of both healthcare and education usually correlates positively with skills development and the level of employment in the country. In a free market economy, where the government does not intervene in the operations of the free market mechanism, people own different levels of factors of production, which often leads to economic inequality. Essentially, economic inequality means unequal distribution of income and/or unequal distribution of wealth.

ATL ACTIVITY 23.1

Conduct a survey in your class on individual preference of economic equity or economic equality. Identify the people with similar views and organize a debate on the importance of economic equity versus economic equality.

Unequal distribution of income (AO2)

Since the eighteenth century, economists such as Adam Smith, David Ricardo and Thomas Malthus have studied how the national income of a country is distributed among its total population. One important meaning or interpretation of economic inequality is the unequal distribution of income in the economy. An **unequal distribution of income** in the economy exists when there are imbalances of income distributions, with very few members of the society enjoying a high concentration of the nation's income, compared to the larger proportion of the society that earns a relatively small percentage of the total income.

◆ **Unequal distribution of income** refers to imbalances of income distributions, that is, very few members of the society account for a disproportionately high concentration of national income.

There are various ways to describe the distribution of the income in an economy, all of which demonstrate some degree of inequality. For example, there could be differences between income distribution between countries, the geographical regions of a country, males and females, age groups, ethnicities or occupations.

ATL ACTIVITY 23.2

Refer to Figure 23.1 and reflect on the image. Using your personal experience, share with someone in your class how you perceive inequality in your everyday life. To what extent does your own understanding and perception of inequality change after hearing the views shared by your peers?

Income distribution can be examined at the individual or the household level (where more than one individual forms a household). Members of a typical household will contribute in different

proportions based on their own level of income. There is likely to be inequality within a family as family members earn different wages and salaries, suggesting that income is not distributed evenly. To compare nations, economists tend to use gross domestic product (GDP) per capita or GDP in terms of purchasing power parity (see Chapter 16) to show income inequalities (also, see Table 23.1).

Theory of Knowledge

On what grounds is inequality justified? What knowledge or evidence would you need to determine whether all households in the economy should receive an equal wage?

Figure 23.1 Unequal distribution of income exists in all societies, to some extent

Table 23.1 GDP for top and bottom 10 countries in purchasing power parity, 2020

Country	GDP per capita ($), PPP	Country	GDP per capita ($), PPP
Qatar	112,532	Madagascar	1,453
Macao	108,831	Sierra Leone	1,425
Luxembourg	93,734	Mozambique	1,180
Singapore	90,091	Malawi	1,163
Brunei	71,802	Liberia	1,161
Ireland	70,361	Niger	932
United Arab Emirates	66,616	Congo	827
Kuwait	65,515	Central African Republic	775
Norway	65,441	Burundi	660
Switzerland	59,019	Mauritania	372
Total (top 10)	803,942	Total (bottom 10)	9,948

Source: Trading Economics https://tradingeconomics.com/country-list/gdp-per-capita-ppp

With reference to Table 23.1, we can see the scale of inequality of income distribution just by comparing the percentage difference between the top and bottom countries in terms of GDP at purchasing power parity. The GDP of Qatar is 30,150 per cent more than that of Mauritania. Even more staggering, perhaps, is the fact that Qatar's GDP is more than 1,000 per cent higher than the combined GDP of the bottom 10 countries. Hence, this is indicative of huge income inequalities across the world.

Unequal distribution of wealth (AO2)

The second important meaning or interpretation of economic inequality is the unequal distribution of wealth in the economy. An **unequal distribution of wealth** (also called a **wealth gap**) refers to imbalances in the spread of a country's wealth among its citizens, that is, very few members of society account for a disproportionately large proportion of the wealth in the country. Wealth is the stock or accumulation of assets with a monetary value, including land, buildings, capital resources, savings and financial holdings (such as shares, stocks and investment funds).

◆ **Unequal distribution of wealth**, also called a **wealth gap**, refers to imbalances in the spread of a country's wealth among its citizens.

In 2010, the UN released data showing that the richest 1 per cent of the world's population owned about 40 per cent of the planet's productive assets. Oxfam released data in 2014 showing 85 of the wealthiest people in the world possessed more cumulative wealth than that of the bottom 48 countries (the equivalent of the wealth of more than 3 billion people). In March 2020, the BBC reported that 3,500 people in South Africa (the equivalent of just 0.01 per cent of the adult population) accounted for 15 per cent of the country's total wealth. It also reported that the bottom 90 per cent of people in South Africa (31.8 million people) account for only 14 per cent of the wealth in the country.

Figure 23.2 Wealth inequality leads to further gaps in economic development

In reality, there is a direct link between income and wealth, and they can affect one another. It is easier for higher income households to accumulate a greater stock of wealth through savings and purchasing assets or investment funds. This higher level of wealth can then lead to a larger income stream from their assets and investments. By contrast, low-income households are likely to incur debts (negative wealth), which reduce their income as households have to pay interest on their loans. For example, the bottom 50 per cent (or 17.7 million people) in South Africa had negative net wealth in 2020 due to their escalating debts. Again, this contributes to the widening wealth gap between the rich and poor.

In practice, data on income is more readily available than on wealth. Public authorities collect information on earnings for purposes such as tax and reporting of gross domestic product. The nature of wealth means that it is more difficult to measure, especially as asset prices (of currencies, gold, land, investment portfolios and so forth) fluctuate. Furthermore, even if the value of some assets might rise over time, they are not worth anything until the asset is sold.

Hence, at any point in time, it can be difficult to determine the actual value of an asset. Table 23.2 lists the mean wealth of adults in the 15 wealthiest and least wealthy countries, suggesting that Icelanders are the richest people in the world.

Table 23.2 Mean wealth of adults ($), 2018

Country	Mean wealth per adult ($)	Country	Mean wealth per adult ($)
Iceland	555,726	Niger	863
Switzerland	530,244	Guinea	816
Luxembourg	412,127	Chad	735
Australia	411,060	Uganda	710
USA	403,974	Guinea-Bissau	701
Belgium	313,045	Rwanda	660
Norway	291,103	Afghanistan	643
New Zealand	289,798	Sudan	530
Canada	288,263	Mozambique	482
Denmark	286,712	Madagascar	432
Singapore	283,118	Sierra Leone	355
France	280,580	Congo, Dem. Rep.	331
United Kingdom	279,048	Burundi	321
Netherlands	253,205	Ethiopia	167
Sweden	249,765	Malawi	141

Source: Credit Suisse, Research Institute's 'Global Wealth Databook', 2018

Different factors lead to the differences in wealth between countries, including:

- **Economic factors** – High national debt and high international loans are two of the key reasons why countries have different levels of wealth. Countries such as Yemen, Mozambique and The Gambia have a high GDP to debt ratio (of over 100 per cent). Although a few high-income countries like Japan, Singapore and Italy have similar level of debts, this is affordable debt. For example, Singapore ranks eleventh in terms of public debt as a percentage of its national income but at the same time also ranks eleventh in wealth per adult (see Table 23.2). Note that this emphasises the fact that a single indicator (like debt as a percentage of GDP) does not show the full picture of why a country does or does not have wealth.
- **Natural resources** – Some countries are endowed with natural resources such as oil, precious minerals or agricultural land conducive for production of cash crops. These can be traded with other countries to garner wealth and achieve economic prosperity. This enables these countries to enjoy a very high net wealth per person, such as the case in Qatar or Kuwait.
- **Environmental factors** – Some countries like Bangladesh and Mauritania often face problems like floods, draughts or extreme temperatures. All things being equal, this decreases their resource productivity and discourages investors, thereby making these countries less prosperous over time. This reduces the per capita wealth in these nations.
- **Physical factors** – Some countries or regions have a hostile or challenging landscape. This can prevent people of the country achieving individual prosperity. Examples of such physical

factors include places with very hot and dry climates or physical conditions that are not conducive to sufficient agricultural production to achieve wealth and prosperity.

- **Social factors** – Various social factors can limit the ability of the people of a country to accumulate wealth. A low level of education, population density, average age of marriage or female participation in the workforce can contribute to a lower level of wealth in the country.
- **Political factors** – The degree of political stability is a key determinant of economic prosperity in a country or region. Political instability caused by civil war, corruption or similar events is highly likely to limit a country's level of domestic economic activity and international trade, both of which are necessary for an increase in per capita wealth.

Measuring economic inequality (AO2, AO4)

The degree of income equality (or inequality) in an economy can be measured by the relative share of national income earned by given percentages of the population. **Deciles** refer to the statistical method of splitting data into tenths, with each part accounting for 10 per cent of the population. For example, if the top decile (10 per cent) of the economy's income earners account for 45 per cent of national income, this would signify huge income inequalities. **Quintiles** are used to divide statistical data sets into fifths, with each part representing 20 per cent of the population. Two interrelated statistical methods are used by economists to measure the degree of inequality in a country: the Lorenz curve and the Gini coefficient.

Lorenz curve and Gini coefficient (AO2, AO4)

The **Lorenz curve** is a graphical representation of income or wealth distribution in a country. It shows the proportion of overall income or wealth accounted for by each quintile or decile of the population. The tool is named after American economist Max Otto Lorenz (1905). The Lorenz curve is used to show the degree of income or wealth inequality, such as the poorest 10 per cent of income earners accounting for just 1 per cent of the nation's income (or real GDP).

◆ The **Lorenz curve** is a graphical representation of income distribution in a country, based on the income or wealth accounted for by each decile or quintile of the population.

In Figure 23.3, the 45 degree line shows perfect equality in income or wealth distribution. For example, at point A, the first four deciles account for 40 per cent of the national income. However, the Lorenz curve shows the actual income or wealth distribution in a country. Point B in Figure 23.3 shows a hypothetical situation where 60 per cent of the population account for just 20 per cent of the nation's income. So, the top four deciles must earn the remaining 80 per cent of the national income. The greater the area between the 45 degree line (showing perfect income equality) and the Lorenz curve, the greater the income inequality in the country.

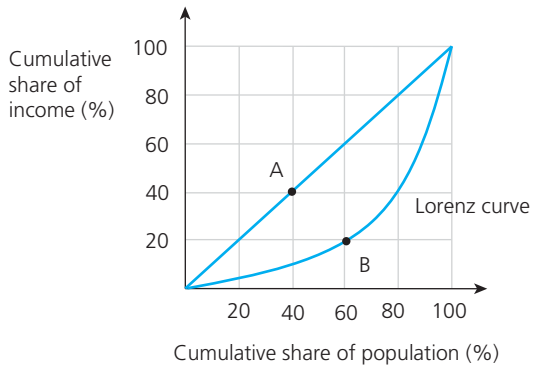


Figure 23.3 The Lorenz curve

CASE STUDY

Inequalities in South Africa

The richest 3,540 people in South Africa (around 0.01 per cent of the adult population) account for 15 per cent of the country's total wealth. South Africa is one of the most unequal countries in the world in terms of wealth distribution:

- The top 10 per cent (3.5 million people) control 86 per cent of all wealth.
- The top 1 per cent (350,000 people) control 55 per cent of all wealth.
- The top 0.1 per cent (35,000 people) control 29 per cent of all wealth.
- The top 0.01 per cent (3,540 people) control 15 per cent of all wealth.
- The bottom 90 per cent (31.8 million people) control just 14 per cent of all wealth.

Source: **https://www.groundup.org.za/article/15-sas-wealth-hands-just-3500-people-study-finds/**

Top tip!

If you are an HL student, you need to be able to construct a Lorenz curve from income quintile data.

The **Gini coefficient** (or **Gini index**) is a statistical tool that measures income or wealth inequality calculating a numerical value of a country's Lorenz curve. The technique is named after Italian statistician Corrado Gini in 1912, and is the most commonly used measure of inequality. The value of the Gini coefficient ranges from 0 (complete equality) to 1 (total inequality). Hence, a higher Gini coefficient means greater income or wealth inequality. At one extreme, if an individual accounted for all of the economy's national income, the Lorenz curve would pass the coordinates (0,0), (100,0) and (100,100) in Figure 23.3, leading to a Gini coefficient equal to 1. At the other extreme, there is total equality as shown by the (45 degree) line of perfect equality, that is, if every person in the country had the same income, the Gini coefficient would be zero.

◆ The **Gini coefficient** is a statistical tool that measures income or wealth inequality, with the outcome ranging from 0 (complete equality) to 1 (total inequality).

Table 23.3 Gini coefficients, 2020

Country	Gini coefficient
Lesotho	0.632
South Africa	0.625
Haiti	0.608
Botswana	0.605
Namibia	0.597
Zambia	0.575
Comoros	0.559
Hong Kong	0.539
Guatemala	0.530
Paraguay	0.517

Source: **https://worldpopulationreview.com/countries/gini-coefficient-by-country/**

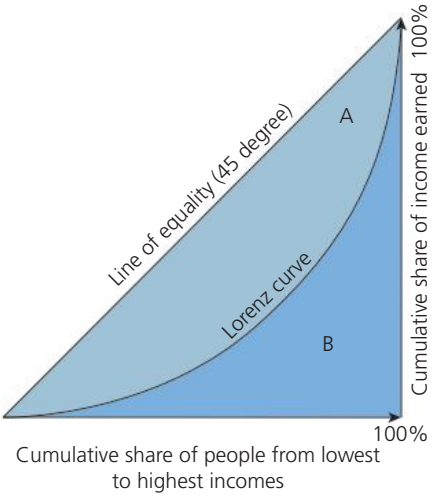


Figure 23.4 The Gini coefficient

The Gini coefficient is calculated by the ratio of the area under the Lorenz curve to the area under the 45 degree line of complete equality (which has a total area of (100 × 100) ÷ 2 = 5,000). With reference to Figure 23.4, the Gini coefficient is calculated by using the formula:

$$\text{Gini coefficient} = \frac{A}{A + B}$$

In general, low-income countries and/or those that suffer from a high degree of corruption have a high Gini coefficient, such as Lesotho (0.632) and South Africa (0.625). High-income countries tend to have a low Gini coefficient, such as Finland (0.272), Belgium (0.259) and Sweden (0.249). Table 23.3 shows the 10 countries with the highest Gini coefficients (largest degree of inequality).

A large degree of income inequality tends to have detrimental social, political and economic impacts on the country, such as greater household debt, political unrest and slower GDP growth. Ultimately, this can lead to higher poverty rates. Therefore, most governments strive to achieve greater income equality over time. They can use the Gini coefficient to measure whether income distribution is increasing or decreasing, and take appropriate policy measures to promote a fairer distribution of income and wealth in the economy.

Common mistake

Do not misinterpret the Gini coefficient as a comparative measure of national income between countries. Different countries with the same Gini coefficient do *not* have the same national income (GDP), that is, a high-income country and a low-income country can have the same Gini coefficients. This statistical tool is used simply to measure the degree of equality or inequality in income or wealth distribution.

PAPER 1 EXAM PRACTICE QUESTION 23.1

Explain how the Lorenz curve and the Gini coefficient are used to measure income inequality. *[10 marks]*

Meaning of poverty (AO2)

Poverty can be defined as the condition of an individual, household, community or country being extremely poor, that is, not having enough money to meet basic human needs such as food, clothing, shelter, healthcare and education. The World Bank describes poverty as a situation that people want to escape. However, definitions of poverty are relative because poverty varies considerably depending on the situation or context and between countries. Feeling poor in Finland or Norway may be different from living in poverty in Sierra Leone or Namibia. Different degrees of poverty also exists within the borders of a country, that is, there are wealthy people living in low-income countries and there are unemployed and homeless people who live in high-income countries. Tackling poverty is a key economic issue because, apart from humanitarian reasons, it represents economic inefficiency, which prevents people and economies from reaching their full potential.

◆ **Poverty** is the condition of an individual, household, community or country being extremely poor, that is, not being able to meet their basic human needs.

ATL ACTIVITY 23.3

Which IB Learner Profile attribute do you personally think is most important in combatting poverty? Now consider this question from the perspective of professional economists – which IB Learner Profile would they be most likely to relate to? Reflect on whether there is or is not a difference to these answers, and why you chose the particular IB Learner Profile.

Difference between absolute and relative poverty (AO2)

Absolute poverty (also known as **extreme poverty** or **abject poverty**) refers to those who are unable to access the basic human needs necessary for survival, such as food, clean water, health, shelter, primary education and information. Those who are in absolute poverty suffer from malnutrition and hunger, a lack of clean water, poor healthcare and inadequate shelter. Communities suffering from absolute poverty also experience a high death rate, including infant and child mortality rates, from preventable diseases such as malaria, cholera and water-contamination related diseases. Hence, life expectancy is low. Absolute poverty is commonly found in low-income countries.

◆ **Absolute poverty** exists when people are deprived of basic human needs for survival. People in absolute poverty suffer from malnutrition, hunger, a lack of clean water, poor healthcare and inadequate shelter.

In the words of Robert McNamara, a former president of the World Bank, absolute poverty is described as:

> *a condition so limited by malnutrition, illiteracy, disease, squalid surroundings, high infant mortality, and low life expectancy as to be beneath any reasonable definition of human decency.*
>
> *Robert McNamara, former president of the World Bank*

The concept of absolute poverty and the way to measure it has been an institutional debate for many decades. Back in 1990, absolute poverty was measured as earnings of $1 or less per day. With changes to inflation and the notion of purchasing power parity (see Chapter 16), this minimum was raised in 2015 to $1.90 or less per day. This amount is referred to as the international poverty line. According to the World Bank, around 736 million people live under the poverty line, so suffer from extreme poverty. Of the 736 million people, half of them live in just five countries: India, Nigeria, Democratic Republic of the Congo, Ethiopia and Bangladesh. Table 23.4 shows the countries with the highest percentage of total population living under $1.90. Of the 15 countries shown, 14 are in Africa. More than 50 per cent of the population from the first 14 listed countries lives in absolute poverty.

Top tip!

Note that while the World Bank uses $1.90 per day as the benchmark for measuring poverty, different measures of the poverty line exist in different countries.

Table 23.4 Percentage of population living in absolute poverty, 2018

Bottom 15 countries	Percentage of total population living on $1.90 or less per day
Madagascar	77.60%
Democratic Republic of the Congo	76.60%
Burundi	71.80%
Malawi	70.30%
Guinea-Bissau	67.10%
Central African Republic	66.30%
Mozambique	62.40%
Uzbekistan	62.10%
Lesotho	59.70%
Zambia	57.50%
Rwanda	55.50%
Nigeria	53.50%
Sierra Leone	52.20%
Turkmenistan	51.40%
Mali	49.70%

Source: World development indicators, World Bank, 2018

Top tip!

Poverty exists in all countries, although is more visible in poorer countries that have mass poverty. Even in wealthy nations, economic recession or natural disasters can cause widespread poverty. It is therefore incorrect to assume that only low-income countries suffer from poverty.

By contrast, the term **relative poverty** refers to incomes, and, hence, consumption levels, below the social norm within a country. It is a comparative measure so relative poverty will differ from country to country. It can lead to damaging effects on individuals and families in relative poverty, including social exclusion, homelessness or even suicide.

◆ **Relative poverty** in a country refers to individuals and households who are unable to earn the minimum amount of income needed for them to maintain the average standard of living in the community or country.

There is no universally accepted way to measure relative poverty as, by definition, it is a comparative idea. Different people belonging to different socioeconomic groups may have different opinions and perceptions of relative poverty. For example, citizens of some Eastern European countries or Balkan Peninsula countries may find themselves in relative poverty if they moved to live in Norway or Luxembourg, as the cost of living is much higher in these two countries. However, high-income earners in Eastern European countries may find living in Norway or Luxembourg is relatively cheap for them, depending on what type of lifestyle is considered.

Whether people live in absolute or relative poverty, there can be major negative consequences for these people and society as a whole. Examples of such problems include:

- Low living standards – Those in poverty, be it absolute or relative, experience a low quality of life as they are unable to meet their basic needs or to experience a better standard of living. Imagine if your household had to survive on $1.90 per day, every day for the foreseeable future!
- Lack of access to healthcare and education – Poverty-stricken individuals and households are unable to afford quality healthcare and education. This hinders the human capital stock of the country and thus compromises prospects of economic growth and development.
- Conflict and war – History has shown that poverty often leads to political instability and social unrest, which can ultimately lead to civil war in extreme cases. This can lead to other issues such as mass emigration of the population, which undermines the country's ability to recuperate from poverty. For example, Syria experienced civil war in 2019–20 causing the displacement of millions of people, mainly women and children.
- Poverty hinders the ability of governments to achieve other macroeconomic objectives such as economic growth and low unemployment.

Top tip!

The causes and consequences of poverty can create a **poverty trap**, that is, the poor become even poorer. For example, low incomes lead to low savings, reduced funds for investments, lower productivity, decreased national output and, hence, a diminished quality of life. This increases poverty even further. Hence, eradicating poverty is a key macroeconomic priority of many governments.

Measuring poverty (AO2)

As mentioned, the definition of poverty is a complex issue and has been debated and argued over many decades. This makes the measurement of poverty challenging, and various methods have evolved over time. Some examples of notable definitions and descriptions of poverty provided by economists, government agencies and international organizations are provided below:

- Adam Smith (1723–90) – The Scottish economist famously referred to as the Father of Economics defined poverty as the inability to afford 'not only the commodities which are indispensably necessary for the support of life, but whatever the custom of the country renders it indecent for creditable people, even of the lowest order.'
- Benjamin Seebohm Rowntree (1871–1954) – The English social pioneer and reformer stated the 'poverty line represents the minimum sum on which physical efficiency could be maintained. It [is] a standard of bare subsistence rather than living.'

CASE STUDY

Poverty in the USA

According to the US Census Bureau, 15% of individuals in the United States live below the poverty line (the minimum level of income judged to be adequate in a particular country). The aftershock of the global financial crisis in 2008 meant rising unemployment and related economic problems for the economy. This resulted in over 46 million people living in poverty in America, more than at any point in the nation's history. In 2020, the US recorded over 44 million unemployed people due to the coronavirus pandemic, causing greater issues of poverty in the country.

- Peter Townsend (1928–2009) – The British Professor of International Social Policy at the London School of Economics referred to poverty as 'individuals or families whose resources are so small as to exclude them from the minimal acceptable way of life of the member state in which they live'.
- Amartya Sen (b.1933) – The Indian economist who holds the post of Professor of Economics and Philosophy at Harvard University and is also the recipient of the Nobel Memorial Prize in Economic Sciences (1998) states that poverty is the failure to achieve certain minimum capabilities, which are not fixed over time or over societies.
- The World Bank – The international financial institution that provides financial services to governments of low-income countries refers to a person being considered poor if his or her income level were to 'fall below some minimum level necessary to meet basic needs'.
- The United Nations, the intergovernmental organization that exists to foster international co-operation, peace and security states that 'poverty is a denial of choices and opportunities, a violation of human dignity. It means [a] lack of basic capacity to participate effectively in society. It means not having enough to feed and clothe a family, not having a school or clinic to go to, not having the land on which to grow one's food or a job to earn one's living, not having access to credit. It means insecurity, powerlessness and exclusion of individuals, households and communities.'
- Sir Tony Atkinson (1944–2017) – British economist and Professor at the London School of Economics who worked on inequality and poverty for more than four decades, showed that poverty has various manifestations, including:
  - lack of income and productive resources sufficient to ensure sustainable livelihoods
  - hunger and malnutrition
  - ill health
  - limited or lack of access to education and other basic services
  - increased morbidity and mortality from illness
  - homelessness and inadequate housing
  - unsafe environments
  - social discrimination and exclusion
  - characterized by lack of participation in decision-making and in civil, social and cultural rights.

CASE STUDY

Poverty in Burundi

Burundi is one of the poorest countries on the planet, with around 65% of its population of 11.1 million people living below the poverty line. GDP per capita in the East African country is a meager $800 (or $2.19 per day), with 45.8% of its GDP coming from agriculture. Burundi has a staggering infant mortality rate of 60.4 deaths per 1,000 live births, one of the highest rates in the world. The median age is just 17 years and life expectancy is 60.5 years.

Source: adapted from CIA World Factbook (https://goo.gl/WdL2a)

Single indicators of measuring poverty (AO2)

The different interpretations of poverty, owing to it being multidimensional, mean that there are different ways to measure poverty. Economists use both single and composite indicators for this purpose. Single indicators of measuring poverty include using international poverty lines and minimum income standards.

International poverty lines

An **international poverty line** (or **poverty threshold**) is the minimum threshold level of income that a person in a given country must earn or have access to in order to meet the basic needs necessary for human survival. Below this poverty line, people are considered to be living in poverty. For example, the first of the United Nations' Sustainable Development Goals is to end poverty in all its forms everywhere. According to the UN, more than 700 million people (or 10 per cent of the world's population) still live in absolute or extreme poverty. This is defined by the World Bank as those living on $1.90 or less per day – which makes this the most widely

◆ An **international poverty line** is the minimum threshold level of income that a person must earn or have access to in order to meet the basic needs necessary for human survival.

recognized and accepted International Poverty Line (IPL). The amount is based on the value of goods needed to sustain the life of one adult.

However, the IPL method does not take into account access to sanitation, water and electricity, and what effect such access has on people's quality of life. Hence, a more useful measure of a poverty line could consider the minimum amount required to meet basic needs, such as food, drinking water, sanitation, shelter, a basic level of education and healthcare. Still, achieving consensus on one single indicator to measure poverty has been continually contested by the likes of the World Bank, the United Nations, economists and policymakers.

After all, determining and calculating a minimum cost to live healthily is not devoid of challenges. First of all, food requirements are linked to energy requirements measured in calories. However, the required food calorie intake will depend on physical attributes of individuals, such as their age, gender, weight or lifestyle. For example, an athlete may have a higher energy requirement than a writer who requires a comparatively lower energy to survive. Also, energy requirements vary between males and females and children or the elderly.

Similarly, in more economically developed regions, higher international poverty lines are more appropriate owing to the higher costs of living. For example, the cost of rent varies in different cities and countries, which makes it difficult to determine a minimum level cost required to sustain living. Economists argue that in general the cost of the shelter is the biggest cost that adds to the minimum cost essential for living so have depended on the real estate market to calculate the poverty line. According to the World Bank, the process of determining an international poverty line starts with estimating the **national poverty line**. Rich countries with higher national income have a higher national poverty line, whereas the poorer countries have a relatively lower poverty line. For example, in the USA, the federal poverty level (FPL) is an annual income level based on the number of members in the household. For a single-person household, the 2020 FPL is $12,760 (or just under $35 a day). For each additional household member, the FPL increases by $4,480 (or just over $12 per day).

◆ The **national poverty line** is the level of income below which a person in the country cannot meet their basic minimum needs in terms of shelter, nutritional level and clothing.

■ **Figure 23.5** Could you live on less than $2 per day for the rest of your life?

As different countries use different ways to define poverty and to determine absolute and relative poverty, it is impractical to collect the national poverty line data for all countries and find a mean (average) for calculating an international poverty line. To overcome this limitation, the World Bank and some consultants studied the national poverty level of some of the poorest nations using purchasing power parity exchange rates to ensure that price of goods and services are standardized. In 1990, from the data of the six poorest countries in the world, the World Bank

determined the international poverty line as $1 a day. In 2008, this was increased to $1.25. The IPL was last raised in 2011 to $1.90 per day.

Minimum income standards

Rising costs caused by inflation cause so many more people to be deprived of a decent standard of living. The **Minimum Income Standard** (MIS) is a research method that determines the lowest amount of income needed for what members of the public think is an acceptable standard of living in the country. The MIS enables individuals and households to be able to live in a socially acceptable way. Within a country or region, adjustments are made to the MIS for those living in large cities (owing to higher costs of living, such as housing) and rural or remote areas of the country.

◆ The **Minimum Income Standard** (MIS) measures the lowest amount of income needed for what members of the public in the country think is acceptable to be able to live in a socially appropriate way.

The Joseph Rowntree Foundation in the United Kingdom, the Centre for Research in Social Policy (CRSP) at Loughborough University and the Family Budget Unit at the University of York were all involved in developing the MIS as a measure of poverty. Traditional measurements of the poverty line made reference to meeting basic needs such as food, shelter and clothing, whereas the MIS approach also emphasizes the use of qualitative research from a variety of focus groups to gather all the necessary information to determine the minimum acceptable income level to ensure a basic standard of living in the country. Similar methods were applied and adapted in other developed countries including Japan, France, Austria and Portugal.

In the United Kingdom, the MIS is used for the following purposes:

- To calculate the Living Wage. In the UK, employers must pay £10.75 per hour (around $14 per hour) for those working in London and £9.30 an hour (just over $12 per hour) for those outside of London, so that employees can reach a minimum acceptable living standard.
- To act as a quantitative benchmark for non-governmental organizations and charities to determine a point of reference for helping people in need, including those who are homeless.
- To calculate the costs of bearing and raising a child. This can help the government to determine the level of social security and transfer payments such as child allowances (child benefits).

Figure 23.6 The MIS is used to calculate the Living Wage, a measure widely used in the UK

Composite indicators of measuring poverty (AO2)

Composite indicators of measuring poverty use a range of indicators (or dimensions) to calculate the degree of poverty for a given population. Although they are more difficult to measure, composite indicators consider both breadth and depth of poverty within a country, that is, the proportion of the population that is deemed to be poor, and the extent to which they suffer from poverty.

The **Multidimensional Poverty Index** (MPI) is a composite measure of poverty that identifies numerous deficiencies of individuals and households in reference to their level of health, education and standards of living. The MPI considers multiple factors that influence or deprive people of their quality of life, ranging from sanitation and the composition of household flooring (dirt, sand or dung) to child mortality rates and the average years of schooling. The MPI was developed by academicians and policymakers from the United Nations Development Programme (UNDP) and Oxford University's Poverty and Human Development Initiative. The MPI comprises a research survey conducted with households, based on three dimensions of poverty measured through ten indicators with different weightings (see Table 23.5).

◆ The **Multidimensional Poverty Index** (MPI) is a composite measure of poverty that identifies numerous deficiencies of individuals and households based on three dimensions or indicators of poverty – health, education and standards of living.

The agenda for the UN's Sustainable Development Goals (see Chapter 33) raises the need to consider a multidimensional approach to measuring global poverty that goes beyond the traditional use of measuring poverty by single indicators (such as income). According to the UNDP's MPI data, around 23.1 per cent of the total population across more than 100 countries is suffering from severe multidimensional deprivations. People fall into the category of being multidimensionally poor if they face hardship in at least one third of the 10 weighted indicators in the MPI (see Table 23.5).

Table 23.5 The Multidimensional Poverty Index

Dimensions of poverty	Indicator	Deprived if living in a household where…	Weight
Health	Nutrition	A child or an adult under 70 years of age is undernourished.	1/6
	Child mortality	Any child has died in the family in the five-year period preceding the survey.	1/6
Education	Years of schooling	No household member aged 10 years or older has completed 6 years of schooling.	1/6
	School attendance	Any school-aged child is not attending school up to the age at which he/she would complete class 8.	1/6
Standard of living	Cooking fuel	The household cooks with dung, wood, charcoal or coal.	1/18
	Sanitation	The household's sanitation facility is not improved (according to the UN's SDG guidelines) or it is improved but shared with other households.	1/18
	Drinking water	The household does not have access to improved drinking water (according to the UN's SDG guidelines), or safe drinking water is at least a 30-minute walk from home, round trip.	1/18
	Electricity	The household has no electricity.	1/18
	Housing	Housing materials for at least one roof, wall or floor are inadequate: the floor is of natural materials and/or the roof and/or walls are of natural or rudimentary materials.	1/18
	Assets	The household does not own more than one of these assets: radio, TV, telephone, computer, animal cart, bicycle, motorbike or refrigerator, and does not own a car or truck.	1/18

Source: **http://hdr.undp.org/en/2018-MPI**

The MPI offers a thorough picture of global poverty (see box below), which helps organizations and governments to monitor the first of the UN's Sustainable Development Goals – to end poverty in all its forms. In particular, the MPI helps policymakers to monitor Target 1.2 of the SDGs, which is to reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions.

The 2019 UNDP Global Multidimensional Poverty Index (MPI) data offer some key findings regarding the global scale of poverty, including:

- Across more than 100 countries, 1.3 billion people (23.1 per cent) are multidimensionally poor.
- Half of the 1.3 billion multidimensionally poor people are children under the age of 18. A third are children under the age of 10 (more than 433 million children are yet to reach their tenth birthday).
- Two-thirds of multidimensionally poor people live in middle-income countries.
- There is wide variation across countries in inequality among multidimensionally poor people.
- There is little or no association between economic inequality (measured using the Gini coefficient) and the MPI value.

Source: United Nations Development Programme 'Illuminating Inequalities', 11 July 2019

Table 23.6 shows the most multidimensionally poor countries among all the economically less developed countries. The lower the MPI value, the more deprived people are in the country.

Table 23.6 Countries identified as most multidimensionally poor, 2019

Country	MPI value
Armenia	.001
Serbia	.001
Turkmenistan	.001
Ukraine	.001
Jordan	.002
Kazakhstan	.002
Montenegro	.002
Trinidad and Tobago	.002
Thailand	.003
Albania	.003

Source: **http://hdr.undp.org/sites/default/files/mpi_2019_publication.pdf**

Difficulties in measuring poverty (AO2)

Economists have tended to use income as the key indicator of measuring standards of living and poverty. There is a clear correlation between a higher income per capita and higher human development indicators. For example, in high-income countries there are significantly lower rates of people living below the national and international poverty line. They also enjoy longer life expectancy, better technological advancements, regulated justice systems, and greater access to healthcare and education. However, a high income does not guarantee all people escape poverty. Inequalities and relative poverty still exist in high-income countries.

Figure 23.7 Poverty can be found in every society

The United Nations' Sustainable Development Goal 1 (No poverty) and Goal 2 (No hunger) directly deal with multidimensional aspects of poverty. For many decades, different governmental and non-governmental organizations have been devising plans, policies and projects to eliminate absolute poverty. To facilitate this, it is essential to measure the extent of poverty in the world to see if the right actions are taken towards helping the poor and to assess the social and economic impacts of these initiatives.

However, as mentioned above, poverty is a multidimensional issue. Along with the income element of the monetary dimension, there are other crucial non-monetary indicators of poverty, such as education, health, sanitation, water, use of power, and so forth. Therefore, a single indicator by measuring only income does not reflect or address the multiple complexities of poverty. This is the reason why more thorough and complex measures such as the International Poverty Line and the Multidimensional Poverty Index were developed.

Nevertheless, the measurement of poverty is complicated even within a country. The whole population of a country does not develop at the same time, and neither do areas or regions within the country. As a result, measuring poverty in a meaningful way is very challenging. Some of the other challenges are:

- There have been discrepancies and debates regarding how the national poverty line of very poor countries would translate to the international poverty line for all nations. The International Poverty Line (IPL) is commonly used to measure if the people of the country have sufficient income to be able to meet their basic survival needs.
- Relative poverty is a highly subjective issue. Although most countries use absolute poverty to measure an individual's or household's ability to meet their day-to-day needs, a few countries prefer to measure this in terms of relative poverty line, that is, the percentage of the population with earnings less than a predetermined percentage of the median income within the country.
- The nature of the low income is often ignored when measuring poverty. While economists and policymakers agree that the value of income is a useful indicator of poverty, a permanently low income (in poor countries) is not same as low income that is only in other countries that might be suffering from a short-term recession. A permanently low level of income creates a poverty trap, possibly for many generations to come.

- There are limitations to the purchasing power parity figures used to compare international measures of poverty. This is because PPP values are likely to differ based on location (such as higher living costs in city centre locations) and time (as exchange rates tend to fluctuate over time). This makes it more difficult to measure both absolute and relative poverty against an international poverty line.

ATL ACTIVITY 23.4

Discuss the extent to which educational attainment and academic qualifications should determine people's economic opportunities. If schools were to ban all external assessments, how should universities and employers know which candidates to select?

Causes of economic inequality and poverty (AO2)

The main causes of economic inequality and poverty include the following:

- inequality of opportunity
- different levels of resource ownership
- different levels of human capital
- discrimination (gender, race and others)
- unequal status and power
- government tax and benefits policies
- globalization and technological change
- market-based supply-side policies.

Inequality of opportunity

Inequality of opportunity means that some individuals and societies have less or no access to employment and consumption opportunities so are deprived of economic prosperity. It exits as a result of the unequal progress between various countries. For example, some countries are able to achieve social, political, environmental and economic improvements while others have not be able to do so. Economists believe that inequality of opportunity, as opposed to inequality that results from personal effort, contributes negatively to economic growth and development.

◆ **Inequality of opportunity** means that some individuals and societies have less or no access to employment and consumption opportunities so are deprived of economic prosperity.

The concept of inequality of opportunity is not a new one: just before the start of the Second World War, US President Franklin Roosevelt argued for creating conditions that offer equal opportunity. British economist Sir Anthony Barnes Atkinson (1944–2017) claimed that 'Inequality of outcome among today's generation is the source of the unfair advantage received by the next generation. If we are concerned about equality of opportunity tomorrow, we need to be concerned about inequality of outcome today.'

The living conditions needed to allow a person to excel and to contribute positively to society largely depends on the factors outside the control of individuals. For example, the economy's provision of healthcare and education overrule individual factors such as effort, initiative, choice and dedication. Therefore, progress in a child's life is primarily dependent on a chance factor, that is, on the place and time a child is born. In general, a child who is born in a sub-Saharan country tends to do worse than a child born in Scandinavian countries at the same time because time and place can affect a child's access to basic opportunities such as water, education, electricity and sanitation.

Figure 23.8 Inequalities of opportunity are a major source of poverty in society

WISE ChoICES

To what extent do you support the view that government **intervention** in dealing with income inequalities is ineffective if **equity** and **sustainability** are compromised?

Different levels of resource ownership

The lack of natural resources and/or the poor management of these scarce resources will generally reduce a country's potential net export earnings. All things remaining equal, this will result in a fall in the level of consumption and investment in the economy, resulting in greater economic inequalities and poverty. Hence, differences in the levels of resource ownership between countries has a direct impact on their level of economic prosperity or poverty.

An increase in an economy's national income may not lead to poverty eradication or the narrowing of the inequality gap because GDP is not shared evenly between skilled and unskilled labourers. The share of national income that is divided between entrepreneurs (capitalists) and the labourers has a direct influence on the level of income inequality of the country. The larger the share of profits of a business that goes to entrepreneurs, the greater income and wealth inequalities will tend to be, and vice versa.

Different levels of human capital

Human capital refers to the stock or accumulation of skills, knowledge and experiences possessed by individuals or the workforce of a country. It is viewed in terms of their value or cost to an organization or the economy as a whole. In general, the lack of sufficient provision of and investment in education and training leads to mass poverty in the country. Without the necessary knowledge and skills base, the workforce will be far less productive, so national income will be significantly lower than its potential. The lack of educational opportunities to develop an economy's human capital is therefore a major factor of economic inequality and poverty.

◆ **Human capital** is the accumulation of skills, knowledge and experience of the workforce, viewed in terms of their value or cost.

The quality of labour resources in the economy is not only determined by the level of education and training, but is also a direct outcome of good healthcare provision. Adequate healthcare provision helps an economy to achieve a higher level of productivity and increased labour mobility. Fundamentally, a healthier workforce leads to a greater participation rate of the workforce, greater national output and employment opportunities (see Figure 23.9).

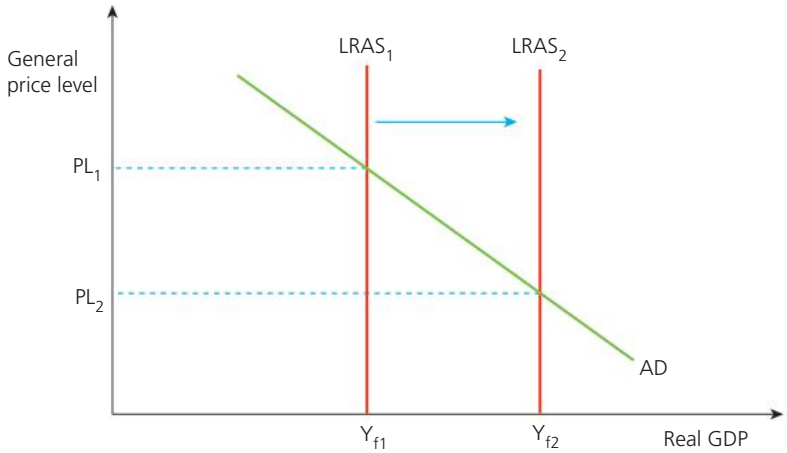


Figure 23.9 The impact of investment in human capital

Greater investment in health and education in order to improve the country's human capital will tend to lead to a favourable change in the productive capacity of the economy. This is shown by an outwards shift of the economy's long run aggregate supply from LRAS$_1$ to LRAS$_2$, thus increasing national income from Y_{f1} to Y_{f2}, and enjoy a lower average price level as prices fall from PL$_1$ to PL$_2$. However, low-income countries lack the financial means to achieve this, so can be trapped in poverty as a result.

Discrimination (gender, race and others)

Discrimination can be described as any form of social exclusion, that is, communities do not accept or respect groups of people based on differences in gender, race, ethnicity, religion, political or ideological views, income level or any other socioeconomic factor. Social exclusion often brings friction and conflict, which creates an environment less conducive for economic growth and development.

Economists regard negative discrimination, in all its forms, as inefficient and detrimental to economic prosperity. Across the world, more women tend to suffer from poverty than men, mainly due to the social prejudice against females. For example, on average, women are less likely to be in paid employment on a full-time basis and tend to earn less than men. This represents an inefficient use of labour resources.

Gender discrimination often results from social norms that have been instilled in cultures and societies. In many cultures, women are perceived as the caretaker of the children and the family, making it difficult for them to participate in the labour workforce. Owing to the imposed restrictions to join the workforce, a nation could be deprived of achieving a higher level of productivity and prosperity simply because of the untapped female labour force. Hence, there are significant economic opportunity costs of discrimination.

The World Economic Forum (WEF) uses the Global Gender Gap Index (GGGI) to quantify gender discrimination across different countries. The value of the GGGI ranges from 0 (indicating the highest possible gender gap) to 1.0 (indicating no gender gap or inequality). The WEF uses the following parameters to calculate the GGGI:

- Economic participation and opportunity
- Educational attainment
- Health and survival
- Political empowerment.

Table 23.7 Top and bottom 10 countries based on GGGI scores

Country	GGGI score
Iceland	0.877
Norway	0.842
Finland	0.832
Sweden	0.82
Nicaragua	0.804
New Zealand	0.799
Ireland	0.798
Spain	0.795
Rwanda	0.791
Germany	0.787

Country	GGGI score
Yemen	0.494
Pakistan	0.564
Syria	0.567
Iran, Islamic Rep.	0.584
Chad	0.596
Lebanon	0.599
Saudi Arabia	0.599
Oman	0.602
Morocco	0.605
Côte d'Ivoire	0.606

Source: Global Gender Gap Report 2020, World Economic Forum

Notice that some low-income countries appear in the top 10 for gender equality (Rwanda and Nicaragua), whereas some high-income countries feature in the bottom 10 for gender equality (Saudi Arabia and Oman).

CASE STUDY

Gender inequalities in India

Increasing women's ownership of land in rural areas can help cut poverty and boost agricultural output in a country like India where dependence on agriculture is high. Nearly two-thirds of the Indian population depends on agriculture for their livelihood. India's Census figures show that 65 per cent of the total female workers in India are engaged in agriculture.

Figure 23.10 Rural farming in India

However, land holdings by women remained restricted to only 15 per cent in north India and between 30 and 35 per cent in south India. The contribution of women towards agricultural output can be substantially improved by increasing access to private property resources and by changing inheritance laws to allow women to inherit land.

Source: https://www.livemint.com/Politics/TOZ8tlmWzwHfXeJ3ffyQMM/Womens-ownership-of-land-in-rural-areas-can-help-cut-povert.html

Unequal status and power

Power and money, in terms of wealth and income, are main driving forces behind status in any society. The concept of status dictates the resources and economic power, signalling a particular group is 'better' or more worthy than others. By contrast, marginalized people are exploited (owing to discrimination, dehumanization and humiliation), and deprived of basic social and economic opportunities, such as access to food, drinking water, sanitation, education and healthcare. Such exploitation over generations widens income, wealth and gender inequalities, all of which deepen the level of poverty in society.

In the long-term, establishment of unequal status and power in an economy deters efficiency and innovation, and instead encourages more social and economic divisions in society. Economies with highly unequal status and power not only widen the imbalance between rich and poor, but also often prevent policymakers from making rational economic decisions.

Government tax and benefits policies

Fiscal policy (see Chapter 25) has long been argued to be one of the main policy instruments that governments can use to influence the level of economic health in the country. Fiscal policy refers to the use of government taxes and spending policies to achieve macroeconomic objectives. Tax revenues are used to fund the expenditures of the government for the day-to-day running of the country (such as spending on social security schemes) and for long-term investments (such as improving infrastructure and human capital).

A major role of taxation is to promote great income equality by redistributing income to help the relatively less well-off in society. Direct taxes are taxes on earnings and are used as a mechanism to redistribute income from the rich to the poor. For example, the highest earners in Belgium pay up to 55 per cent income tax. To create incentives to work, individuals are granted a personal tax allowance, that is, they can earn up to a certain amount before they are taxed. Countries with a poor or ineffective tax system and do not have sufficient welfare benefits tend to create a wider income and wealth gap between the rich and poor. For example, in Hong Kong, top income earners pay no more than 15 per cent income tax yet there is only a very basic welfare system to cater for those in extreme poverty. The combination of low taxes and low welfare benefits creates poverty and inequalities in the economy, ceteris paribus. However, this does not necessarily mean that high rates of taxes create less poverty as the use of overly high rates of income tax can create disincentives to work.

Essentially, the tax raised and the welfare benefits offered to the community can significantly reduce the gap between rich and poor, if the transfer payments are used for the direct benefit of the marginalized members of society. However, not all taxes reduce inequality. Regressive taxes, such as sales taxes imposed on the purchase of goods and services, can increase the inequality gap in society.

PAPER 1 EXAM PRACTICE QUESTION 23.2

Using real-world examples, evaluate the view that government tax and benefits policies are the most effective way of addressing inequality and poverty in an economy. *[15 marks]*

Globalization and technological change

Globalization is the process of greater interaction and integration of goods, services, cultures, languages, social behaviours among people, business operations and governments worldwide. Globalization is an all-encompassing concept of cultural, political, social, technological and economic exchange of tangible and non-tangible products. It also signifies the growing degree of interdependence of the world's economies.

◆ **Globalization** is a multidimensional process of greater interaction and integration of goods, services, cultures, languages, social behaviours, business operations and governments worldwide.

However, globalization has been criticized for not benefiting many poor people living in different parts of the world. There are no guarantees that the social and economic benefits of globalization are distributed fairly in society. Globalization has also caused market instability due to the collapse of local businesses that cannot compete with large multinational companies. This has led to greater inequality and poverty in many low-income countries. For example, trade liberalization

policies have led to an increased level of wage differentials (wage gaps) between the educated and uneducated, and the skilled and unskilled not only in economically less developed countries but also high-income countries.

Similarly, technological change has also caused greater discrepancies in equality and poverty. The Industrial Revolution helped some economies to achieve an unprecedented level of economic growth and prosperity. However, the use of capital-intensive production methods led to huge job losses and therefore an increase in inequalities. While technological improvements drive economic growth and development in almost every country, they strip the opportunity for low-skilled labourers to earn incomes sufficient to eradicate poverty.

Market-based supply-side policies

Market-based supply-side policies (see Chapter 26) focus on freeing up markets and improving incentives to work and invest, thereby increasing the potential output of the economy. Examples include policies to encourage competition and efficiency in industries, labour market reforms to increase aggregate supply, and tax reforms to improve incentives to work, productivity and investments.

Supply-side policies such as deregulation, privatization, trade liberalization and anti-monopoly measures (all covered in Chapter 26) help to increase the economy's long run aggregate supply by increasing the ability and propensity of firms to produce and labourers to supply their services. However, these market-based supply-side policies can create more inequality and poverty in the following ways:

- **Deregulation** – the reduction or elimination of institutional powers in a country is intended to stimulate more competition in the economy, including competition from overseas companies. However, this could cause domestic businesses to collapse, thereby leading to further job losses (unemployment) and creating greater inequalities in the country.
- **Privatization** – the transfer of ownership of property or businesses from the public sector to the private sector can also lead to unemployment and an uneven wealth distribution within the country. This is because privatized businesses are driven by profits, so will aim to cut costs in order to make efficiency gains.
- **Trade liberalization** – freer and fairer international trade increases economic efficiency, but often leads to a wage gap between labourers because of the greater opportunities presented to the highly skilled workers. The resulting increase in wage differentials causes a widening of income inequalities.
- **Anti-monopoly regulation** – these are laws against anti-competitive behaviours in order to prevent abuse of monopoly power and to encourage competition. However, governments in low-income countries are not always able to prevent monopoly abuse, which then limits competition. Hence, the exploitative behaviour of large multinational companies with monopoly power often leads to increased inequalities in overseas countries. For example, many multinational companies have been accused of exploiting workers in low-income countries by underpaying them and making them work long hours.

CASE STUDY

Privatization makes life harder for the garbage sifter

The world is producing ever more waste, with serious health and environmental consequences. In urban areas, domestic waste is accumulating fast and landfills fill up quickly. The informal recycling sector of materials (such as metals, plastics and paper) is a way of making a living for up to 24 million people in the global South, according to the International Labour Organization (ILO). However, waste – which was once freely available to the poor – is increasingly being appropriated for business purposes. The policy shift towards privatizing of waste management automatically limits waste pickers' access to recyclable materials at landfill sites. As a result, private corporations obtain large profits, while waste pickers lose their livelihoods. Meanwhile society at large loses the environmental benefits of recycling.

Figure 23.11 Child separating recycled garbage

Source: https://www.usnews.com/news/cities/articles/2020-03-02/life-is-getting-even-harder-for-garbage-sifters-in-the-developing-world

The impact of income and wealth inequality (AO3)

The impact of income and wealth inequalities can be examined with reference to economic growth, standards of living and social stability.

The impact of income and wealth inequality on economic growth

The impact of income and wealth inequality on economic growth is debatable. Research suggests an increase in income and wealth inequality can create incentives to work and invest. This can lead to improved economic performance, more spending on human capital and an increase in risk-taking. These growth-promoting effects enable economically poor countries to achieve a higher rate of economic growth.

The following are reasons why greater income inequality may be beneficial for a country:

- High income and wealth inequality bring the possibility of higher income for those who work hard, which can create incentives for people to work harder thereby improving labour effort.
- A prospect of earning higher incomes can encourage people to invest in their own education and skills development, thereby improving labour productivity.
- Entrepreneurial instincts are encouraged as a result of the potential to earn higher profits. For example, firms may increase their investments in research and development and be more innovative.
- Greater incentives and wealth creation can lead to a higher savings ratio. The increased savings ratio in the country can be used to fund investments, which generates an increase in the economy's long-term growth and development.

On the other hand, an increase in income and wealth inequality can dampen economic growth prospects for some countries. An increase in income and wealth inequality can result in demotivating incentives (disincentives), greater social tensions and increased political tensions.

The following are reasons why greater income inequality may be detrimental to a country's economic growth:

- Greater income and wealth inequalities create more social tensions in the form of demonstrations, protests, political unrest and crime, which could lead to less investment and labour participation rates.
- Greater inequalities usually lead to governments spending more on transfer payments in the form of unemployment benefits, employment generation schemes and other relevant expenditures to sustain the economy. However, transfer payments add to the burden of government debts and do not directly add to the national income of the country.
- Inequalities can discourage workers from joining the labour force and entrepreneurs of small and medium-sized businesses from investing in the economy. People are more likely to be voluntarily unemployed if there are large inequalities in incomes and a good welfare system that supports the unemployed. The inability of smaller firms to compete with large multinationals also causes demotivation for these entrepreneurs.

The impact of income and wealth inequality on standards of living

The **standard of living** in an economy indicates the extent to which people have the ability to lead a life that includes, but is not limited to, personal wealth and material comfort derived from having access to goods and services. Economists tend to use real gross domestic product (GDP) per capita to measure the standard of living. This is often equated to the economic well-being of the society. In general, the ability to access more goods and services increases an individual's economic well-being and utility from consumption, thereby increasing their standard of living. However, there are other quantitative and qualitative factors that also influence the degree of economic well-being in society, such as the quality or degree of political and social stability, religious freedom, environmental quality, safety and security, shelter, longevity, numeracy and literacy rates, and the freedom of speech.

◆ The **standard of living** in an economy indicates the extent to which people have the ability to lead a life that includes, but is not limited to, personal wealth and material comfort derived from having access to goods and services.

Figure 23.12 The quality of housing is a key determinant of standards of living

With an increase in real disposable income, individuals and households can increase their overall consumption. However, the wealthier members of society will be able to afford a disproportionately larger amount on education and healthcare (with long-term impacts on human capital and earning potential) and luxury goods and services (which tend to improve their standards of living). Over time, this unequal level of spending leads to further increases in income and wealth inequalities.

The impact of income and wealth inequality on social stability

Researchers mostly agree that there is a positive correlation between an equal society and the ability of its people to live congruently. Citizens in some of the most equal countries, like Iceland, Denmark, Norway and Canada, tend to live harmoniously. This means people tend to trust each other (and their government), communities involve citizens in a more meaningful way, people have mutual empathy and tend to foster an environment of mutual respect, goodwill, fellowship and social connectedness.

By contrast, the most unequal countries are characterized by political and social unrest, which negatively impacts on the economy. Nobel laureate economist Joseph Stiglitz (b.1943) asserted that economic inequalities prompt mutual distrust of business and government. For example, the Yellow Vest Movement (or Yellow Jacket Movement), which began in France in October 2018, was the result of discontent with tax reform policies that had caused widening income and wealth gaps and caused higher costs of living for low- and middle-income households. This led to years of social altercations and public riots, communal disharmony and criminal activities. In general, unequal countries are often seen as undesirable places to live and travel by expatriate workers and families, and unattractive locations for businesses to invest in.

An unequal society, caused by large income and wealth gaps, also tends to have higher rates of crime and homicide (see Table 23.8). All the countries listed in Table 23.8 score very low in terms of income and gender inequality. There is a direct relationship between income and wealth inequality and social stability (including crime rates), although other factors might also affect this like gun control laws. The homicide rate is measured as the number of murders per 100,000 inhabitants of the country.

Table 23.8 Countries with the highest homicide rates

Rank	Countries	Homicide rate
1	El Salvador	61.80
2	Jamaica	57.0
3	Venezuela	56.33
4	US Virgin Islands	49.26
5	Honduras	41.70
6	Lesotho	41.25
7	Belize	37.90
8	Saint Vincent and the Grenadines	36.46
9	South Africa	35.90
10	Saint Kitts and Nevis	34.24

Source: **https://dataunodc.un.org/GSH_app**, *United Nations Office on Drugs and Crime, 2019*

The role of taxation in reducing poverty, income and wealth inequalities (AO3, AO4)

One of the most important instruments a government uses to redistribute income and wealth is taxation. Policymakers continually deal with the issues of equity and equality – should the wealthiest members of society be taxed more to redistribute to the poor, or should people be taxed less to create better incentives in society? While taxation provides necessary finance to fund government expenditure (see Chapter 25), it puts extra pressure on taxpayers and can cause disincentives to work.

Whether a tax is a 'good' tax is purely contextual and depends on the economic, social and political situation of the country. Adam Smith in his seminal book, *The Wealth of Nations* (1776) indicated that the tax structure of a country must have the following characteristics in order to be effective:

- Certainty – taxpayers must know what and when they need to pay, that is, the amount and the deadline of the tax due should be unquestionably clear, thereby limiting late payments and the number of tax evaders.
- Convenience – the procedures and processes of paying taxes should be simple, that is, the methods and timing of tax payments should be made as easy as possible for people to pay.
- Economic – taxes should be cost-effective for the authorities to collect, that is, the cost of collecting the tax should be a relatively small proportion of the tax yield.
- Equitable – taxes should be levied according to the individual's ability to pay, otherwise they create huge disincentive effects.

Governments can use any of the following methods, funded by tax revenues, to reduce the level of inequalities in the economy:

- Tax revenues enable the government to directly provide socially desirable goods and services, making these merit goods available to everyone, such as public education and healthcare services. Alternatively, the government may choose to use subsidies to enable firms to provide these merit goods and services.
- The provision of essential infrastructure (such as sanitation and clean water supplies) helps to promote equity to those on low incomes and to help eradicate absolute poverty.
- Transfer payments can be used to reduce income and wealth inequalities by making payments to the less affluent without any corresponding exchange or change in national output. Examples include:
  - Retirement pensions (state pensions) – this is an income support scheme for elderly people beyond retirement age. The transfer payment is paid monthly to support this vulnerable group who may not otherwise have any other form of money.
  - Unemployment benefits – these are transfer payments made to unemployed people to enable them to meet their basic human needs, while they actively seek employment opportunities.
  - Child allowances (also known as child benefits) – these are welfare payments made to parents or guardians of children and teenagers in full-time education, for spending on items such as food, clothing, schooling, transportation and health check-ups.

Figure 23.13 Transfer payments are made to provide healthcare services for the general public

Progressive, regressive and proportional taxes (AO3, AO4)

Taxes can be categorized as being progressive, regressive or proportional in nature. These have varying impacts on the redistribution of income and wealth.

Progressive taxes

A **progressive tax** on income means that earning higher incomes will result in a higher percentage of tax paid, which means an increasing proportion of the income is paid in taxes. Figure 23.14 shows that as the level of an individual's income increases from Y_1 to Y_2, the rate of tax rises from T_1 to T_2.

Most countries use a progressive tax system as this is an effective way to address inequalities, by redistributing income and wealth from the rich to improve the economic well-being of the poor. In addition, such a system includes a **tax threshold** (or **tax allowance**) that enables all workers to earn a certain amount of income per year before they become liable for tax. This helps to avoid taxing the economically poor, so that their employment income helps to meet their daily needs.

◆ A **progressive tax** charges a higher *percentage* tax as an individual's income rises, that is, those who earn more pay a greater proportion of their income in tax

◆ The **tax threshold** (or **tax allowance**) refers to the amount of income that workers can earn per year before they become liable for tax.

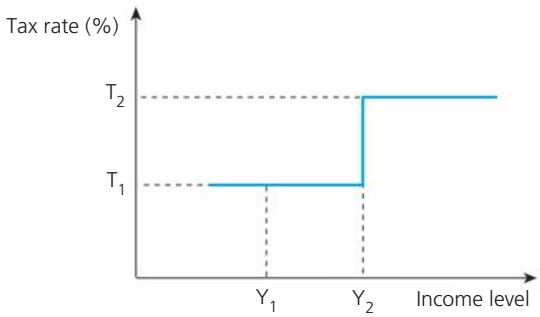


Figure 23.14 Progressive taxes

Common mistake

Be careful not to incorrectly define progressive taxes as a policy whereby individuals pay more taxes as their income level rises. This is true for proportional taxes too, where the more you earn, the more you pay in taxation. Instead, progressive taxation imposes a larger marginal tax rate as incomes rise, which is where the more you earn, the greater the average rate of tax paid as a *percentage* of the individual's total income.

For example, suppose a country does not tax individuals with earnings up to $12,000 per year. This amount is called the tax threshold. If a person in this country earns $20,000 per year, then that person will be taxed only on $8,000 (the amount over and above the tax allowance). The $8,000 that is liable for tax is known as the **taxable income**. So, if the tax rate is 10 per cent of incomes above the tax threshold, the individual pays 10% × $8,000 = $800 per year.

◆ **Taxable income** is the amount of income that an individual or firm is liable to pay tax on, after taxable allowances have been accounted for.

The progressive tax rates (or progressive tax brackets) for a single person in the USA are shown in Table 23.9.

■ **Table 23.9** Progressive tax rates (USA), 2019–20

Tax rate	Single
10%	$0 to $9,700
12%	$9,701 to $39,475
22%	$39,476 to $84,200
24%	$84,201 to $160,725
32%	$160,726 to $204,100
35%	$204,101 to $510,300
37%	$510,301 or more

Top tip!

If you are an HL student, you need to be able to calculate the total tax paid from a set of data. See exam practice questions in the remainder of this chapter.

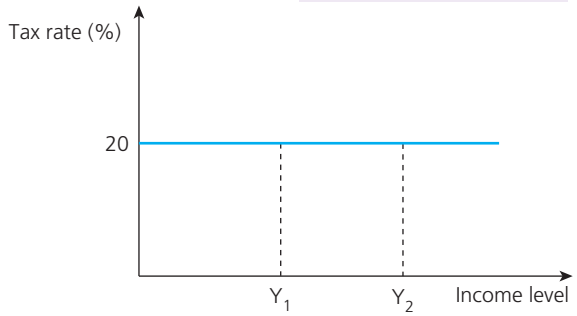
Note that progressive taxes impose different rates of tax as incomes for an individual rises, but is not applied on the total income. For example, someone who earns $80,000 per year will be taxed as follows:

- The first $9,700 is taxed at 10%, that is, **$970** in tax
- The next tier of tax is 12% on total income up to $39,475, that is, $29,775 is taxable income, which equates to **$3,573** (that is, $29,775 × 12%)
- The remaining $40,525 of earnings is taxed at 22%, which equates to **$8,915.50** (that is, $40,525 × 22%)
- Therefore, on personal income of $80,000 in a year, the person pays total income tax of $970 + $3,573 + $8,915.50 = **$13,458.50**

◆ A **proportional tax** (or a **flat rate tax**) charges the same percentage rate of tax, irrespective of the taxpayer's level of income, wealth or profits.

Proportional taxes

Under a system of **proportional taxes**, the percentage of tax paid stays the same, irrespective of the taxpayer's level of income, wealth or profits (see Figure 23.15). Examples include capital gains taxes (levied when an asset is sold for more than it was bought for) and sales taxes such as value added tax (VAT) and goods and services tax (GST) by using a fixed percentage rate of tax. For example, Denmark has a 25 per cent GST whereas sales taxes in India and Japan are as low as 5 per cent. In Figure 23.15, irrespective of the income level (Y_1 or Y_2), the same rate of tax is charged, such as the 20 per cent VAT in Austria, Morocco, Senegal and the UK. As the proportional tax



■ **Figure 23.15** Proportional taxes

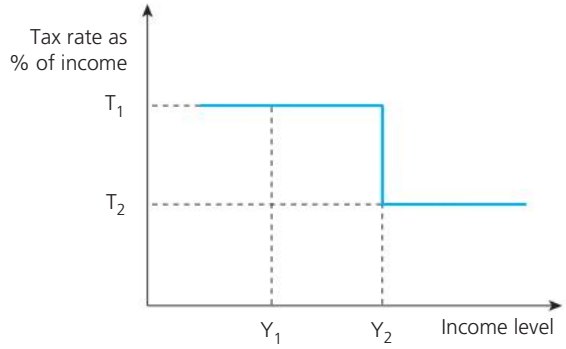
is charged at a flat rate for everyone, whether they are high or low earners, progressive taxes are also called **flat rate taxes**. The underlying basis for charging proportional taxes is the idea that everyone is equal, so taxes should also be charged the same way.

Since the rate of tax is unchanged, there is no marginal tax effect associated with a rise in income. Hence, there is no change in the average tax rate as a person's income increases. For example, if the tax rate is set at 20 per cent, an individual who earns $10,000 will pay $2,000 in tax whereas a person who earns $50,000 will pay $10,000 in tax. Under a proportional income tax system, all taxpayers pay the same fixed percentage rate on their annual taxable income, regardless of the amount they earn. This means the fixed rate does change if the level of income changes.

◆ A **regressive tax** applies a uniform tax rate to all taxpayers, thereby taking a larger percentage of income from low-income earners than from high-income earners.

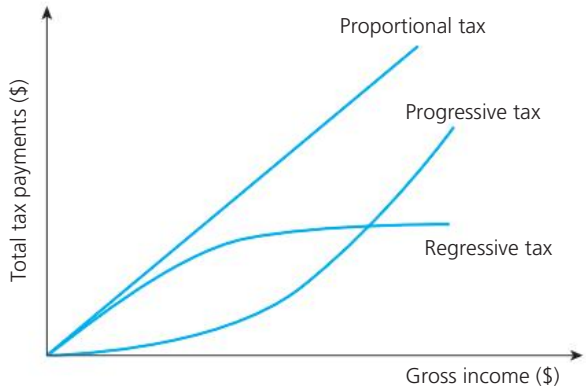
Regressive taxes

Under a system of **regressive taxes**, those with a higher ability to pay are actually charged a lower rate of tax, that is, the wealthier the individual, the lower the tax rate paid as a percentage of their income level (see Figure 23.16). For example, although a high-income earner pays the same absolute amount of airport tax or road licence fee (to drive on public roads and highways) as a less wealthy person does, the value of the tax paid is a smaller proportion of the wealthier person's income. Hence, under this system, the rate of tax decreases as the amount subject to taxation increases.



■ **Figure 23.16** Regressive taxes

With a specific amount of tax paid, high-income earners (those at Y_2) pay a smaller proportion of tax (T_2) compared to low-income earners (those on Y_1 who pay T_1), as shown in Figure 23.16. This means that the average rate of taxes paid by a person goes down with an increase in income. However, this also means that lower income earners are hit harder if a government imposes indirect taxes. For example, taxes on demerit goods such as cigarettes and alcohol take a greater toll on low-income individuals than on high-income earners.



■ **Figure 23.17** Total tax payments under different tax systems

Figure 23.17 shows total tax payments in dollars for the different tax systems given an increase in gross incomes (that is, incomes before tax deductions). Although proportional tax yields more total tax payments as gross incomes rise, the average rate of tax remains the same.

Table 23.10 shows the tax burdens for different type of tax systems for three individual citizens of the country.

■ **Table 23.10** Tax payments under progressive, regressive and proportional tax systems

Tax payments based on different income levels, per year						
	$20, 000		**$50, 000**		**$100, 000**	
	Tax paid ($)	**Average tax rate (%)**	**Tax paid ($)**	**Average tax rate (%)**	**Tax paid ($)**	**Average tax rate (%)**
Regressive tax (flat rate tax of *$2,000*)	*$*2,000	10%	*$*2,000	4%	*$*2,000	2%
Proportional tax (10%)	*$*2,000	10%	*$*5,000	10%	*$*10,000	10%
Progressive tax (10%, 15%, 20%)	*$*2,000	10%	*$*7,500	15%	*$*20,000	20%

Top tip!

Proportional taxes are a type of regressive tax because the tax rate is fixed so places a higher financial burden on low-income earners. For example, flat rate sales taxes are also considered to be regressive because the tax carries a larger impact on low-income earners.

PAPER 3 EXAM PRACTICE QUESTION 23.3 (HL ONLY)

Study the data below and answer the questions that follow.

Income ($ per year)	Tax paid per year ($)		
	Tax A	Tax B	Tax C
10,000	1,000	1,000	1,000
15,000	1,000	1,800	1,500
20,000	1,000	3,000	2,000
25,000	1,000	4,500	2,500

a Identify the tax (A, B or C) that is progressive. *[1 mark]*

b Identify the tax (A, B or C) that is proportional. *[1 mark]*

c Distinguish between a *regressive tax* and a *proportional tax*. *[2 marks]*

PAPER 3 EXAM PRACTICE QUESTION 23.4 (HL ONLY)

Suppose a country taxes an individual $5,500 on earnings of $55,000 a year and charges another person $1,500 on earnings of $15,000 a year. Explain what type of tax system the country uses. *[2 marks]*

Average and marginal tax rates (AO3, AO4) (HL only)

The role of taxation in reducing poverty and inequalities in income and wealth can also be examined by using the principles of the average and marginal rates of tax. The **marginal tax rate** (MTR) refers to the percentage of direct tax paid on the last dollar of an individual's income, that is, the change in tax rate paid from a given change in income. It is calculated by using the formula $\Delta T \div \Delta Y$, where T refers to the tax paid and Y is gross income. The **average tax rate** (ATR) refers to the total amount of tax paid as a proportion of the amount of income earned. It is calculated by the total tax paid to the government divided by the total gross income of the individual, that is, $T \div Y$.

PAPER 3 EXAM PRACTICE QUESTION 23.5 (HL ONLY)

Suppose in a country the progressive tax rates are 10% (for those earning between $10,001 and $40,000 per year) and 15% (for those earning over $40,000 per year).

a Use the table below to calculate the total amount of tax paid by an individual who earns $75,000 per year. *[4 marks]*

Income level ($)	Tax rate (%)	Amount of tax paid ($)
Up to $10,000	0%	
$10,001–$40,000	10%	
$40,001 and above	15%	
Total tax paid:		

b Using your answers from above, show how the average tax rate paid by the individual is 11%. *[2 marks]*

Top tip!

When making or evaluating recommendations to reduce income inequalities, remember that every government decision has an opportunity cost. For example, redistributing income and wealth by using progressive taxes involves administrative costs to the government.

◆ The **marginal tax rate** (MTR) refers to the percentage of tax paid on the last dollar of an individual's income, that is, the change in tax rate paid from a given change in income.

◆ The **average rate of tax** refers to the amount of tax paid compared to the amount of income earned, that is, the total tax paid divided by the total income for an individual, or $T \div Y$.

PAPER 3 EXAM PRACTICE QUESTION 23.6 (HL ONLY)

Use the data below to answer the following questions for an individual who earns $40,000 per year.

a Identify the marginal rate of income tax. *[1 mark]*

b Calculate the total amount of tax paid by the individual. *[2 marks]*

c Calculate the average rate of income tax. *[2 marks]*

d Calculate the average rate of tax for another individual who earns $80,000 per year. *[3 marks]*

Income tier	Tax rate (%)
First $10,000	0
Next $20,000	10
Next $20,000	20
Thereafter	30

PAPER 3 EXAM PRACTICE QUESTION 23.7 (HL ONLY)

Calculate the average tax rate of individuals who earn $50,000 per year in a country that charges a flat rate 20% income tax and grants individuals with a tax allowance of $15,000. *[2 marks]*

PAPER 3 EXAM PRACTICE QUESTION 23.8

Calculate the change in the amount of tax paid if an individual faces a marginal tax rate of 45% when her income changes from $47,000 to $55,000 per year. *[2 marks]*

PAPER 3 EXAM PRACTICE QUESTION 23.9 (HL ONLY)

From the data below, calculate the tax paid by an individual who earns $35,000 a year. *[3 marks]*

Income tier	Tax rate (%)
$9,000	0
$9,001–$20,000	10
$20,001–$30,000	20
$30,001 and above	30

Top tip!

If you are a HL student, you should remember under a regressive tax system, the average tax rate exceeds the marginal tax rate. By contrast, in a progressive tax system, the marginal tax rate exceeds the average tax rate.

Direct taxes (AO3, AO4)

The use of direct and indirect taxes can also be implemented to reduce poverty and inequalities in income and wealth. A **direct tax** is a levy imposed on income (such as wages, salaries and profits), rather than on expenditure. As the name suggests, it is imposed directly on taxpayers who are liable to pay the tax. For example, personal income tax is levied on an individual's incomes, such as wages, salaries, interest, rent and dividends. Businesses are also liable for corporation tax, a direct tax on their profits. These taxes help to ensure that wealthy individuals, households and firms contribute to the wider community through paying their fair share of taxes.

◆ A **direct tax** is a levy imposed on income, rather than on expenditure. It is imposed directly on the person or firm liable to pay the tax.

Direct taxes that can be used to reduce the impact of poverty and inequalities include taxing personal income, corporate income and wealth.

Personal income taxes

Table 23.11 shows the countries with the highest personal income tax rates. Notice that all these countries are economically developed nations. Many of these countries are known for the quality of their provision of public services, healthcare and top-class education systems, all of which need to be funded by tax revenues.

Table 23.11 Countries with the highest rates of personal income tax

Country	Highest marginal tax rate (%)
Aruba	59.0
Sweden	57.0
Finland	56.0
Japan	55.95
Denmark	55.85
Austria	55.0
Canada	54.0
Luxembourg	52.45
Ireland	52.0
Netherlands	51.95

Source: https://www.heritage.org/index/explore?view=by-variables

However, it is important to note that governments should not overtax as this creates disincentive effects in labour markets and deters domestic and foreign businesses from investing in the economy. For example, if the tax threshold is minimal and tax bands (or tax brackets) are too high, this can put off some people from seeking employment. This is because these people might even experience a financial loss from working, due to forgone welfare benefits and the liability of direct taxes once they start earning a wage or salary. Hence, the use of high direct taxes would act as a deterrent for the unemployed to join the workforce, especially if they currently receive state benefits. Furthermore, overly high direct taxes encourage tax avoidance or even tax evasion owing to an increase in unreported incomes and transactions taking place in unofficial or parallel markets, that is, there is a greater incentive for workers to avoid reporting tax liabilities to the tax authorities, thus possibly leading to widespread corruption. This leads to the country undervaluing or underreporting its national income (see Table 23.12).

◆ **Corporate income tax** (or **corporation tax**) is a direct tax imposed on the annual profits of private sector businesses.

Corporate income taxes

Corporate income refers to the profits that private sector businesses earn in a year. **Corporate income tax** (also called **corporation tax**) is a direct tax imposed on the annual profits of firms operating in the private sector of the economy.

Table 23.12 Countries with the highest tax avoidance (tax revenue loss as a percentage of GDP)

Country	Revenue loss as a % of GDP
Guyana	6.97
Chad	6.97
Malta	4.59
Comoros	4.42
Guinea	4.42
Zambia	4.42
Pakistan	4.42
Argentina	4.42
Eritrea	3.69
Namibia	3.69

Source: https://www.weforum.org/agenda/2017/04/which-countries-are-worst-affected-by-tax-avoidance/

Businesses and policymakers often disagree on the optimum level of corporation tax. Firms argue for lower rates of corporate income tax so that the burden of the tax does not fall heavily on the owners and shareholders. This would improve the profitability of businesses, thus enabling them to hire more workers and invest a greater amount in the long-term. By contrast, higher rates of corporate taxes ensure a more effective redistribution of income and wealth in a country (see Table 23.13).

As with personal income tax, overtaxing company profits also creates disincentives and can cause tax avoidance or even encourage tax evasion and corruption. Data from the World Economic Forum (WEF) show that governments lose hundreds of billions of dollars each year from unethical business behaviour in an attempt to reduce corporation tax liability (see Table 23.14). Hence, collecting corporate taxes in an effective way requires good governance, policy reforms and flexibility. This would then help to ensure a better reallocation of income and wealth in the country.

■ **Table 23.13** Countries with the highest corporate taxes, 2019

Country	Corporate income tax rate (%)
France	32.00
Portugal	31.50
Australia	30.00
Mexico	30.00
Germany	29.90
Japan	29.70
Belgium	29.60
Greece	28.00
New Zealand	28.00
South Korea	27.50

Source: https://stats.oecd.org/

■ **Table 23.14** Countries with the highest corporate tax losses due to tax avoidance

Country	Estimated annual corporate tax losses ($bn)
USA	188.8
China	66.8
Japan	46.8
India	41.2
France	19.8
Germany	15.0
Australia	6.1
Spain	5.5
South Korea	1.1
United Kingdom	1.1

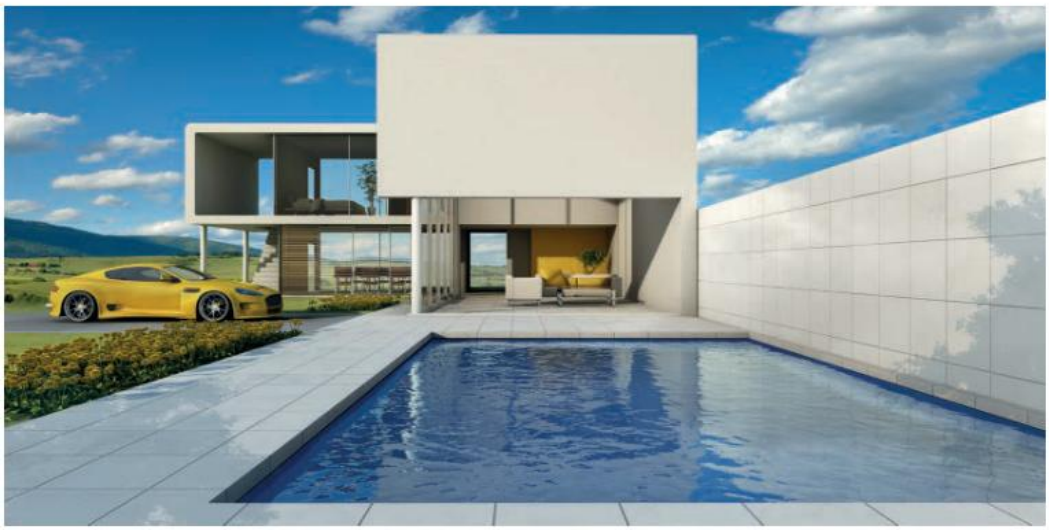
Source: https://www.weforum.org/agenda/2017/04/which-countries-are-worst-affected-by-tax-avoidance/

Wealth taxes

A third and final type of direct tax that can be used to tackle poverty and reduce inequalities in income and wealth is the use of wealth taxes, such as capital gains taxes. A **wealth tax** is a direct tax based on the market value of assets owned by an individual, household or firm.

◆ A **wealth tax** is direct tax based on the market value of assets owned by an individual, household or firm.

In the USA, the Federal Reserve Board published data in 2018 that indicated the wealthiest 10 per cent of people in the country owned 70 per cent of the nation's wealth, while the richest 1 per cent owned 32 per cent. Hence, there is a strong case for using wealth taxes to minimize the income inequality gap and to redistribute income and wealth in order to reduce the poverty level.



■ **Figure 23.18** Wealth taxes are used to redistribute income and wealth

For example, some governments impose a 'vacant tax' on second homes and properties owned by individuals that are not owner-occupied. Other assets that may be liable for wealth taxes include inheritance, pension plans and investment funds. To be more equitable, governments usually impose wealth taxes based on a person's net worth, measured as the difference between the value of the individual's assets and liabilities (such as the outstanding mortgage on real estate assets).

Theory of Knowledge

To what extent should people of a country have the right to know where and how their share of tax revenues is spent? How far should governments be required to justify their spending with evidence and reasoning?

Indirect taxes (AO3, AO4)

In addition to the use of progressive, regressive and proportional taxes and direct taxes, indirect taxes can also be used to reduce poverty and inequalities in income and wealth. An indirect tax (also known as an expenditure tax) is a government levy on consumption, rather than being imposed on incomes. Examples of indirect taxes include Goods and Services Tax (GST), customs duties (imposed on imported products) and excise duties (imposed on demerit goods).

Figure 23.19 Duties are imposed on petroleum as a form of expenditure tax paid by drivers

It can be effective in that wealthier people spend more money, so end up paying a greater value of indirect taxes. For example, stamp duty is an indirect tax paid on the sale of commercial or residential property. In many countries, stamp duty is also progressive, that is, the more expensive the purchase of a property, the higher the rate of indirect tax. Hence, indirect taxes placed on the purchase of expensive items can help to redistribute income and wealth in the country. In addition, major advantages of using indirect taxes are that they are easier to collect and harder to evade.

However, as indirect taxes – and sales taxes in particular – are applied to everyone across the country irrespective of their level of income, indirect taxes are often criticized for being regressive. A person with a low income pays the same amount of tax as a person with a higher income for the purchase of cigarettes, alcohol or fuel. The poor would be disadvantaged as they need to spend a greater proportion of their earnings compared to wealthier people. Therefore, indirect taxes can further widen income inequalities.

Further policies to reduce poverty, income and wealth inequality (AO3)

Governments across the world, with the help of independent policymaking experts and consultants, continually work on designing policies that cater for the needs of their respective countries to address the short-term and long-term issues related to poverty and income and wealth inequalities. Apart from using taxation policies, other policies can be used to reduce poverty, and income and wealth inequalities. These measures include policies to reduce inequalities of opportunities and investment in human capital, transfer payments, targeted spending on goods and services, universal basic income, policies to reduce discrimination, and the use of minimum wages.

Policies to reduce inequalities of opportunities/investment in human capital

Professor John A Powell (b.1947), the director of the Othering & Belonging Institute at UC Berkeley, suggests several evidence-based policy solutions to reduce inequalities of opportunities. These policies include:

- Increase the minimum wage in a country, which could lead to the eradication of poverty owing to higher incomes for low-waged labourers. This could also help to improve their household's access to education and healthcare.
- In many countries, the Earned Income Tax Credit (EITC) scheme gives tax benefits for working people with low to moderate income, who have children. The EITC scheme reduces the amount of tax these households pay, and some households even qualify for a tax refund. This has been an effective tool to increase income and reduce poverty. Policymakers suggest that other countries could also adopt such a policy, offering tax credits to low- and middle-income families who do not have any children.
- An investment in early education could help to facilitate labour productivity and economic mobility in the long-term. Essentially, investments in the country's human capital contribute to reducing or removing inequalities of economic opportunities.
- Appropriate tax reforms including more progressive taxes could lead to higher government tax revenues without hurting lower income families. The extra revenue would be used to ensure a reduction in income and wealth inequalities, and greater access to economic opportunities for lower- and middle-income families.
- Policies to encourage housing for low-income families could lower residential and social segregation by income and race, which can boost economic mobility for all members of society. Again, this helps to reduce inequality of opportunities.

Transfer payments

A **transfer payment** is a sum of money from the government to another party for which no goods or services are being paid, that is, there is no corresponding output or economic activity for such payments. Examples include unemployment benefits, disability allowances, subsidies to domestic producers, student loans on favourable terms (low interest rates and delayed repayments), pension payments for elderly people who have retired and housing allowance for low-income families.

◆ A **transfer payment** is a sum of money from the government to another party for which no goods or services are being paid, that is, there is no corresponding output or economic activity for such payments.

Furthermore, government spending on merit goods, such as healthcare and education, create more economic opportunities for low-income families. Without government support, these merit goods would be under-consumed and under-provided. Hence, spending on such merit goods funded by tax revenues would increase the economic welfare of the country in the long-term.

However, the use of transfer payments also has its drawbacks. Government expenditure of transfer payments is costly and can drain their limited budgets. Austerity measures following the global financial crisis of 2008 caused severe economic problems, even for high-income countries such as Portugal, Ireland, Greece and Spain. They can also cause laziness among the workforce and an overreliance on government bailouts. There are also opportunity costs in the direct provision of goods and services aimed at redistributing income and wealth. For example, subsidies, apart from being costly to taxpayers, can encourage firms to become overly reliant on government funding, thus hindering economic efficiency and innovation. In reality, low- and middle-income countries struggle to raise sufficient tax revenues to fund the necessary amount of transfer payments to improve economic opportunities.

Targeted spending on goods and services

Governments often budget for targeted spending on certain goods and services in the country. Targeted spending on goods and services is argued as one of the effective ways to reach out to the individuals and groups who need the requisite support. Examples of targeted spending include the budgeted expenditure on preprimary and primary level education, aimed at attaining full child literacy.

Another example is the use of conditional cash transfers (CCTs). CCT programmes give money to low-income families on the condition that they comply with certain requirements to break the cycle of poverty for the next generation through the development of human capital. These conditions may include regular visits to a healthcare facility, getting vaccinations and regular school attendance by children. Hence, as with all forms of targeted spending, CCTs offer greater economic opportunities to the poor and disadvantaged.

However, targeted spending on goods and services is hard to allocate because of the many multiple and competing needs. This is further complicated because of limited government funds, especially in low-income countries that naturally have fewer financial resources.

Figure 23.20 Targeted spending on education can improve living standards in a country

This means governments tend to allocate only a limited amount of their budgets to targeted spending on specific goods and services, as their priority will be on macroeconomic objectives such as growth, employment and price stability.

Figure 23.21 Dr Martin Luther King Jr

Universal basic income (UBI)

Nobel Peace Prize winner and civil rights visionary Dr Martin Luther King Jr (1929–68) proposed a guaranteed income to the poor and unprivileged for a complete eradication of poverty. A **universal basic income** (UBI) refers to the guaranteed minimum level of income that a government guarantees to each individual of the country. In many societies, the increasing use of automation in production processes means that menial jobs have been replaced by robots or technology-driven mechanisms. The loss of jobs leads to a loss in income, which is a precursor of poverty and income inequality. In order to prevent this from happening, governments intervene to offer everyone in the country a UBI.

◆ A **universal basic income** (UBI) refers to the guaranteed minimum level of income that a government guarantees to each individual of the country.

American economist and Nobel Prize winner in Economic Sciences in 1976, Milton Friedman (1912–2006) argued for a negative income tax, where tax credits are provided to the poor if their income falls below a certain minimum level. This would be funded by the tax collected from wealthier individuals in the country.

There are certain advantages that the provision of a universal basic income could offer. Since the provision of a UBI is quite straightforward, in theory, the government could administer it quite easily. During a recessionary period, the universal basic income could help cushion the fall in aggregate demand as all households can maintain their standard of living. Whereas traditional welfare systems can create a poverty trap, a universal basic income could help to break it.

However, there are complications to the introduction of a UBI. The economy would suddenly experience a significant rise in the aggregate demand for goods and services if all households were to have a UBI. This leads to a rise in the general price level of goods and services (inflation), which can cause its own problems (see Chapter 20). Furthermore, funding the universal basic income would be costly, so the government may face a budget deficit and mounting government debts. Eradication of poverty and reducing inequalities by providing a UBI in the country may therefore be extremely costly. Also, as the UBI is essentially free income for households, it could easily discourage individuals from taking up a full-time job and contributing towards the economy.

Policies to reduce discrimination

Discrimination refers to any form of social exclusion based on socioeconomic differences such as gender, ethnicity, race, religion, disability or sexual orientation. Discrimination is a major cause of income and wealth inequalities and poverty. To combat such prejudices and economic injustices, governments develop and implement anti-discrimination policies. These laws make all forms of discrimination within the country illegal.

Anti-discrimination legislation therefore helps the country to tackle issues of poverty, social deprivation and unemployment. There are three distinct steps in achieving this:

1. Setting a commonly agreed standard for individual and collective behaviour towards any form of discrimination.
2. Educating the public about the standards expected within a society.
3. Establishing deterrents and consequences for those who violate the standards.

All three steps require government initiatives, intervention and participation. Good intentions of individuals and firms will not be enough to create an institutional framework that supports and encourages non-discrimination. However, many governments have not been proactive active enough to design and enforce policies to reduce discrimination.

Minimum wages

A final method that policymakers can adopt to reduce poverty and income and wealth inequalities is the use of minimum wages. The International Labour Organization (ILO) defines a **minimum wage** as the lowest amount of remuneration that an employer is required to pay wage earners for the work performed during a given period, which cannot be reduced by collective agreement or an individual contract. Setting a minimum wage for a country could be challenging as costs of living differ in different parts of the nation. There would also be a need for regularly reviewing the minimum wage and making adjustments based on changes in the cost of living for the average household in the country.

◆ A **minimum wage** is the lowest rate of pay that firms must pay to their employees per time period, as stipulated by the government.

There are some clear advantages to a minimum wage in eradicating poverty and minimizing the income and wealth gap. A minimum wage would encourage more labour participation, which would ensure less spending on transfer payments by the government over the long-term. This also enables the government to redistribute its resources for achieving other priorities or reduce the burden on taxpayers. For many countries, having a minimum wage law helps to bring greater transparency in the job market as all employers need to comply by paying the minimum wage to all workers. This is particularly the case for workers in low-income countries.

However, the main disadvantage and limitation of using minimum wage legislation to tackle the problems of poverty and inequality is the resulting higher labour costs. A minimum wage could create unemployment as a result, because it forces employers of low-skilled workers to pay more than the market equilibrium wage, for every hour they work. Hence, the minimum wage can disrupt the price mechanism in the labour market and can cause individuals from low-income households to be unable to find employment.

Theory of Knowledge

Under what situations would it be fairer for one person to be paid more than another?

Chapter summary

- Equity differs from equality in the distribution of income. Equity is based on the argument that income inequalities (such as wage differentials) are needed to create incentives for people to study and work harder. The same applies to businesses that strive for greater levels of profit.
- Equality means the equal distribution of income in the economy. Economies face unequitable distribution of income owing to the naturally unequal ownership of factors of production in a free market economy.
- There are various ways to describe the distribution of the income in an economy. There could be differences in the income distribution between countries, between geographical regions of a country, between males and females, between different age groups, ethnicities or occupations (professions).
- Income and wealth can affect one another. It is easier for higher income households to accumulate a greater stock of wealth through savings or the accumulation of assets. The higher level of wealth can then lead to a larger income stream from assets and investments. Low-income households may incur debts (negative wealth), which reduces their income as households have to pay interest on loans.
- The degree of income equality (or inequality) can be measured by the relative share of national income earned by given segments of the population.

- The Lorenz curve shows the degree of income inequality in a country. The Gini coefficient is a numerical representation of a country's Lorenz curve.
- Quintiles are used to divide statistical data into fifths, with each part representing 20 per cent of the population.
- Poverty refers to the state of an individual, household or country being extremely poor, that is, not having enough money to meet their basic human needs, such as food, water, clothing, shelter, healthcare and education.
- Absolute poverty is the lack of access to the basic goods and services needed for survival; for example, food, shelter, clean water, sanitation, healthcare, primary education and information.
- Relative poverty in a country is determined by examining the percentage of the population with earnings less than a predetermined percentage of the median income.
- Single indicators, such as the International Poverty Line or the Minimum Income Standard, and composite indicators, such as the Multidimensional Poverty Index (MPI), are used to measure poverty.
- Along with the income element of the monetary dimension of measuring poverty, there are other crucial non-monetary indicators such as education, healthcare, sanitation, clean water and access to the use of power. These multidimensional factors make it difficult to measure poverty in an accurate and objective way.
- There are different causes of economic inequality and poverty such as inequality of opportunity, different levels of resource ownership, different levels of human capital, discrimination (gender, race and others), unequal status and power, government tax and benefits policies, globalization and technological change, and consequences of market-based supply-side policies.
- Research suggests an increase in income and wealth inequality can lead to stronger performance incentives in some communities. However, it can also act as a growth-dampening agent in other countries due to a widening income and wealth gap.
- One of the most important instruments a government employs for the redistribution of income and wealth is the use of taxation.
- Progressive taxation charges a higher percentage tax as an individual's income rises, that is, those who earn more pay a greater proportion of their income in tax.
- Regressive taxation charges a greater proportion of tax on lower-income earners. For example, sales taxes, such as Value Added Tax (VAT) and Goods and Services Tax (GST), account for a greater proportion of tax paid by the relatively poor.
- Proportional taxation charges the same flat rate percentage tax, irrespective of how much an individual earns. While more tax is paid in absolute terms as an individual's income rises, the percentage tax paid is fixed.
- Governments across the world with the help of independent policymaking experts and consultants have designed a range of policies to tackle the problems of poverty and inequalities. These include the use of minimum wages, investment in early education, tax reforms, transfer payments, targeted government spending, offering a universal basic income and policies to reduce discrimination.

24 Demand management (demand-side policies) – monetary policy

Real-world issue

How do governments manage their economy and how effective are their policies?

Conceptual understandings (WISE ChoICES)

- Government **intervention** attempts to achieve macroeconomic objectives through a **choice** of policies.
- Political, economic, social and environmental factors are **interdependent** and will influence the effectiveness of government policies.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- monetary policy, with reference to control of the money supply and interest rates by the central bank (AO1)
- the goals of monetary policy (AO2)
- the process of money creation by commercial banks (AO2) (HL only)
- the tools of monetary policy (AO2) (HL only)
- the demand and supply of money, with reference to the determination of equilibrium interest rates (AO2) (HL only)
- a diagram to show the determination of equilibrium interest rates (AO4) (HL only)
- the difference between real and nominal interest rates (AO2)
- interest rates from given data (AO2)
- the use of expansionary and contractionary monetary policies to close deflationary/recessionary and inflationary gaps (AO3)
- a diagram using AD–AS curves to show expansionary and contractionary monetary policy (AO4)
- the effectiveness of monetary policy, with reference to constraints on monetary policy and the strengths of monetary policy (AO3)
- the strengths and limitations of monetary policy in promoting growth, low unemployment, and low and stable rates of inflation (AO3).

◆ **Monetary policy** is the use of interest rates and the money supply to influence the level of aggregate demand and economic activity, in order to achieve the country's macroeconomic objectives.

Monetary policy (AO1)

Monetary policy refers to the government's control and use of interest rates and the money supply to influence the level of aggregate demand and economic activity. Like fiscal policy (see Chapter 25), monetary policy as a demand-side policy is discretionary in nature, that is, it is purposively carried out by the government to influence the level of aggregate demand and economic activity. In practice, monetary policy is overseen by the nation's **central bank** or designated monetary authority.

◆ The **central bank** is the monetary authority responsible for regulating the country's financial system and implementing monetary policy, including regulation of commercial banks and the country's money supply.

Interest rates are the price of money. Interest rates can refer to the price of borrowing money (the interest rate charged to borrowers) or the return from saving money (the interest rate paid to savers) at financial institutions, such as commercial banks. A **demand-side policy** refers to any government strategy or plan to influence the level of aggregate demand, such as reducing interest rates to reduce the costs of borrowing money to finance household consumption expenditure (C) and corporate investments (I). The **money supply** refers to the entire quantity of money circulating in an economy, including notes and coins, loans and savings deposits at financial institutions and banks.

◆ The **interest rate** is the price of money. Expressed as a percentage measure, interest rates represent the cost of borrowing money or the return for savers.

◆ A **demand-side policy** refers to any government strategy or plan to influence the level of aggregate demand, such as reducing interest rates to reduce the costs of borrowing money to finance household consumption expenditure (C) and corporate investments (I).

◆ **Money supply** is the amount of money in circulation within the economy at a particular point in time, as determined by the central bank. The money supply comprises legal tender (banknotes and coins), loans, credit, bank deposits and central bank reserves.

Functions of a central bank (monetary authority)

- **Executor of monetary policy** – In most countries, the central bank or the country's monetary authority is responsible for managing interest rates and the exchange rate (for its currency) in order to achieve macroeconomic objectives.
- **Government's bank** – The central bank is responsible for the money of the government, including its foreign currency reserves. It maintains the accounts of the government in the same way that commercial banks maintain the accounts of their private and corporate customers.
- **Bankers' bank** – The central bank is the regulator of the country's commercial banking system. For example, commercial banks must keep a certain percentage of their cash reserves at the central bank so it can control the money supply in circulation and/or use the reserves in times of financial emergencies.
- **Sole issuer of legal tender** – As the supreme bank, the central bank or monetary authority is the sole issuer of legal tender (bank notes and coins) within the country. This helps to control the money supply and brings uniformity and confidence to the monetary system.
- **Lender of last resort** – The central bank provides loans to commercial banks when necessary to prevent the risks of a financial crisis caused by limited cash reserves and liquidity problems. Thus, as lender of last resort, the central monetary authority helps to ensure the banking system runs smoothly.
- **Credit control** – By controlling the cash reserves that commercial banks must hold at the central bank, credit creation is managed more effectively. For example, the central bank could raise the cash reserve ratio during an economic boom to limit over-lending by commercial banks.

Figure 24.1 The European Central Bank is the sole issuer of euros

The central bank will consider a range of factors before making any decision about changing interest rates in the economy. For example, it will consider the current and expected level of economic activity, the level of inflation in the economy, and the outlook for the global economy. When the economy is doing well or expected to grow too rapidly, the central bank will tend to increase interest rates, and vice versa. By controlling the money supply, the central bank can control the level of economic activity.

Goals of monetary policy (AO2)

The goals (objectives) of monetary policy include: achieving a low and stable rate of inflation (inflation targeting), low unemployment, reducing fluctuations in the business cycle (see Chapter 16), promoting a stable economic environment for long-term growth, and harmonizing the external balance.

Low and stable rate of inflation (inflation targeting)

Like fiscal policy, governments use monetary policy to either expand or contract economic activity to achieve their macroeconomic goals. For instance, interest rate policy and manipulation of the money supply can be used to achieve a low and stable rate of inflation.

An **inflation rate target** (or **inflation targeting**) refers to the practice of central banks in some countries (such as Canada, Finland, New Zealand, South Africa and the UK) to use monetary policy to achieve a specific rate of inflation. An inflation target is used to provide a transparent goal in order to help control inflation to enable sustainable economic growth and employment. This is because price stability will enhance consumer and business confidence in the economy.

◆ **Inflation targeting** refers to the practice of central banks using monetary policy to achieve a specific rate of inflation.

Top tip!

Note that central banks use monetary policy to *influence* rather than to directly *determine* the rate of inflation. In reality, interest rates are only one (significant) factor that affects the general price level in the economy.

Low unemployment

It is ultimately the role of the government, rather than a monetary authority or central bank, to focus on achieving and maintaining full employment. However, monetary policy can be used to help achieve low unemployment. Lower interest rates, in theory, should stimulate economic activity by increasing aggregate demand. This is because lower interest rates reduce borrowing costs for households and firms, thereby helping to boost consumption and investment, ceteris paribus. Figure 24.2 shows that lower interest rates encourage consumption spending (C) and investment expenditure (I), boosting aggregate demand from AD_1 to AD_2. As national output increases from Y_1 to Y_2, this should reduce the number of people in unemployment.

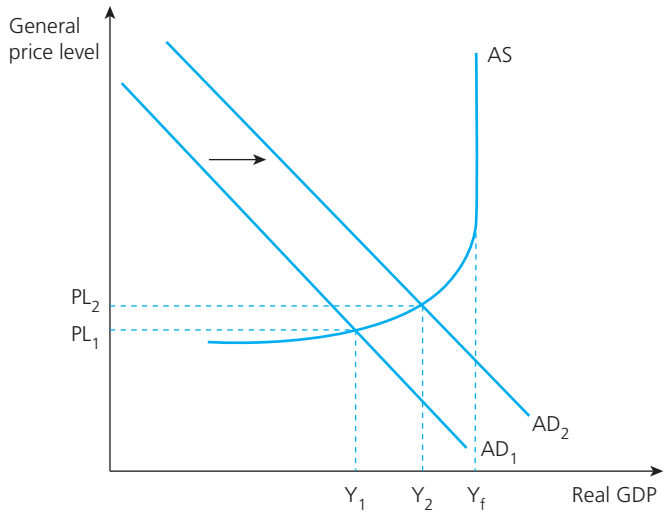


Figure 24.2 Effects of lowering interest rates

Reduce business cycle fluctuations

Monetary policy is a key demand-side policy used to influence the level of economic activity. Lower interest rates can be used during an economic downturn in the business cycle (see Chapter 16), as illustrated in Figure 24.2. By contrast, if the economy is booming, as shown in Figure 24.3, higher interest rates can be used to reduce the impact of inflationary pressures especially if the country is near to or at full employment (Y_f). Higher interest rates, in theory, shift aggregate demand from AD_1 to AD_2, thereby helping to keep inflation under control by reducing the general price level to fall from PL_1 to PL_2.

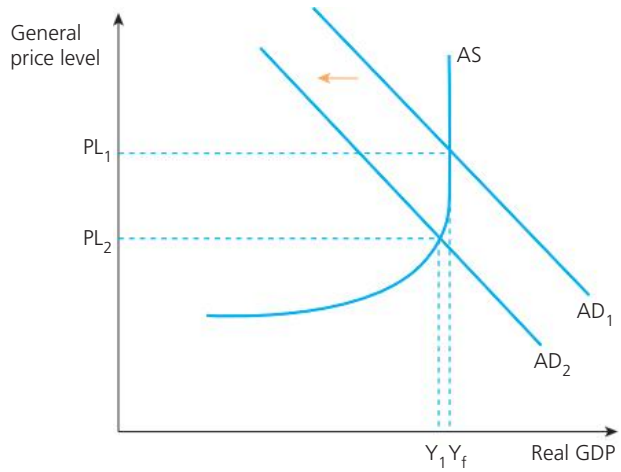


Figure 24.3 Effects of raising interest rates

Promote a stable economic environment for long-term growth

Economic stability, through the use of monetary policies, makes it easier to achieve macroeconomic objectives, such as stable prices, lower unemployment and sustainable economic growth. This is because economic stability creates a greater degree of certainty and confidence for consumers and firms. Hence, by stabilizing fluctuations in the business cycle, effective use of monetary policy helps to encourage investments in physical and human capital for long-term growth of the economy.

External balance

The **external balance** refers to value of a country's export earnings being equal to or approximately equal to the value of its import expenditure. Interest rates can be used to influence the exchange rate (see Chapter 31), thereby affecting the value of a country's imports and exports. For example, lower interest rates tend to make a currency less attractive for foreign buyers (as the rate of return is lower), thereby reducing the exchange rate. This will tend to increase the demand for exports. If export earnings (X) exceed import expenditure (M), then this leads to inflationary pressures in the long run as more money flows into the domestic economy. By contrast, if M > X the country will experience a negative external balance, with a net outflow of money from the economy. As with any individual, countries cannot spend more than they earn in the long run. Hence, an external balance is considered essential for sustainable economic growth and development.

◆ The **external balance** refers to value of a country's export earnings being roughly equal to the value of its import expenditure.

The process of money creation by commercial banks (AO2) (HL only)

One of the key functions of commercial banks is to bring together savers and borrowers in the process called **credit creation**. This refers to the process by which banks create money from deposits of savers and borrowers. Savers will deposit money in banks and gain a rate of interest (or return) from the bank (see Figure 24.4). Banks will then create loans from the savings deposited and lend some of this money to borrowers and charge an interest rate accordingly. The interest rate charged to borrowers (5 per cent, for example) will be greater than the return paid to savers (1 per cent, for example) to ensure the bank earns a profit and also to safeguard the bank from a fall in the real value of loans due to inflation.

◆ **Credit creation** refers to the process by which commercial banks create money from deposits from savers and use these funds as loans to borrowers.

In Figure 24.4, an individual saver deposits $1,000 into the bank. At any point in time, it is unlikely that the saver will need to withdraw all of the money from the account so the bank may decide to lend 90 per cent of this in the form of loans to a borrower. In reality, the central bank or monetary authority may decide how much commercial banks can lend in an attempt to limit growth in the money supply. This requirement from the central monetary authority is called the **minimum reserve ratio**, that is, the 10 per cent that is not used as loans in this particular example.

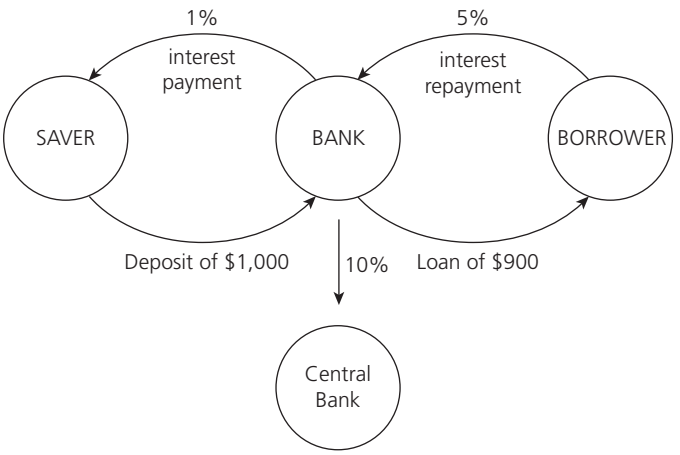


Figure 24.4 Credit creation: the process of money creation

Borrowers use this money for consumption, thereby contributing $900 of spending in the economy, which creates income for another business. If the firm then deposits this money in the bank, there is an increase in the money supply by $900. With this extra $900 deposited by the firm, the bank again retains only 10 per cent and lends the remaining $810 to another borrower, and so the process continues. Hence, deposits generate loans and come back into the banks as a form of further deposits thereby increasing the money supply.

Economists can work out how much the initial deposit increases the money supply by using the **money multiplier** equation:

Money multiplier = 1 ÷ Reserve ratio

In the above example, the money multiplier is 10 (that is, 1 ÷ 0.1). To work out by how much the money supply increases from the initial $1,000 deposit, we need to remember to deduct this amount because the money already existed. So, the initial $1,000 of savings increases the money supply by $9,000 (or creates credit of $9,000):

$1,000 × (1 ÷ 0.1) = $10,000 – $1,000 = $9,000

If the government or central bank feels the need to limit economic activity and suppress inflationary pressures, it can increase the minimum reserve ratio to limit growth in the money supply. For instance, if the reserve ratio is raised to 20 per cent, an initial $1,000 of savings deposited at the bank would now increase the money supply by only $4,000:

$1,000 × (1 ÷ 0.2) = $5,000 – $1,000 = $4,000

◆ The **minimum reserve ratio** (MRR) is the lowest amount that commercial banks are required to keep as reserves in the central bank, thereby limiting how much they can lend and how much credit they can create.

◆ The **money multiplier** is a formula used to calculate by how much an initial deposit (savings at a bank) increases the money supply. The formula is: Money multiplier = 1 ÷ Reserve ratio.

Figure 24.5 Banks increase the money supply through credit creation

Tools of monetary policy (AO2) (HL only)

The IB Economics syllabus specifies four tools of monetary policy, that is, methods used by a central bank or monetary authority to influence the money supply or interest rates. These tools are: open market operations (OMO), minimum reserve requirements (MRR), changes in the central bank minimum lending rate and quantitative easing (QE).

Open market operations (OMO)

Open market operations (OMO) refers to the buying and selling of government securities by a country's central bank. Government securities are a type of public sector debt, in order to finance the government's operations. These securities include the sale of bonds issued by a government, with the promise to repay borrowed money from investors who also receive a fixed rate of interest during the credit period. Since the securities are backed by the government that issues them, they are usually regarded as one of the safest (lowest risk) investments.

When the central bank wants or needs the money supply to fall, in order to slow inflation and/or stabilize economic growth, it can sell government securities to banks or other investors. The higher interest rate (the return paid to purchases of the securities) attracts buyers or investors. This is an example of contractionary monetary policy because it withdraws money from the economy as buyers of the government securities opt not to spend this money on consumption or investment.

By contrast, the central bank may need or want to increase the money supply to lower the price of money – or the interest rate – in order to stimulate higher aggregate demand in the economy. It can do this by purchasing government securities. This would be an example of an expansionary monetary policy as this injects money into the economy, thereby helping to lower unemployment and encourage economic growth.

In summary, OMOs occur when the central bank buys or sells government securities in order to influence interest rates. It either injects money into the economy to stimulate economic activity or withdraws money from the economy in order to reduce inflation. Hence, OMOs are a key tool used by the central banks across the world in the implementation of monetary policy.

Minimum reserve requirements

As explained in the section about credit creation above, banks would generally want to lend as much of their deposits as possible in order to increase their profits. However, the central bank in most countries requires commercial banks to keep a certain percentage of their deposits at the central bank – the amount of which is called the minimum reserve ratio or minimum reserve requirement (MRR). The MRR acts to ensure that commercial banks have enough cash on hand for their daily transactions, given that most customers will not want to withdraw all of their cash deposits on any given day – technically called a *bank run*. These reserve funds held by commercial banks must be kept at the central bank (recall that one of the key roles of a central bank is that it is the bankers' bank). Unless there is a bank run, the MRR ensures that there is enough liquidity or money in the economy to cover the daily operations of commercial banks.

As explained above, there is an inverse relationship between the MRR and the money multiplier. This means, the higher the MRR, the lower the money multiplier and therefore the lower the amount of money that is created by the banking system. Hence, raising the MRR as a form of contractionary monetary policy enables the central bank or government to limit growth in the money supply in order to combat inflationary pressures. This is because a reduction in the money supply causes interest rates to increase, thereby lowering consumption and investment expenditure. By contrast, the central bank can reduce the MRR as a form of expansionary

Theory of Knowledge

The process of money creation by commercial banks is entirely based on the assumption that customers will not all withdraw their money from the banks on any given day. To what extent does faith play an important role in economics?

◆ **Open market operations** refer to the buying and selling of government securities by a country's central bank.

monetary policy if it wants to inject liquidity into the economy and increase the money supply (and, hence, reduce interest rates) to stimulate growth and employment.

Changes in the central bank minimum lending rate

The **minimum lending rate** (MLR) is the official rate of interest charged by the central bank on loans to commercial banks. This means the central bank will not lend money below this rate as this is based on their own cost of funds. The minimum lending rate is also known by several other terms, including the *base rate, discount rate* and *refinancing rate*.

◆ The **minimum lending rate** (MLR) is the official rate of interest charged by the central bank on loans to commercial banks. The central bank will not lend money below this rate.

As the MLR operates as the base rate for the banking system, it influences the interest rates charged on credit transactions, bank loans and mortgages. If the central bank or government decides to implement contractionary monetary policy (to achieve price stability) then it can raise the MLR, which tends to cause commercial banks and other financial institutions to also increase their lending rates. Hence, the higher lending rates mean that households and firms are likely to be discouraged from borrowing money while the higher interest rates can encourage people to save more. This will therefore reduce consumption (C) and investments (I), and therefore reduce the level of aggregate demand.

By contrast, if the central bank decides to use expansionary monetary policy, it can lower the minimum lending rate. This lower base rate means that commercial banks and other financial institutions will lower their refinancing rates and also reduce the interest rate paid to savers. This will tend to encourage consumers and firms to borrow more money and discourage them from saving (owing to the lower return). Therefore, a lower MLR will tend to increase consumption and investment expenditure, thereby increasing aggregate demand.

Quantitative easing

Following the global financial crisis of 2008, the UK and US central banks made significant cuts in interest rates to encourage people and businesses to spend money, rather than to save it. However, with interest rates at their lowest levels in history, near zero per cent, they could not really go any lower. This situation is known as a **liquidity trap** – when the MLR cannot be cut any further, thereby making traditional expansionary monetary policy ineffective. This is because despite the low interest rates, banks are reluctant to lend and/or consumers are reluctant to borrow due to the lack of confidence in the economy. The Bank of England and the Federal Reserve had one other option – quantitative easing.

◆ The **liquidity trap** exists when the minimum lending rate cannot be cut any further, as it is close to or at zero, thereby making traditional expansionary monetary policy ineffective.

Quantitative easing (QE) is a form of monetary policy that injects money directly into the economy (that is, increasing the money supply) via the central bank purchasing corporate bonds. Bonds are a form of government debt security. The institutions selling these bonds to the government, such as commercial banks or insurance companies, then have 'new' money in their accounts, that is, banks see an increase in their liquidity (or cash reserves). This helps to boost the money supply and to promote lending (and, hence, investment spending, and therefore boost aggregate demand). Higher lending should, in theory, help to improve economic growth. As QE is used to stimulate economic activity, it also helps to avoid a situation of deflation.

◆ **Quantitative easing** (QE) is a monetary policy tool that injects money directly into the economy via the central bank purchasing bonds (government-backed securities with a promise to pay later).

As the central bank creates new money (increasing bank reserves electronically) the effect is similar to printing new money without the hassle or costs. As the supply of money in the economy increases, interest rates will fall, so this reduces the value of real debts of households and firms with existing debts. Therefore, QE should increase consumer and business confidence, and increase aggregate demand. Furthermore, lower interest rates will cause exchange rates to fall, thereby making exports less expensive and imports relatively more expensive. This leads to an increase in export earnings and a fall in important expenditure, which again increases domestic aggregate demand, ceteris paribus.

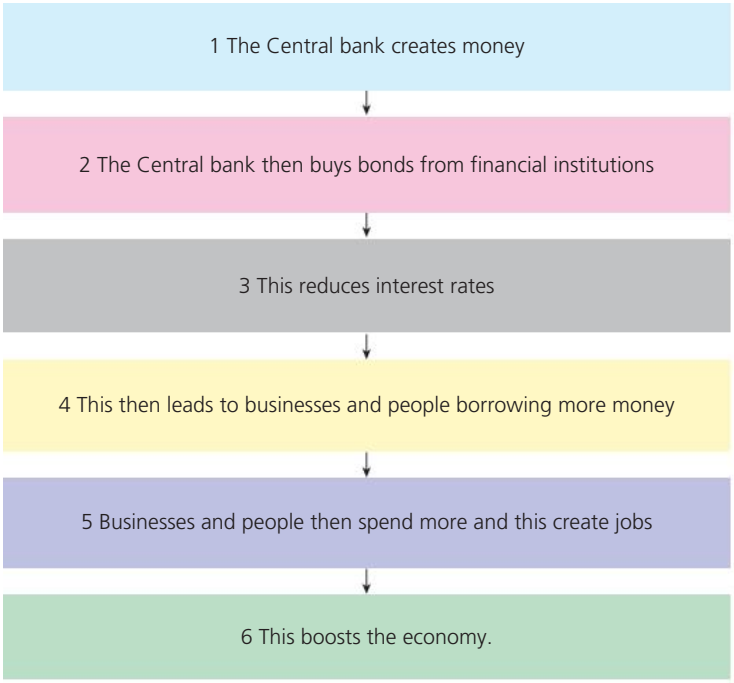


Figure 24.6 Quantitative easing: the theory

◆ The **demand for money** refers to the desire of households and firms to hold money (rather than saving it) in order to finance spending and investment.

Demand and supply of money (AO2, AO4) (HL only)

The interest rate can be described as the return for lenders of money or the price of borrowing money. It is expressed as a percentage of the money loaned to or borrowed by others. The equilibrium interest rate is determined by the intersection of the demand for and supply of money (see Figure 24.7).

The **demand for money** (D_m) refers to the desire to hold money (rather than saving it) to finance consumption and current expenditure. Interest rates tend to rise when the quantity of money demanded exceeds the quantity supplied. The opportunity cost of holding money varies directly with the level of interest rates, that is, a fall in interest rates will reduce the opportunity cost of holding money (there is not much of an opportunity cost if, like in the UK, Japan and Hong Kong, interest rates have been close to 0 per cent for many years).

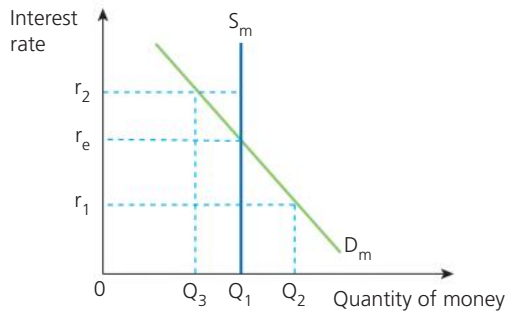


Figure 24.7 The demand for and supply of money

The **supply of money** (S_m) refers to the total amount of money circulating in the economy at any point in time. It will include bank notes and coins, bank deposits, loans and credit. An increase in the money supply will tend to decrease interest rates, and vice versa. The price of money is called the interest rate (shown on the *y*-axis) and the quantity of money in the economy is called the money supply (shown on the *x*-axis).

The supply of money curve (S_m) is vertical at Q_1 quantity of money because the supply of money is fixed by the central bank at any one point in time. At interest rates below the equilibrium, there is excess demand for money. For example, at an interest rate of r_1 excess demand is shown by the distance $Q_2 - Q_1$. Hence, there will be a tendency for the price of money

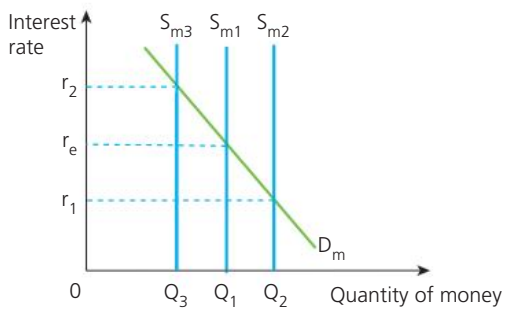


Figure 24.8 Shifts in the supply of money curve

(the interest rate) to rise towards its equilibrium. By contrast, when interest rates are higher than the equilibrium, more money is available than people wish to hold. Hence, at r_2, excess supply of money is $Q_1 - Q_3$. In this case, there will be a tendency for the interest rate to fall towards its equilibrium.

A change in either the demand for money or the supply of money will change the equilibrium interest rate. For example, in Figure 24.8, an increase in the money supply from S_{m1} to S_{m2} due to a lower minimum reserve requirement or quantitative easing, reduces the interest rate, ceteris paribus. This is a form of expansionary monetary policy. By contrast, if the government wants to reduce the money supply from S_{m1} to S_{m3} in order to raise interest rates, it might choose to sell government securities to corporate institutions. This is essentially a form of contractionary monetary policy.

While the central bank does not control the demand for money, it has a key role in influencing the supply of money by manipulating interest rates. There are several factors a central bank will consider when setting interest rates, including:

- The state of the economy – for example, a deflationary gap (see Chapter 17) may require a reduction in interest rates to prevent the economy from going into a deep recession.
- The rate of growth of nominal wages – for example, higher costs of labour usually mean that firms will increase prices. Higher interest rates might then be used to combat inflationary pressures (cost-push inflation).
- Business confidence levels – lower interest rates tend to create incentives for investment expenditure due to the lower costs and, hence, risks of investment.
- House prices – in many countries, house prices (the most valuable asset of a typical household) have a direct impact on the level of consumer confidence and, hence, the value of consumption expenditure and potential economic growth in the economy.
- The exchange rate – for example, lower interest rates might be needed to reduce the demand for the currency on the foreign exchange market (see Chapter 31). This will encourage the sale of exports, ceteris paribus.

Direct control of the money supply is relatively difficult as the definition of money is quite loose, and banks can create credit fairly easily. In reality, central banks or monetary authorities face difficulties in accurately controlling the supply of money in the economy. Hence, most governments rely on interest rate policy to achieve economic stability.

In practice, there is no single equilibrium interest 'rate' in an economy but a structure of different interest rates. This is because there are separate markets for different kinds of loans, such as short-term bank overdrafts, ongoing credit card debts and long-term mortgages; all of which charge different rates of interest. Borrowers also have different levels of risk. For example, lending to governments and large multinationals tends to be less risky than lending to private individuals. The riskier the client in terms of not being able to repay loans, the higher the interest rate will tend to be.

◆ The **interest rate** is the price of money. Expressed as a percentage measure, interest rates represent the cost of borrowing money or the return for savers.

Real versus nominal interest rates (AO2)

The **interest rate** is the price of money. Expressed as a percentage rate, interest rates represent the cost of borrowing money or return for savers. The European Central Bank defines interest rates as '*the cost of credit or the return on savings*'.

Economists distinguish between real and nominal interest rates. The **nominal interest rate** is the actual rate that is agreed between the bank and the customer, that is, it is the rate borrowers pay on their loans or the return savers receive on their cash deposits.

◆ The **nominal interest rate** is the actual rate that is agreed between a bank and the customer, that is, it is the rate borrowers pay on their loans or the return savers receive on their cash deposits.

However, the nominal interest rate is of relevance to economists who consider the **real interest rate** to both borrowers and savers. The real interest rate accounts for the impact of inflation on the return to savers and the cost of debts to borrowers. It is by subtracting the inflation rate from the nominal interest rate that determines the real interest rate on loans and savings.

◆ The **real interest rate** accounts for the impact of inflation on the return to savers and the cost of debts to borrowers: Real interest rate = Nominal interest rate – Inflation rate.

Real interest rate = Nominal interest rate – Inflation rate

For example, if a bank pays savers a nominal interest rate of 2 per cent per year, but inflation is 1.5 per cent during this time, the real return is only 0.5 per cent (as inflation has depleted the real return on the value of savings).

WORKED EXAMPLE

A person who saves $10,000 in a bank account for a year receives a nominal interest rate of 2.5%. The annualized inflation rate is 2%, that is, average prices rise by 2% during the year. Calculate the nominal and real return on the person's savings at the bank. *[2 marks]*

Solution

- The nominal interest rate = 2.5% per annum, so the nominal return = $10,000 × 2.5% = **$250** interest paid at the end of the year (before any income tax deductions)
- The real interest rate = 2.5% – 2% = 0.5%, so the real return = $10,000 × 0.5% = **$50**

It is possible for the real interest rate to be negative. **Negative real interest rates** occur when the nominal (actual) interest rate is less than the inflation rate. For example, if the nominal interest rate is 3 per cent while the inflation rate is 4.2 per cent, then the real interest rate is –1.2% (that is, 3 – 4.2). This means that an individual with $1,000 savings would earn $50 interest, that is, a total of $1,050 at the end of the year. However, inflation causes a good or service that cost $1,000 a year ago to increase to $1,060, so in real terms the money saved has *fallen* in value (in real terms).

◆ **Negative real interest rates** occur when the inflation rate is greater than the nominal interest rate, meaning any return on savings are completely diminished by the effects of higher prices in the economy.

Top tip!

You may be asked to calculate real interest rates in Paper 2 (SL and HL) and/or Paper 3 (HL only). The calculation is straightforward, but make sure you understand the meaning of real interest rates, so you know how to calculate these.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 24.1

Use the data below to calculate the real interest rate in 2018, 2019 and 2020. *[3 marks]*

Year	Consumer price index (CPI)	Nominal interest rate (%)
2017	100	5.0
2018	110	5.5
2019	120	5.0
2020	115	4.5

Expansionary and contractionary monetary policies (AO3, AO4)

Expansionary and contractionary monetary policies can be used by a government to close deflationary and inflationary gaps respectively. As a form of demand-side policy, monetary expansionary and contractionary policies are used to influence the level of economic activity. For example, changes in interest rates are used to influence the level of aggregate demand (recall from Chapter 17 that AD = C + I + G + X – M). In the case of expansionary monetary policy, lower interest rates will tend to shift the aggregate demand curve to the right in order to close a deflationary (or recessionary) gap, because:

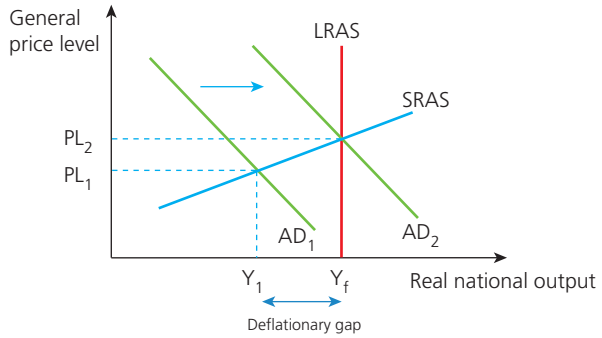


Figure 24.9 Expansionary monetary policy and the deflationary gap

- Consumption (C), investment (I) and government expenditure (G) are likely to rise due to the cheaper costs of borrowing money to spend and invest. Many households and firms with existing loans or mortgages will also benefit from lower interest repayments.
- Net exports (the value of exports minus imports) are likely to increase because lower interest rates tend to cause a fall in the exchange rate (see Chapter 31). This should make exports more relatively attractive to foreign buyers, and imports relatively more expensive for domestic households and firms, thereby helping to increase the level of aggregate demand.

Expansionary monetary policy (also referred to as **loose** or **easy monetary policy**) aims to increase the level of economic activity by reducing interest rates and/or expanding the money supply. This makes borrowing for consumption and investment purposes more attractive owing to the lower interest repayments charged on credit, loans and mortgages. Therefore, expansionary monetary policy will shift the aggregate demand curve to the right, ceteris paribus (see Figure 24.9), thereby helping to close a deflationary (recessionary) gap.

◆ **Expansionary monetary policy** aims to boost economic activity by expanding the money supply. This is achieved mainly by lowering interest rates, to stimulate aggregate demand, thereby closing a deflationary gap.

If the economy operates at less than the full employment level, it is operating along its SRAS curve. Expansionary monetary policy (such as through a cut in the base interest rate) will tend to increase the level of aggregate demand, shifting the AD curve from AD_1 to AD_2, ceteris paribus. This increases real national output from Y_1 to Y_f, thereby closing the deflationary gap (recall that a recessionary gap is the difference between equilibrium national output and the full employment level of national output).

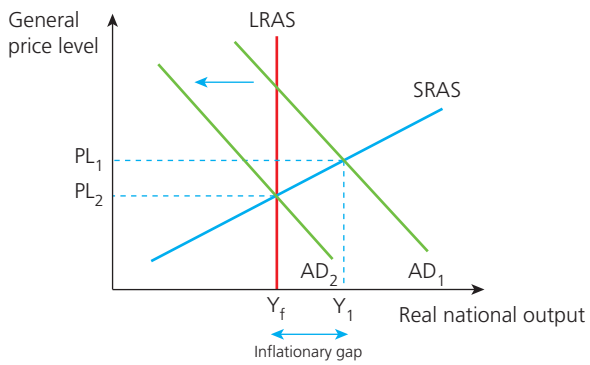


Figure 24.10 Contractionary monetary policy and the inflationary gap

However, a consequence of loose monetary policy is the potential emergence of inflationary pressures. This can be seen in Figure 24.9 with the general price level rising from PL_1 to PL_2. This is particularly the case if the economy is operating near full capacity, as shown by the vertical part of the economy's aggregate supply curve.

By contrast, the use of **contractionary monetary policy** (or **tight monetary policy**) can help to close an inflationary gap. An increase in interest rates, for example, tends to reduce consumption and investment in the economy, thereby reducing aggregate demand and real national output (see Figure 24.10).

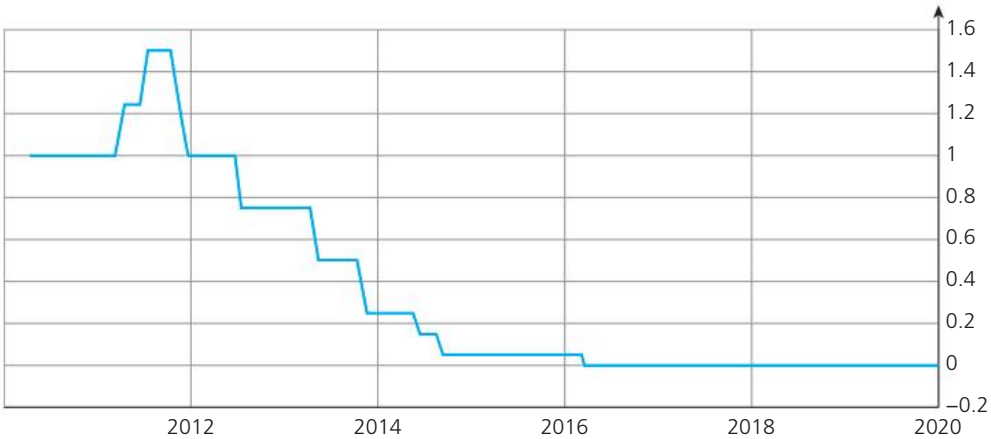
◆ **Contractionary monetary policy** aims to reduce the level of economic activity by restricting the money supply. This is achieved mainly by higher interest rates to suppress aggregate demand, thereby closing an inflationary gap.

Tight monetary policy, such as the use of higher interest rates shifts the aggregate demand curve leftwards from AD_1 to AD_2, ceteris paribus. This shrinks real national output from Y_1 to Y_f, so that actual output no longer exceeds potential output. Thus, contractionary monetary policy helps to close the inflationary gap, restoring prices back to PL_2 from PL_1.

Common mistake

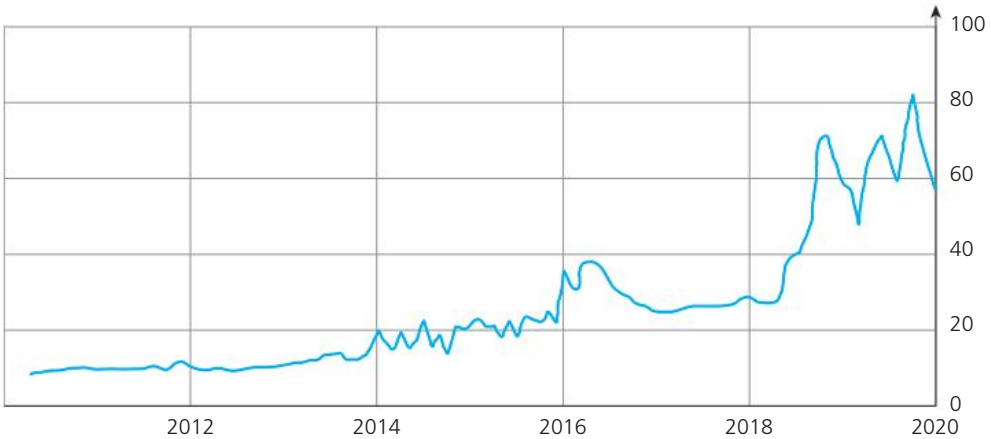
Beware of saying that higher interest rates reduce aggregate demand because they create incentives to save. While this is partially the case, the impact on borrowing is far more significant. Interest rates tend to change by 0.25 per cent at a time, which is barely an incentive for most individuals to save more. However, an extra 0.25 per cent interest charge on people's outstanding mortgages will certainly have an impact on their spending ability. The same applies to firms with existing loans and mortgages when interest rates rise.

PAPER 3 EXAM PRACTICE QUESTION 24.2 (HL ONLY)



Source: Trading Economics **https://tradingeconomics.com/spain/interest-rate**

Figure 24.11 Interest rates in Spain (%), 2010–20



Source: Trading Economics **https://tradingeconomics.com/argentina/interest-rate**

Figure 24.12 Interest rates in Argentina (%), 2010–20

With reference to each of the graphs above for Spain (Figure 24.11) and Argentina (Figure 24.12), explain why the respective governments have adopted contrasting interest rate policies. *[4 marks]*

Top tip!

Each policy decision comes with its drawbacks and opportunity cost. For example, contractionary monetary policy can be used to control the threat of inflation by imposing higher interest rates although can harm economic growth and therefore cause job losses in the long run.

Common mistake

When defining or referring to **demand-side policies**, be sure to always refer to these as macroeconomic policies in terms of both fiscal and monetary policy. This is where both policies are used to influence the level of aggregate demand in the economy.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 24.3

With the aid of an AD–AS diagram, explain how the use of expansionary monetary policy can help to close a deflationary (recessionary) gap. *[4 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 24.4

With the aid of an AD–AS diagram, explain how the use of contractionary monetary policy can help to close an inflationary gap. *[4 marks]*

Effectiveness of monetary policy (AO3)

The effectiveness of monetary policy can be evaluated through consideration of factors, including the constraints on monetary policy and strengths of monetary policy in achieving a country's macroeconomic objectives, which include economic growth, low unemployment, and a low and stable rate of inflation.

Constraints on monetary policy include limited scope of reducing interest rates when they are close to zero and low consumer and business confidence.

Limited scope of reducing interest rates, when close to zero

It is incorrect to assume that low interest rates *will* stimulate aggregate demand and pull an economy out of recession. For example, Japan's interest rates have been close to zero since the mid-1990s yet businesses have struggled to survive despite the easy access to money at almost zero cost. Similarly, Banco de España, Spain's central bank, kept interest rates at historic lows during the aftermath of the 2008 global financial crisis, but it took the country many years to get out of recession. Globally, central banks cut interest rates close to zero per cent in 2020 in an attempt to minimize the social and economic impacts of the coronavirus pandemic. On its own, expansionary monetary policy does not necessarily solve macroeconomic problems such as high unemployment and low economic growth, especially as there is limited scope to reduce interest rates further when they are already close to zero.

Low consumer and business confidence

Consumption and investment are not entirely dependent on interest rates. The 2008 financial crisis meant interest rates were close to zero per cent in Japan, the USA and Hong Kong, but the lack of business and consumer confidence led to a prolonged economic recession. A similar pattern occurred in 2020 during the global coronavirus pandemic, which brought consumer and business confidence to an all-time post-war low. Even at the best of times, changes in interest rates and the money supply can be destabilizing to the consumers and businesses. For example, while the property (housing) market in many countries is a key determinant of consumer and business confidence levels, this is highly vulnerable to regular changes in interest rates. Similarly, firms will not borrow money to invest, even at very low rates of interest, if the demand for their products remains low due to low consumer confidence in the economy.

Figure 24.13 Low confidence in the economy reduces the effectiveness of monetary policy

ATL ACTIVITY 24.1

Compare and contrast the economic costs of the coronavirus outbreak in any two countries of your choice. To what extent was monetary policy used to combat the economic problems of the crisis, and how effective were these policy measures?

WISE ChoICES

The nine concepts in economics can be a useful tool for students to consider topics from different perspectives. For example, disparities in **equity** mean that households and firms have different interest elasticity of demand, that is, some stakeholders are more affected by changes in interest rates than others. Not all homeowners have an outstanding mortgage, for example, while those with minimal savings do not benefit from an increase in interest rates. These disparities can make it difficult for the government to estimate the extent to which monetary policy is effective in influencing macroeconomic objectives.

Strengths of monetary policy include it being incremental, flexible and easily reversible, plus the relatively short time lags.

Incremental, flexible and easily reversible

The ability to adjust interest rates incrementally means policymakers can monitor the effectiveness of monetary policy. It is common for interest rates to change in increments of 0.25 percentage points, thus reducing the risks of causing huge disruptions to the economy. Central banks review the state of the economy and adjust interest rates accordingly, often on a monthly basis. Independence of the central bank allows decision makers, in theory, to have the flexibility to act in the best interest of the economy, without political interferences. The decision is reversible

if the government feels the need. For example, it may have underestimated the value of the money multiplier, which creates more spending in the economy than planned, thus fuelling inflation. The flexibility to implement changes in interest rates relatively quickly also means monetary policy can be used to influence and fine-tune macroeconomic objectives.

Short time lags

Monetary policy may be preferred to fiscal policy because it can be implemented more quickly. For example, following the UK government's 'Coronavirus budget' in March 2020, the Bank of England cut interest rates twice within a week to their lowest levels in history, in order to stabilize the economic downturn in the country. By contrast, fiscal policy (such as tax changes) requires careful planning and legislating. Hence, fiscal policy is relatively time consuming. In addition, using government spending to bring the economy out of a recessionary gap can be highly costly, thereby increasing the national debt (see Chapter 21). Nevertheless, there are time lags to the reaction of households and firms to changes in interest rates in the economy, especially if the levels of consumer and business confidence are low. This would make the effectiveness of monetary policy less certain.

Top tip!

When discussing the effectiveness of monetary policy as a form of demand-side policy, remember to consider the policy in terms of its ability to promote economic growth, achieve low unemployment, and foster a low and stable rate of inflation.

Top tip!

When discussing the strengths and limitations of monetary policy, remember to consider the potential conflicts among government macroeconomic objectives that may exist. For example, a cut in interest rates or an increase in the money supply to stimulate growth can cause inflationary pressures (price instability). Another example is that the use of contractionary monetary policy restricts economic activity so discourages foreign direct investment in the country (due to higher interest rates), which negatively impacts on profits, jobs and economic growth in the long run. Hence, the use of monetary policy can be counterproductive.

Top tip!

When evaluating the effectiveness of different macroeconomic policies, consider the impact in the long run, and not just the short run. For example, in the short run, monetary policy is generally more effective in dealing with demand-pull inflation than in getting an economy out of a deep economic depression, which might require the use of expansionary fiscal policy (see Chapter 25) and supply-side policies (see Chapter 26).

PAPER 1 EXAM PRACTICE QUESTION 24.5

Explain how monetary policy can be used to influence the level of economic activity. *[10 marks]*

WISE ChoICES

With reference to the use of monetary policies, examine how the **interdependence** of social, economic and environmental factors influence the effectiveness of government **intervention** in achieving macroeconomic objectives.

PAPER 1 EXAM PRACTICE QUESTION 24.6

Evaluate the effectiveness of monetary policy in achieving macroeconomic objectives. *[15 marks]*

Chapter summary

- As a form of demand-side management, monetary policy is the use of interest rates by the central bank to influence aggregate demand and the level of economic activity.
- Goals of monetary policy include: a low and stable rate of inflation (inflation targeting), low unemployment, reducing business cycle fluctuations, promoting a stable economic environment for long-term growth and achieving an external balance.
- Banks are able to create money through the credit creation process, which can be worked out by using the money multiplier process.
- Open Market Operations (OMO) occur when the central bank buys or sells government securities in order to influence interest rates. It either injects money into the economy to stimulate economic activity or withdraws money from the economy in order to reduce inflation.
- Expansionary monetary policy is used to stimulate aggregate demand by expanding the money supply or lowering interest rates, thus increasing economic activity.
- By contrast, contractionary monetary policy involves increasing interest rates or restricting the money supply to discourage over-consumption and limit investment expenditure in the economy, due to inflationary pressures.
- Real interest rates can be negative. For example, if the nominal (that is, actual) interest rate is 2 per cent but the inflation rate is 3 per cent, then the real interest rate is –1 per cent.
- In reality, there is no single interest 'rate' in an economy. Instead, a structure of different interest rates exists for different financial services, such as bank loans, credit cards and mortgages.
- The effectiveness of monetary policies can be determined by the extent to which they help to achieve the macroeconomic objectives (economic growth, low unemployment and price stability). However, monetary policy is limited by the liquidity trap and low consumer and producer confidence levels.

25 Demand management – fiscal policy

Real-world issue

How do governments manage their economy and how effective are their policies?

Conceptual understandings (WISE ChoICES)

- Government **intervention** attempts to achieve macroeconomic objectives through a **choice** of policies.
- Political, economic, social and environmental factors are **interdependent** and will influence the effectiveness of government policies.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- the main sources of revenue and government expenditure (AO2)
- the goals of fiscal policy (AO2)
- the role of expansionary and contractionary fiscal policies in order to close deflationary and inflationary gaps (AO3)
- an AD–AS diagram to show expansionary and contractionary fiscal policy for both Keynesian and monetarist/new classical schools of thought (AO4)
- the Keynesian multiplier (AO2) (HL only)
- the Keynesian multiplier (AO4) (HL only)
- the effect on GDP of a change in an injection, using the Keynesian multiplier (AO4) (HL only)
- the effectiveness of fiscal policy with reference to the constraints on and strengths of fiscal policy, with reference to macroeconomic objectives (AO3)
- the crowding-out effect as a constraint on fiscal policy (AO3) (HL only)
- a diagram to show the crowding-out effect (AO3) (HL only)
- the role of automatic stabilizers – progressive taxes and unemployment benefits (AO3) (HL only).

Fiscal policy (AO2)

Fiscal policy is the use of taxation and government expenditure strategies to influence the level of economic activity to achieve the macroeconomic aims of low unemployment, sustainable economic growth and low inflation (price stability). For example, taxation can be used to redistribute income and wealth to improve standards of living in the country, such as redistributing income to lower income groups, or by building schools, hospitals and transportation networks. As a type of demand-side management, fiscal policy directly influences the level of aggregate demand in the economy. Fiscal policy is discretionary, meaning that it is purposively carried out by the government to influence the level of aggregate demand. The government can use expansionary fiscal policy to stimulate aggregate demand and use contractionary fiscal policy to reduce the level of economic activity.

◆ **Fiscal policy** is the use of taxation and government expenditure policies to influence the level of economic activity and macroeconomic objectives.

Fiscal policy can be used to directly promote long-term economic growth and low unemployment in three main ways:

- Government spending on physical capital goods – capital goods are tangible assets used to produce other goods and services. Examples include spending on machinery, buildings and commercial vehicles to increase the productive capacity of the economy.
- Government spending on human capital formation – human capital formation is the transformation process of using education and/or training to create knowledge and skills to create a more skilled labour force. Again, this helps to increase the economy's production capacity in the long run.
- Provision of incentives for firms to invest – government provision of tax breaks and tax incentives helps to create an economic environment that is conducive to investment for long run economic growth and development. Government spending on infrastructure can also help to encourage foreign direct investment, thereby promoting long-term economic growth and low unemployment in the economy.

Sources of revenue

Before the government can spend money on the economy, it must first take the money from taxpayers (both from individuals and firms). Taxation is the main source of government revenue. These can be categorized as direct and indirect taxation. Other sources of government revenue include the sale of goods and services from state-owned enterprises and privatization proceeds (from the sale of government assets).

Direct and indirect taxation

A **tax** refers to a government levy on income or expenditure. Taxes can be categorized as direct or indirect tax. Direct taxation is a type of tax imposed on the income, wealth or profit of individuals and firms. Examples include taxes on wages and salaries, inheritance and company profits. Indirect taxation refers to expenditure taxes imposed on the spending of goods and services in the economy. For example, countries such as Australia and Singapore impose a 'goods and services tax' (GST), whereas the European Union uses 'Value Added Tax' (VAT). Other examples include taxes on the sale of petrol, alcohol and cigarettes.

◆ A **tax** refers to a government levy on income directly or indirectly through expenditure. Therefore, taxes can be categorized as direct or indirect tax.

Examples of direct taxes

- **Income tax** is levied on personal incomes, such as, wages, interest, rent and dividends. In most countries, this is the main source of tax revenue for the government.
- **Corporation tax** is a direct tax on the profits of businesses.
- **Capital gains tax** is a levy on the earnings made from investments such as buying shares and private property.
- **Inheritance tax** is a tax on the transfer of income and wealth, such as money or property bequeathed (passed on to) another person.
- **Windfall tax** is charged on individuals and firms that gain an unexpected one-off amount of money, such as a person winning the lottery or the gains from a takeover bid for a firm.

Examples of indirect taxes

- **Sales tax** is an indirect tax, such as Value Added Tax (VAT) or Goods and Services Tax (GST), charged on the manufacturing, sale and consumption of goods and services.
- **Excise duties** are indirect inland taxes imposed on certain products. Depending on the country, these might include alcohol, tobacco, petrol (gas), soft drinks and gambling.
- **Customs duties** are indirect cross-border taxes on foreign imports.
- **Stamp duty** is a progressive tax paid on the sale of commercial or residential property.
- **Carbon tax** is a tax on vehicle manufacturers or firms that produce excessive carbon emissions.

Top tip!

In addition to taxation (direct and indirect taxation), the sale of goods and services from state-owned enterprises and sale of government assets, there are other sources of government finance not mentioned in the IB Economics syllabus. These include loans (public sector borrowing), sovereign wealth funds (state-owned investment funds, such as stocks and shares, investments in property and gold reserves) and donations (financial aid). ATL Activity 25.1 will provide further examples for you.

ATL ACTIVITY 25.1

Investigate the government budget for a country of your choice. Construct a table to show the main sources of tax revenues for the country and its main items of government expenditure. Be prepared to share your results in class.

Figure 25.1 All governments have annual budgets

The sale of goods and services from state-owned enterprises

The sale of goods and services is another source of government revenue. These come from state-owned enterprises (also known as **nationalized industries**), and will vary between countries, but could include national providers of postal services, the national broadcasting corporation, the airport authority and nationwide railway operators. As the government does not strive to primarily earn a profit, these revenue sources usually go towards paying for the costs of providing the good or service.

The sale of government assets

The sale of state-owned assets and enterprises is also a source of government revenue. This practice is known as **privatization** (see Chapter 26). Privatization proceeds can be earned by selling government-owned assets and enterprises to shareholders in the private sector. Many governments have privatized industries such as telecommunications, postal services, railroads, airlines, steel works, utilities (gas, water and electricity companies), and even banks. Privatization was popularized by Margaret Thatcher, the UK's first female prime minister, who administered the sale of major British businesses, including British Airways, British Gas, British Telecom and British Steel.

◆ **Privatization** is the transfer of the ownership of government assets from the public sector to owners (shareholders) in the private sector.

Thatcher (1925–2013) was heavily influenced by economist FA Hayek, known for his defence of free-market capitalism. However, privatization is a short-lived policy as state-owned assets can be sold only once to the private sector, and there is only a finite supply of such government assets.

CASE STUDY

Hong Kong's MTRC

The MTR Corporation Limited is a publicly listed company on the Hong Kong Stock Exchange. The MTRC operates Hong Kong's railway services, as well as being a major property developer. The company was privatized on 5 October 2000, having raised HK$9.4 billion (US$1.22bn) in its initial public offering. The MTRC also has investments in railways in Beijing, Hangzhou, London, Melbourne, Shenzhen, Stockholm and Sydney.

Figure 25.2 Sha Tin MTR Station, Hong Kong

Expenditures

Government expenditure is a key component of aggregate demand (see Chapter 17). All things being equal, an increase in government spending (G) helps to boost real national output, employment and economic growth. Government spending therefore helps to improve standards of living, such as by building and improving schools, hospitals and transportation and infrastructure networks. The government uses its sources of revenue to finance its expenditure, which can be categorized as current expenditures, capital expenditures or transfer payments.

Current expenditures

Current expenditure refers to government spending on goods and services consumed within the (current) year. Hence, it is sometimes referred to as **consumption expenditure**. This includes items of spending on a reoccurring basis, often for day-to-day purposes. Current expenditure is for immediate operations and benefits to satisfy the needs of individuals or groups in society. Examples of such spending include:

- wages and salaries for public sector employees, such as civil servants, diplomats, politicians, firefighters, police officers, public healthcare workers and tax inspectors
- current expenditure on healthcare and education, such as supplies for government hospitals, clinics and schools
- provision of subsidies and grants from the government to support targeted industries
- interest repayments on loans made to the government, that is, public sector debts.

Capital expenditures

Capital expenditure refers to long-term items of government spending (public sector investments) that boost the economy's productive capacity. It involves the government spending a large amount of money to increase the nation's capital stock, so this needs to be planned and funded carefully. Hence, capital expenditures (sometimes referred to as **fixed capital formation**) are intended to create future benefits for all members of society. Examples of government capital expenditure include spending on physical infrastructure, such as new roads, motorways (highways), tunnels, harbours, ports, airports, state schools and public hospitals. Ideally, the government would borrow money only to fund capital expenditure (that is, to fund investment expenditure in the economy), rather than current expenditure.

◆ **Current expenditure** refers to government spending on goods and services consumed within the (current) year. It is sometimes referred to as consumption expenditure.

◆ **Capital expenditure** refers to long-term items of government spending (public sector investments) that boost the economy's productive capacity.

Top tip!

Both current expenditure and capital expenditure contribute to the G component (government spending) in the expenditure method of calculating national income, given that AD = C + I + G + (X – M).

Transfer payments

Transfer payments are welfare expenses from the government, used to redistribute income to others in society by funding essential public services such as the provision of state education, housing, healthcare, social housing and postal services. Unlike current and capital expenditure, there is no corresponding exchange of goods or services when the government spends money on transfer payments. Examples of such expenditures on essential services to low-income groups include unemployment benefits (job seekers' allowance), state pensions (for retirees), housing benefits, disability allowances and child allowances.

◆ A **transfer payment** is a sum of money from the government to another party for which no goods or services are being paid, that is, there is no corresponding output or economic activity for such payments.

Figure 25.3 National defence is a large item of government spending

Top tip!

Ensure you understand the meaning of transfer payments. They are considered to be part of government spending but do not appear as part of the national income (GDP) statistics as there is no corresponding output for transfer payments.

WISE ChoICES

To what extent do transfer payments help to achieve greater **choice** and **equity** in society?

Goals of fiscal policy (AO2)

The goals of fiscal policy are to tackle macroeconomic issues that can be remembered by the acronym **BUDGIE** (see Figure 25.4).

Figure 25.4 The goals of fiscal policy

Low and stable inflation

Taxation policies can also be used to help control the rate of inflation in the economy, thereby promoting price stability and affecting the country's international competitiveness.

In Figure 25.5, contractionary fiscal policy, such as the use of higher tax rates and/or deliberately running a budget surplus reduces consumption (C), investment (I) and government spending (G), thereby shifting the aggregate demand curve leftwards from AD_1 to AD_2. This shrinks real national GDP from Y_1 to Y_f. Hence, contractionary fiscal policy helps to close an inflationary gap and control inflation by reducing the general price level from PL_1 from PL_2.

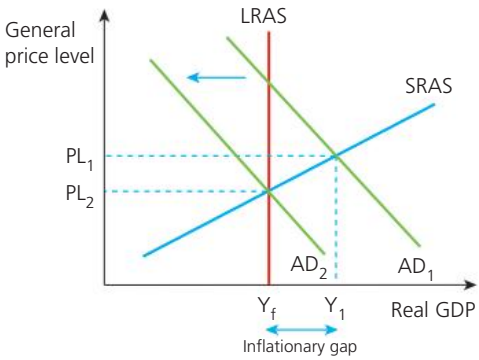


Figure 25.5 Fiscal policy and the control of inflation

Low unemployment

Governments can deal with the problems of unemployment in various ways, depending on the types and causes of unemployment in the economy. In the case of fiscal policy, this is best suited to tackle cyclical unemployment (see Chapter 19). Cyclical unemployment (or demand-deficient unemployment) occurs during a recession, with falling aggregate demand causing mass job losses across most if not all industries. Hence, the government might choose to use expansionary fiscal policy to stimulate economic activity and, hence, to reduce the level of unemployment.

Expansionary fiscal policy involves a reduction in taxes and/or increased government expenditure. This should, in theory, increase the level of aggregate demand and hence the derived demand for labour (see Figure 25.6).

In Figure 25.6, tax cuts help to increase the level of consumption (C) and investment expenditure (I), and along with increased government spending (G) help to increase the level of aggregate demand from AD_1 to AD_2. Therefore, this raises the real national income of the economy from Y_1 to Y_f. In turn, this will lead to more employment opportunities thereby lowering unemployment in the economy.

Promote a stable economic environment for long-term growth

Fiscal policy can help to create a stable economic environment that promotes sustainable economic growth. Higher taxes tend to reduce incentives to work and to produce, which may therefore have a negative impact on the economy as a whole. However, tax cuts can boost spending in the domestic economy, thus benefiting individual customers, businesses and the economy (through job creation). Lower taxes can also attract foreign direct investments, which helps to create an economic environment conducive to long-term growth (see Table 25.1). Essentially, fiscal policy can be used to promote long-term economic growth by creating an environment of low taxation that is favourable to private sector investment by domestic and foreign firms.

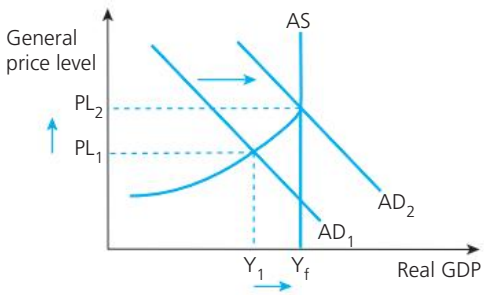


Figure 25.6 Expansionary fiscal policy to reduce unemployment

Nevertheless, taxes cannot be too low as tax revenues are essential to fund government spending (such as the construction of schools, hospitals, railways, airports and roads), which is also vital to promote long-term economic growth.

Table 25.1 Tax rates in selected countries

Country	Corporate tax rate (%)	Sales tax rate (%)	Highest marginal tax rate (%)
United Arab Emirates	0.0	5.0	0.0
Qatar	10.0	0.0	0.0
Hong Kong	16.5	0.0	17.0
Sweden	21.4	25.0	57.2
Indonesia	25.0	10.0	30.0
Canada	26.5	5.0	33.0
Sri Lanka	28.0	8.0	24.0
Australia	30.0	20.0	45.0
Mozambique	32.0	17.0	32.0
Chad	35.0	18.0	60.0

Sources: https://tradingeconomics.com/country-list/corporate-tax-rate, https://tradingeconomics.com/country-list/sales-tax-rate, https://tradingeconomics.com/country-list/personal-income-tax-rate

CASE STUDY

France's mass demonstrations

France tried to introduce a 75 per cent income tax rate in 2012 for individuals earning incomes in excess of €1 million ($1.09m per year). However, the proposals were overturned, with some economists arguing that the government would actually receive more tax revenues by *cutting* tax rates. This is because lower rates of tax can create incentives to work and also help to reduce tax avoidance and tax evasion. Mass demonstrations began in Paris starting in November 2018 due to protests against the government for its perceived unfair tax system.

Figure 25.7 Mass protests in Paris, France

Reduce business cycle fluctuations

The **government budget** is a country's planned (forecasted) revenues and its planned expenditures over a given period of time (usually one year). To reduce the impacts of a recession in the business cycle (see Chapter 16), the government can run a budget deficit, spending more than it gets from its various sources of government revenue such as tax revenues. For example, social welfare payments tend to rise during an economic downturn, while tax revenues fall due to rising unemployment. The budget deficit is needed to stabilize the economy and is financed (paid for) by using government borrowing.

◆ The **government budget** is a country's planned (forecasted) revenues and its planned expenditures over a given period of time (usually one year).

By contrast, if the economy experiences an economic boom, tax revenues will be higher due to increased earnings and consumption by the average household, while government spending on transfer payments will tend to fall. This also helps to reduce fluctuations in the business cycle. Hence, as government revenues exceed public sector expenditure, the country has a **budget**

surplus. This will generally help to reduce the overall value of government debt, caused by budget deficits in previous years.

Equitable distribution of income

Despite a country being able to achieve economic growth, inequality in disposable income and wealth has grown in many advanced and high-income economies. Fiscal policy can be used to redistribute income and wealth to benefit poorer members of society, such as the unemployed or retirees. This can be achieved by using high marginal tax rates in a progressive tax system (see Chapter 22) and wealth taxes to fund the provision of essential goods and services such as universal education and public healthcare. Another example is the use of transfer payments to improve the well-being of poorer members of society. Such payments include unemployment benefit, child benefit (financial assistance for parents), social housing and state pensions. Hence, fiscal policy can be used to greatly improve the economic conditions of poorer individuals and households.

Table 25.2 Highest marginal tax rate in selected countries

Country	Highest marginal tax rate (%)	Country	Highest marginal tax rate (%)
Ivory Coast	60.0	Saudi Arabia	0.0
Denmark	55.8	Romania	10.0
Ireland	48.0	Russia	13.0
Germany	45.0	Hungary	15.0
South Korea	42.0	Singapore	22.0

Source: Trading Economics https://tradingeconomics.com/country-list/personal-income-tax-rate

Theory of Knowledge

To what extent can we rely on statistical data collected by economists to determine and justify an economic policy decision?

External balance

The external balance refers to value of a nation's exports (X) being equal to the value of its imports (M), that is, X = M. Indirect taxes imposed on imports and/or government subsidies for domestic exporters will generally increase the external balance as exports become relatively more appealing in terms of price. Hence, this will tend to increase the demand for exports, so that X > M causing an external disequilibrium and leading to inflationary pressures in the long run as more money flows into the economy. In addition, such protectionist measures are likely to cause retaliation from governments of other countries (see Chapter 29), which can therefore lead to economic conflicts.

By contrast, if the value of M > X, the economy will face net withdrawals from the circular flow of income, so will experience a negative external balance. This means the country spends more than it earns from international trade, which is not sustainable in the long run. Hence, fiscal policies can be used to achieve an external balance, which is considered essential for the country's long-term economic growth and development.

Top tip!

When evaluating the effectiveness of demand-side policies in reducing unemployment, do note that fiscal and monetary policies might not deal with the root cause(s) of unemployment, such as incentives to work and a reluctance to accept paid work at prevailing wage rates.

Top tip!

While fiscal policy is typically considered as a demand management policy, the use of taxation and government spending can also be classified as supply-side policies as they can be used to improve incentives and the economy's productive capacity in the long run. For example, tax cuts can improve incentives to work, while increased government expenditure on healthcare and education help to improve the quality of the economy's resources.

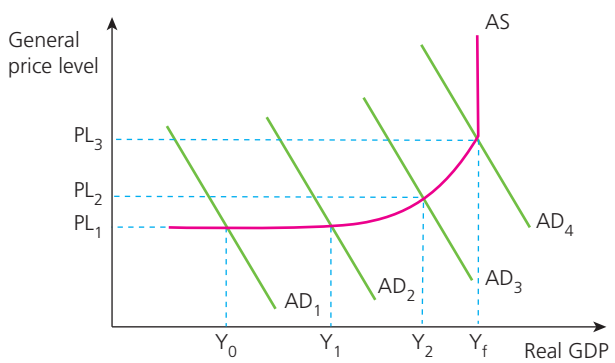
Expansionary and contractionary fiscal policies (AO3, AO4)

Expansionary fiscal policy (sometimes referred to as **reflationary fiscal policy**) is a demand-side policy used to stimulate the economy during an economic recession, by increasing government expenditure and/or lowering taxes to boost consumption and investment, thereby helping to close a deflationary (recessionary) gap. Government expenditure (G) itself is a key component of aggregate demand (AD = C + I + G + X – M). Thus, changes in the level of government spending will directly influence the level of AD in the economy, ceteris paribus.

◆ **Expansionary fiscal policy** is the use of increased government spending and/or reduced taxes in order to stimulate the level of economic activity.

Keynesian economists believe there is a need for direct government intervention through demand management policies in order to influence the level of economic activity, especially during an economic downturn and times of high unemployment. For example, the government could increase social security and welfare payments to support low-income households (such as unemployment benefits, child benefits or state pensions). It can also reduce income taxes in order to raise household disposable income or corporate profits. Hence, domestic consumption and investment should increase thereby increasing aggregate demand and real GDP. In turn, this will create jobs in the economy (see Figure 25.8).

In Figure 25.8, expansionary fiscal policy through income tax cuts and/or an increase in the level of government spending increases aggregate demand from AD_1 to AD_2, thus boosting real GDP from Y_0 to Y_1. As the Keynesian AS curve is perfectly price elastic due to spare capacity in the economy, there is no pressure on the general price level, which remains at P_1. If these expansionary policies are continued, then aggregate demand will increase from AD_2 to AD_3, which will mean an increase in real GDP from Y_1 to Y_2 and, hence, an increase in economic growth and lower unemployment. However, this will add inflationary pressure as the general price level rises from PL_1 to PL_2. Similarly, if aggregate demand is allowed to continue to expand from AD_3 to AD_4, the change in real GDP from Y_2 to Y_f would be less than the proportional change in the general price level from PL_2 to PL_3 due to the increasing scarcity or factors of production, which leads to general levels of inflation. Hence, there is a trade-off between real national output and the level of inflation. However, Keynesians believe that expansionary fiscal policies are effective in getting the economy out of a recession and stimulating growth up to the point where it operates at the full employment level of national output (shown at Y_f in Figure 25.8).



■ **Figure 25.8** The impact of expansionary fiscal policy (Keynesian view)

By contrast, monetarists/new classical economists are more critical of the ability of expansionary fiscal policy to stimulate economic growth. They argue that fiscal policies are ineffective in the long run as they have no real impact on the level of national output. This is illustrated in Figure 25.9 where the use of demand-side policies simply creates inflationary pressures as general prices rise from PL_1 to PL_2, without any corresponding change in the full employment level of real national output (Y_f).

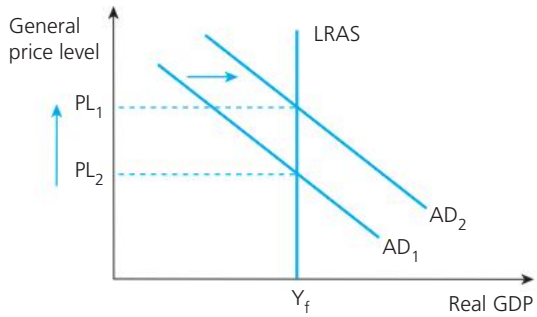


Figure 25.9 The impact of expansionary fiscal policy (Monetarist view)

Therefore, monetarists/new classical economists prefer to emphasize the use of supply-side policies (see Chapter 26) to increase the economy's long run aggregate supply (LRAS) and reduce the natural rate of unemployment (see Chapter 19). Thus, monetarists argue there is no long-term trade-off between lower unemployment and inflation. They also emphasize the importance of controlling the money supply to keep inflation low (see Chapter 23) rather than relying on demand-side policies.

Contractionary fiscal policy is used to reduce the level of economic activity by decreasing government spending and/or raising taxes to limit consumption (C) and investment (I). For example, countries such as China and the USA have used property taxes to slow down escalating house prices in certain cities and provinces. Contractionary fiscal policies tend to be used to reduce inflationary pressures during an economic boom, thus helping to close an inflationary gap (see Figure 24.10 on page 374).

◆ **Contractionary fiscal policy** is used to reduce the level of economic activity by decreasing government spending and/ or raising taxes to limit consumption (C) and investment (I).

Top tip!

Ensure you can explain the difference between a deflationary gap and an inflationary gap as these are featured in **four** different sections of the IB Economics syllabus: 3.2, 3.3, 3.5 and 3.6. The former occurs when the economy is operating below full employment. This means fiscal stimulus (tax cuts and/or increased government spending) is needed to close the recessionary gap. The opposite applies to an inflationary gap (when real GDP exceeds potential GDP), so contractionary fiscal measures are needed to combat the subsequent inflationary pressures.

PAPER 1 EXAM PRACTICE QUESTION 25.1

Explain how fiscal policy can be used to influence the level of aggregate demand in an economy. *[10 marks]*

Keynesian multiplier (AO2, AO4) (HL only)

The **Keynesian multiplier** shows that any increase in the value of injections into the circular flow of an economy results in a proportionately larger increase in aggregate demand. Recall from Chapter 16 that injections are the sources of spending that increase spending in the economy, namely export earnings (X), government spending (G) and investments (I). All things being equal, an increase in any of the injections will increase the value of the Keynesian multiplier.

◆ The **Keynesian multiplier** shows that any increase in the value of injections into the circular flow of an economy results in a proportionately larger increase in aggregate demand.

The Keynesian multiplier comes about because injections of money into the circular flow of income stimulate further rounds of spending, that is, increased spending by an individual becomes income for another individual. For example, if the government spends $500 million on a

new social housing project, other individuals, firms and industries will also benefit, such as supermarkets, restaurants, schools, real estate agents, taxi firms and so forth. Hence, the initial $500m injected by the government generates a far greater value of final output and, hence, creates further job opportunities in the economy.

By contrast, leakages or withdrawals reduce the value of the Keynesian multiplier. These comprise savings (S), taxes (T) and import expenditure (M), which take money out of the domestic circular flow of income. This creates what is called a **negative multiplier effect,** when an initial leakage from the circular flow of income leads to a greater than proportionate fall in final real GDP.

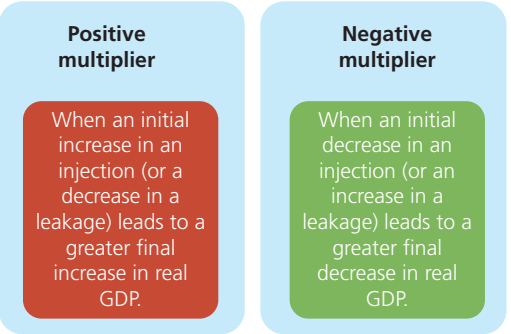


Figure 25.10 The positive and negative Keynesian multiplier

The key to understanding the Keynesian multiplier is knowing the following four determinants of the size of the multiplier effect:

- The **marginal propensity to consume** (MPC) – this measures the proportion of an increase in household income that is spent on goods and services rather than saved. The value of the MPC = ΔC ÷ ΔY, where C = consumption and Y = income. The greater the MPC (spent on domestically produced goods and services), the greater is the multiplier effect, ceteris paribus. For example, a reduction in income tax rates will increase the amount of extra disposable income that households can spend on goods and services.
- The **marginal propensity to import** (MPM) – this measures the proportion of an increase in household income that is spent on imports rather than on domestically produced goods and services. The value of the MPM = ΔM ÷ ΔY. The larger any leakage is, the lower the value of the Keynesian multiplier effect. So, if households spend a larger proportion of any extra income on imports, the value of the Keynesian multiplier will fall, ceteris paribus. Essentially, a high MPM, MPS and MPT will reduce the value of the Keynesian multiplier and reduce the impact of any attempt by the government to boost spending in the economy.
- The **marginal propensity to save** (MPS) – this measures the proportion of an increase in household income that is saved rather than spent on consumption or imports. The value of the MPS = ΔS ÷ ΔY. The more that people save of any increase in their income, the lower the Keynesian multiplier will tend to be as more money leaks from the circular flow of income and expenditure.
- The **marginal propensity to tax** (MPT) – this measures the proportion of each extra dollar of income earned that is taxed by the government. It will be dependent on the marginal rate of tax paid by the household. The value of the MPT = ΔT ÷ ΔY. The higher the marginal tax rates (see Table 25.2), the lower the Keynesian multiplier will be as more money is withdrawn from the economy, ceteris paribus.

◆ A **negative multiplier effect** occurs when an initial leakage from the circular flow of income leads to a greater than proportionate fall in final real GDP.

The value of the Keynesian multiplier can be calculated by using one of the following formulae:

$$\text{Keynesian multiplier} = 1 \div (1 - \text{MPC})$$

$$\text{Keynesian multiplier} = 1 \div (\text{MPS} + \text{MPT} + \text{MPM})$$

WORKED EXAMPLE

Suppose the government injects $400 million into the economy to get it out of a negative output gap, and that the following are known:

- The marginal propensity to save (MPS) = 0.1
- The marginal rate of tax on income (MPT) = 0.2
- The marginal propensity to import goods and services from overseas (MPM) = 0.3

a Calculate the value of the Keynesian multiplier. *[1 mark]*

b Using your answer from above, show how much the initial injection of money increases the value of final real GDP in the economy. *[2 marks]*

Solutions

a The Keynesian multiplier = 1 ÷ (MPS + MPT + MPM) = 1 ÷ (0.1 + 0.2 + 0.3) = 1 ÷ 0.6 = **1.67**

b $400m × 1.67 = $666.67 million (to 2 dp)

Hence, the initial change of an extra $400m injected into the economy leads to a final rise in real GDP by $266.67m to $666.67m.

Note that the shape of the aggregate supply curve will affect the size of the Keynesian multiplier, and, hence, the effectiveness of macroeconomic policies. If the economy is operating on the flat section of the SRAS curve (see Figure 25.8), there is a lot of spare capacity and mass unemployment, so the multiplier would significantly increase real national output (thereby reducing unemployment). This would happen without much, if any, increase in the general price level. By contrast, if the economy is on the vertical section of the (long run) AS curve, with the economy operating at the full employment level of national output, the Keynesian multiplier will have no impact on real GDP, but there will be an increase in the general price level (see Figure 25.9). If the economy operates on the upward sloping section of the AS curve, the multiplier effect will impact both real national output and the general price level.

Top tip!

The formulae for calculating the value of the Keynesian multiplier are essentially the same, because what is not spent on consumption is saved, taxed and/or spent on imports. The value of '1' represents 100 per cent of income. So, 1 – MPC = MPS + MPT + MPM. This means that the Keynesian multiplier can also be calculated as **1 ÷ MPW**, where the marginal propensity to withdraw (MPW) is the sum of MPS, MPT and MPM. Which formula you will have to use in the exam will depend on the information you are presented with.

PAPER 3 EXAM PRACTICE QUESTION 25.2 (HL ONLY)

a Explain why calculating the value of the Keynesian multiplier can be useful for government macroeconomic policymaking. *[2 marks]*

b Calculate the expected change in the value of real national output if the government spends $9.5 billion on infrastructure and the country's marginal propensity to consume (MPC) is known to be 0.75. *[2 marks]*

c Calculate the amount of government expenditure needed to restore macroeconomic equilibrium if the economy faces a $15 billion recessionary gap and its MPC is 0.8. *[2 marks]*

d Calculate the impact of real national output if the government increases tax revenues by $20 billion due to inflationary pressures in the economy, and the MPC is known to be 0.75. *[2 marks]*

Top tip!

If you are an HL student, you should be able to include the concept of the Keynesian multiplier when evaluating the effectiveness of contractionary or expansionary fiscal policies.

Effectiveness of fiscal policy (AO3)

The effectiveness of fiscal policy can be judged against the extent to which it achieves specific macroeconomic objectives, such as economic growth, a stable rate of inflation and low unemployment.

Constraints on fiscal policy

The constraints on fiscal policy in achieving its macroeconomic goals include political pressures, time lags and sustainable debts. For HL students, a fourth limitation exists – the crowding-out effect.

Political pressures

Political constraints and political pressures mean that government spending and taxation policy decisions are often influenced by political rather than economic factors. In many countries, the political cycle (of re-electing political leaders and political parties) can cause artificial shocks to the business cycle, such as large tax cuts being introduced prior to a general election simply in order to gain political votes. This may happen instead of tackling fundamental economic problems such as price instability and a drop in the country's international competitiveness due to rising inflation in the domestic economy. Hence, political pressures prevent the use of deflationary fiscal policies, such as higher taxes, due to the fear of losing votes from the electorate. Monetary policy, by contrast, is reviewed and changed by the central bank of the country (see Chapter 24), without direct political interference.

Time lags

Changes to fiscal policies take time to plan, get approval and execute before their impact can be measured. For example, the introduction of a consumption tax or adjusting the rate of income tax cannot be changed quickly in most countries owing to the legislative processes and administrative procedures that must take place prior to gaining final approval.

There are three time-related problems with the use of fiscal policy:

- **Recognition time lags** – There is a delay in recognizing that government intervention is needed to influence the level of economic activity. This is because governments do not necessarily know if the economy is growing too fast (or declining too quickly). Furthermore, supply-side shocks may occur at any time, so are difficult to plan and budget for, such as natural disasters or the widespread outbreak of infectious diseases.
- **Administrative time lags** – There is a time delay between recognizing the need for fiscal policy intervention and the time to actually implement an appropriate course of action, such as getting approval for tax changes or alterations to the government budget.

Figure 25.11 Fiscal policy changes incur time lags

- **Impact time lags** – There is a time lag from the implementation of fiscal policies to seeing the actual effects take place in the economy. A small reduction in income tax, for example, is likely to take some time to filter into the economy and to have any significant impact on the spending habits and consumption level of the average household.

Therefore, time lags (delays) slow down the implementation and, hence, effectiveness of fiscal policy.

Sustainable debt

Another constraint on the effectiveness of fiscal policy being able to achieve macroeconomic objectives is budgetary constraints. The effectiveness of fiscal policy will depend on the extent to which the government can afford to sustain a budget deficit. In the short run, governments may be able to run budget deficits (when government expenditures exceed public sector revenue) in order to fund expansionary fiscal policies to get the economy out of a recession. However, this is not sustainable in the long run. Austerity measures (see Chapter 21) in Italy, Portugal, Greece and Spain were inevitable following the huge public sector debts incurred by these countries and many others during the global financial crisis. Eventually, the government will need to repay its national debts, which will limit the extent to which fiscal policies can be used to influence the level of economic activity, because it lacks the necessary funds for public sector spending.

Top tip!

Recall that the effectiveness of expansionary fiscal policy will partly depend on the (shape of the) AS curve. It is also debatable whether demand-side policies alone are sufficient to achieve long-term economic growth and low unemployment; there is certainly a case for the use of monetary and supply-side policies.

Common mistake

You may think that a budget deficit is 'bad' for the economy. It may seem like common sense for a government to aim for a budget surplus or at least a balanced budget, but there may be valid reasons for the government to deliberately spend more than it collects in revenue in order to stimulate the level of aggregate demand, especially during a deep economic recession.

ATL ACTIVITY 25.2

What do you think is the value of the USA's national debt? Visit this website **http://www.usdebtclock.org** to see how close (or far off) you are from the actual figure! Then use the internet to find out the value of the USA's nominal gross domestic product (GDP). To what extent do you feel that US national debt is affordable?

Crowding out (HL only)

When a government runs a budget deficit, it needs to borrow funds. **Crowding out** (or **financial crowding out**) occurs when increased government borrowing (to finance its public sector spending) causes interest rates to rise owing to the higher demand for loanable funds, which causes a reduction in private sector investment expenditure due to the higher costs of borrowing. Hence, although government spending (G) might increase, the consequential increase in interest rates causes private sector investment (I) to fall. Hence, this limits the effectiveness of expansionary fiscal policies.

◆ **Crowding out** occurs when increased government borrowing causes interest rates to rise, thereby reducing private sector investment expenditure.

In Figure 25.12, the supply of loanable funds is determined by the amount of savings in the economy. The price of these loanable funds is represented by the interest rate. At any given point in time, the supply of loanable funds is fixed, that is, perfectly inelastic (unresponsive to a change in the interest rate). Hence, an increase in government demand for borrowed funds to finance its budget deficit leads only to a large increase in interest rates from r_1 to r_2, with no overall increase in total borrowing in the economy. In fact, financial crowding out and the resulting higher interest rates at r_2 mean that private sector businesses are able to borrow only Q_2 loanable funds yet are charged the higher interest rate of r_2.

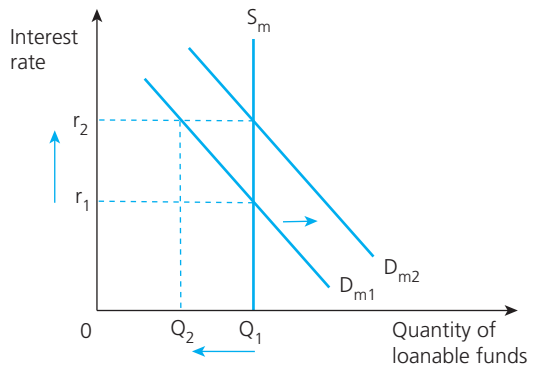


Figure 25.12 The crowding-out effect in the loanable funds market

Top tip!

While reducing government spending might be important at times in order to reduce overall public sector debt, there can be detrimental effects if the cuts are too large, as this can cause the economy to contract or even go into a deep economic depression.

WISE ChoICES

Does the government really know best? To answer this question, consider the extent to which government **intervention** in the economy promotes **efficiency** or is detrimental to **economic well-being**.

Figure 25.13 Government funding can be targeted to support those with health issues

Strengths of fiscal policy

The strengths of fiscal policy include being able to target specific economic sectors and government spending being effective in getting the economy out of a deep recession. HL students also need to be aware of a third strength – the use of automatic stabilizers.

Targeting of specific economic sectors

Whereas monetary policy is indiscriminate, fiscal policy can be used to target specific sectors of the economy. This means that fiscal policy measures can have a greater impact on some areas of the country than others. This can be due to regional disparities in income and spending habits, or simply due to wage differentials between different individuals and households. For example, tax cuts or tax breaks might be applied only to the lowest income earners, while marginal tax rates can be increased only for the highest income earners in the country in order to redistribute income and wealth. Similarly, government spending can be targeted at the unemployed, retirees, those with disabilities or those with young children, for example. By contrast, the use of interest rate policy does not target any specific economic sectors but any household or firm with existing loans or mortgages, or those with savings.

Government spending effective in deep recession

Another key strength of fiscal policy is its effectiveness in promoting economic activity during a recession. As Keynes famously said, governments can 'spend their way out of a recession'. Government spending (G) is a major component of aggregate demand, so even if this means having to run a budget deficit, injecting money into the circular flow of income, combined with the Keynesian multiplier, mean that economic activity increases, ceteris paribus.

For example, the global financial crisis of 2008 saw governments across the world inject trillions of dollars to stimulate their economies. However, the effectiveness of such policies was hindered by the severity and duration of the global recession. The same happened again in 2020 as governments around the world dealt with the largest recession in living memory, caused by the coronavirus pandemic. Government intervention was inevitable to minimize the impacts of the global recession, which caused more than 1.2 billion children to stay at home owing to school closures, and caused hundreds of thousands of businesses to collapse. Very importantly, expansionary fiscal policy can help to increase both consumer and business confidence levels during a recession. This creates further incentives for firms to continue hiring their workers and to have faith in the economy. Government spending may also be required to deal with supply-side causes of instability to the economy, such as oil crises and major natural disasters.

■ **Figure 25.14** Government bailouts were vital during the financial crisis of 2008

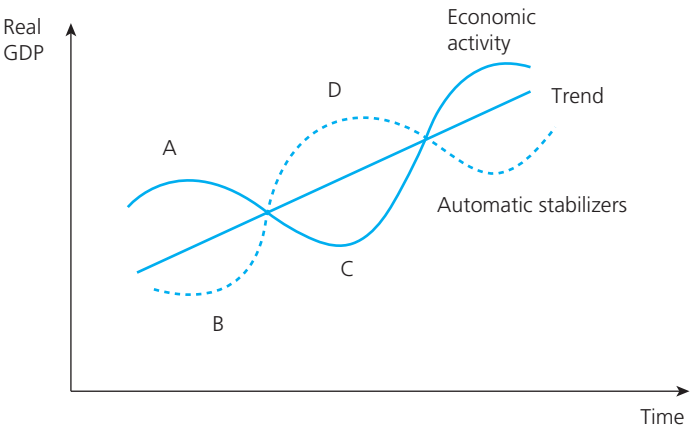
In the short-term, expansionary fiscal policy has a negative impact on the government's budget as revenues fall (from lower tax revenues) while government spending increases. However, as the economy improves in the long-term, households and firms earn and spend more money (helping to raise revenues from income tax and corporation tax) while there is less of a need for the government to increase its own expenditure to stimulate the economy. Therefore, interventionist economists argue that fiscal policy is effective and necessary to get the economy out of a deep recession.

Overall, the effectiveness of fiscal policy can therefore be evaluated by consideration of the constraints and strengths of the policies in meeting the main macroeconomic objectives.

Automatic stabilizers (HL only)

Automatic stabilizers are aspects of fiscal policy that naturally reduce fluctuations in the level of economic activity, thereby stabilizing the rate of growth in real national output. Hence, automatic stabilizers help to even out short-term fluctuations in the level of economic activity, as depicted in the business cycle (see Figure 25.15). The two main types of automatic stabilizers are progressive taxes and unemployment benefits.

◆ **Automatic stabilizers** are aspects of fiscal policy that naturally reduce fluctuations in the level of economic activity, thereby stabilizing the rate of growth in real national output.



■ **Figure 25.15** Automatic stabilizers and the business cycle

In Figure 25.15, if the economy experiences an economic boom (point A) in the business cycle, the use of progressive taxes and lower spending on unemployment benefits (implicitly shown by point B) automatically withdraws money from the circular flow, thereby stabilizing economic activity along the trend line. Likewise, if the economy is experiencing a recession or is at the trough of the business cycle (point C), an automatic increase in government spending on unemployment benefits helps to inject money into the economy (implicitly shown by point D), again stabilizing the economy around the trend line.

Progressive taxes

A progressive tax system charges higher income earners a proportionately higher rate of tax, that is, those who earn more pay a higher rate of income tax. During an economic boom, a progressive tax system means the government automatically receives more tax revenue, from increased tax revenues on earnings (direct taxes) and spending (indirect taxes). Progressive taxes help the government to earn even more money, thereby limiting the extent to which economic activity can grow.

Therefore, progressive taxes as automatic stabilizers help to reduce the rate of economic growth, thereby avoiding the risks of an unsustainable economic boom and uncontrollable inflation in the economy. The resulting budget surplus (from higher tax revenues and lower government spending on unemployment benefits) can be used as reserves for when the economy might go into a recession.

Unemployment benefits

Unemployment benefits are a form of transfer payments or welfare benefits for the unemployed, so they have some money to meet their consumption expenditure needs. Hence, negative economic growth in a recession is minimized by automatic stabilizers, such as more government spending on unemployment benefits helps to limit the fall in economic growth. In addition, both low-income earners and high-income earners might automatically qualify to pay a lower marginal rate of income tax during a recession, thereby minimizing the reduction in consumption expenditure. Hence, increased transfer payments from the government during an economic downturn as automatic stabilizers help to reduce the impact of falling incomes and spending in the economy, thereby avoiding the risks of a deep recession and high unemployment.

By contrast, government spending on welfare benefits automatically falls during an economic boom owing to rising levels of employment. Hence, lower government spending during a boom helps to dampen the increase in the level of economic activity and prevent inflationary pressures as a result.

The effectiveness of automatic stabilizers in reducing fluctuations in economic activity is dependent on:

- the size of the public sector, that is, government spending as a percentage of the country's GDP – the greater the level of government involvement in the economy (through the use of progressive taxes and provision of unemployment benefits), the more effective automatic stabilizers are
- the extent to which the tax regime is progressive – the more progressive the tax system of the country, the more effective automatic stabilizers tend to be
- the scope of the unemployment benefits system – automatic stabilizers will be more effective in countries where social welfare systems are wide-reaching, such as in Sweden, the UK and Australia.

ATL ACTIVITY 25.3

The coronavirus outbreak that began in 2019 caused the world's deepest recession in living memory, with business closures and unemployment in literally every industry in every country in the world. For a country of your choice, investigate the scale of unemployment caused by the coronavirus pandemic and evaluate the extent to which unemployment benefits helped the economy as a whole during the crisis.

Nevertheless, Keynesian economists argue that automatic stabilizers are important in helping to cushion a fall in aggregate demand in an economy and to support national output and jobs during a recession.

Top tip!

When evaluating the effectiveness of fiscal policies, remember why these are being used in the first place – in terms of the strengths and limitations of these policies in promoting economic growth, achieving low unemployment, and reaching a low and stable rate of inflation.

Theory of Knowledge

Keynesians and monetarists disagree about the outcomes of demand-side policies such as the effectiveness of fiscal policies. On what basis might we make judgements about their relative conclusions about the use of demand-side policies?

PAPER 3 EXAM PRACTICE QUESTION 25.3 (HL ONLY)

Study the tax system for the three countries below. Calculate the percentage of tax paid on each level of income and identify whether the tax systems are progressive, regressive or proportional. *[4 marks]*

Tax system A		Tax system B		Tax system C	
Annual income ($)	**Amount of tax paid ($)**	**Annual income ($)**	**Amount of tax paid ($)**	**Annual income ($)**	**Amount of tax paid ($)**
4,500	900	10,000	1,500	8,000	800
10,000	1,800	20,000	3,000	20,000	3,000
20,000	2,800	30,000	4,500	45,000	11,250

Common mistake

It is incorrect to assume that fiscal policy, as a demand-side policy, does not impact the supply-side of the economy. For example, income tax cuts can create incentives for people to seek employment and to work harder. Lower rates of corporation tax can help to attract foreign direct investment in the country, thereby boosting the economy's potential output in the long run.

Top tip!

A useful way to evaluate the effectiveness of fiscal policy in achieving macroeconomic objectives is to consider alternative perspectives. For example, new classical economists argue that the private sector is more efficient in spending money than the government, so demand-management policies cause a decline in economic welfare. They argue that aggregate supply is perfectly price inelastic in the long run, so attempts to increase AD would simply be inflationary. By contrast, Keynesians argue markets are too rigid to correct themselves and require government intervention, especially during a deep economic recession.

PAPER 1 EXAM PRACTICE QUESTION 25.4

Explain how the effectiveness of expansionary fiscal policies depends on the shape of the aggregate supply curve. *[10 marks]*

PAPER 1 EXAM PRACTICE QUESTION 25.5

Explain why expansionary fiscal policy can cause a budget deficit for the government. *[10 marks]*

PAPER 1 EXAM PRACTICE QUESTION 25.6

Using real-world examples, evaluate the extent to which fiscal policy is effective in achieving long-term economic growth. *[15 marks]*

WISE ChoICES

In many countries, individuals must declare and register their earnings (incomes). From the perspective of **economic well-being**, can society ever benefit from tax avoidance or tax evasion? Should people ever have a **choice** over how much tax they pay?

Top tip!

When evaluating the use of fiscal policy, it is worth noting potential conflicts in the government's macroeconomic objectives. For example, the government might choose to use contractionary fiscal policy to deal with cost-push inflation, but unemployment may occur as a consequence. The effectiveness of the policy then depends on the aims (or purpose) and priorities of such policies.

WISE ChoICES

Given fiscal policy involves the use of taxation and government expenditure policies, discuss how **interdependence** and **equity** influence economic decision-making.

Chapter summary

- Fiscal policy is the use of taxation and government expenditure strategies to influence the level of economic activity and to achieve macroeconomic objectives. Government spending (G) and consumption (C) are the two components of aggregate demand most directly affected by fiscal policy.
- The main source of government revenue is from taxation. These can be classified as either direct taxes (levies imposed on an individual's income and wealth or on the profits of businesses) or indirect taxes (charges imposed on expenditure).
- Government spending can be classified as current expenditure, capital expenditure or transfer payments.
- Capital expenditure can help to promote long-term economic growth, such as investments in the country's infrastructure (railways, road networks, airports and telecommunications networks).
- The main reason for taxation is to raise funds for government expenditure, such as the current and capital expenditure on schools, hospitals and transportation networks.
- Government expenditure is also a key component of aggregate demand, so an increase in government spending helps to boost real national output, employment and growth.
- Expansionary fiscal policy is used to reduce the effects of an economic recession by boosting real gross domestic product and reducing unemployment.
- The government budget refers to its financial plans in terms of planned revenues and expenditure. The budget can be in balance, deficit or surplus.
- The Keynesian multiplier shows that any increase in the value of injections into the circular flow of an economy results in a proportionately larger increase in aggregate demand. Hence, an injection into the circular flow will tend to lead to a bigger eventual final effect on national output and, hence, employment. (HL only)
- The theory of crowding out means that government attempts to increase spending through increased borrowing raises interest rates and reduces the loanable funds available for private sector investments, which limits the effectiveness of such expansionary fiscal policies. (HL only)
- Progressive tax systems and unemployment benefits act as automatic fiscal stabilizers of short-term fluctuations in the economy. (HL only)

26 Supply-side policies

Real-world issue

How do governments manage their economy and how effective are their policies?

Conceptual understandings (WISE ChoICES)

- Government **intervention** attempts to achieve macroeconomic objectives through a **choice** of policies.
- Political, economic, social and environmental factors are **interdependent** and will influence the effectiveness of government policies.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- the goals of supply-side policies (AO2)
- the meaning of market-based policies (AO2)
- the following market-based policies to encourage competition: deregulation, privatization, trade liberalization and anti-monopoly regulation (AO2)
- an AD–AS diagram including the LRAS curve to show the effect of supply-side policies (AO4)
- labour market policies as a form of market-based policies (AO2)
- the minimum wage diagram (AO4)
- interventionist supply-side policies (AO2)
- the demand-side effects of supply-side policies (AO2)
- the supply-side effects of supply-side policies (AO2)
- the effectiveness of supply-side policies, with reference to the constraints on and strengths of supply-side policies in promoting growth, low unemployment, and low and stable rate of inflation (AO3).

Goals of supply-side policies (AO2)

Supply-side policies are long-term government strategies used to increase the potential productive capacity of the economy by increasing the quality and/or quantity of factors of production. Recall from Chapter 16 that the potential output of the economy refers to the country's productive capacity if all factors of production are used efficiently, that is, the maximum possible national output if there are no unemployed resources and if they are all used efficiently. Supply-side policies are used because increases in the productive capacity of the economy can be achieved only through an increase in the economy's long run aggregate supply.

◆ **Supply-side policies** are long-term government strategies used to increase the potential productive capacity of the economy by increasing the quality and/or quantity of factors of production.

Supply-side policies are designed to make markets and industries more productively efficient in the long run, thereby contributing to an increase in the country's real national output. Hence, many economists believe that improved supply-side performance is vital for the economy to achieve sustained economic growth and development without the threat of inflation (see Chapter 20). Examples of such policies include spending on education, healthcare and training to improve the economy's human capital, which enhances labour productivity and labour mobility.

There are five main goals of supply-side policies: long-term economic growth by increasing the country's productive capacity, improving competition and efficiency, reducing labour costs and unemployment through labour market flexibility, reducing inflation to improve international competitiveness, and increasing incentives to invest in innovation by reducing costs. These goals can be remembered by the acronym **i-CLIP**:

Figure 26.1 The goals of supply-side policies

Long-term growth by increasing the economy's productive capacity

Supply-side policies aim to increase the production side of the economy, thereby shifting its production possibility curve (PPC) or the long run aggregate supply (LRAS) curve outwards. This is achieved by improving the institutional framework (the country's system of established structures and contexts that affect economic behaviour, such as the nation's culture and its legal system) and increasing the economy's potential productive capacity to produce goods and services (by increasing the quantity and/or quality of the factors of production) to achieve long-term growth. Recall that the LRAS refers to the productive capacity of an economy (or its potential output, as shown in Figure 26.2). This can also be shown by a PPC diagram (see Figure 26.3).

In Figure 26.2, supply-side policies are used to shift the long run aggregate supply of the economy from $LRAS_1$ to $LRAS_2$, thereby increasing the productive capacity of the economy from Y_{f1} to Y_{f2}. A major advantage of such growth through increasing the economy's productive capacity is that the general price level also falls, from PL_1 to PL_2.

In Figure 26.3, supply-side policies are used to shift the economy's production possibility curve from PPC_1 to PPC_2. This increases the economy's capacity to produce more goods and services, from point A on the PPC_1 curve to point B on the PPC_2 curve.

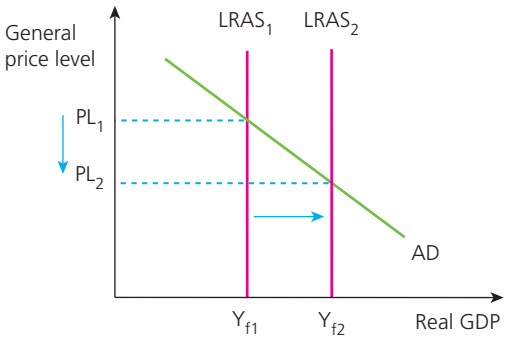


Figure 26.2 Increasing the economy's productive capacity (LRAS)

Improving competition and efficiency

Supply-side policies also aim to improve the degree of competition and therefore efficiency in certain markets. For example, trade liberalization policies can be used to encourage freer and fairer international trade, which leads to more competition in the market. Trade liberalization is the removal or reduction of trade restrictions or barriers (see Chapter 27). This creates incentives for firms to be efficient in order to remain attractive to consumers, for example, by keeping prices relatively low and/or improving the quality of the goods and service they supply. Competition can also stimulate investments in research and development, which increases the potential output of the economy, as illustrated in Figure 26.2. The benefits of improving competition and efficiency in markets are discussed in Chapter 15.

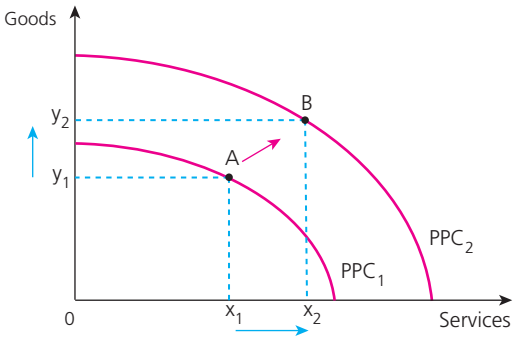


Figure 26.3 Increasing the economy's productive capacity (PPC)

Reducing labour costs and unemployment through labour market flexibility

Labour market reforms as a form of supply-side policies are used to increase labour market flexibility and efficiency so that the workforce is more productive and competitive. For example, labour market imperfections are created by labour unions that push the government to introduce or raise the national minimum wage, above the market equilibrium wage rate (see Figure 26.6 on page 408). This increases the labour costs for businesses, which can cause unemployment in the economy. Hence, supply-side policies aim to reduce labour costs and unemployment through methods that promote greater labour market flexibility, such as by reducing or removing minimum wages and the power (influence) of labour unions.

Reducing inflation to improve international competitiveness

As shown in Figure 26.2, supply-side policies help to increase the productive capacity of the economy, thereby increasing the economy's LRAS curve and allowing the general price level to fall, from PL_1 to PL_2. Thus, supply-side policies aim to reduce domestic inflation in order to improve the country's international competitiveness. This might be achieved, for example, by removing minimum wages to cut labour costs or by trade liberalization policies to increase the degree of competition in export and import markets.

Increasing firms' incentives to invest in innovation by reducing costs

Supply-side policies aim to create or increase incentives for businesses to invest in innovation. This helps the economy to achieve greater levels of potential output in the long run, which can be illustrated by the outwards shift of the PPC from PPC_1 to PPC_2 in Figure 26.3). Driving forces that motivate firms to increase their investment in innovation include lower business taxes, including reduced corporation tax on company profits and capital gains tax on the returns from investment projects. Tax incentives and/or government subsidies to encourage research and development can also be used to encourage foreign direct investment (FDI) in the economy.

Market-based policies (AO2, AO4)

Supply-side policies are often categorized as either **market-based** or **interventionist**, although both aim to increase the economy's potential national output. Market-based supply-side policies are designed to make markets work more efficiently. This approach gives the private sector more autonomy in the allocation of resources. By contrast, interventionist policies favour direct government involvement in order to correct market imperfections and market failures thereby improving productivity.

Market-based supply-side policies focus on freeing up industries and improving market incentives in order to increase aggregate supply in the economy. Examples of market-based supply-side policies include policies to encourage competition and efficiency, labour market reforms to increase aggregate supply, and incentive-related policies, including income tax reforms to improve investments, productivity and incentives to work.

◆ **Market-based supply-side policies** focus on freeing up markets and improving market incentives in order to increase aggregate supply, thereby improving incentives to increase investments and productivity and incentives to work.

Policies to encourage competition

In theory, competition helps to improve economic efficiency and innovation, thereby increasing the potential output of the economy. The four main policies to achieve this are deregulation, privatization, trade liberalization and anti-monopoly regulation.

Deregulation

Deregulation is the reduction or removal of statutory barriers to entry into certain industries thereby allowing markets to operate more competitively and efficiently. For example, getting rid of excessive rules and directives in a certain industry helps to promote healthy competition (thereby increasing supply) and greater productive efficiency. For example, the USA deregulated its airline industry in the late 1970s. The UK deregulated many of its industries in the 1980s, including gas, steel, transportation and telecommunications. Financial markets around the world were deregulated in the 1990s, allowing financial institutions greater flexibility in the types of financial products they offered and the fees they charged for their services.

◆ **Deregulation** is the reduction or removal of government rules and directives in an industry, with the intention of creating more competition and greater efficiency.

Deregulation reduces the market power of firms that face little or no competition, such as state-owned industries. For example, the UK's Royal Mail had a legal monopoly on the delivery of letters and parcels up to 2006 when the industry was deregulated, allowing courier firms such as DHL and Parcelforce to compete. The process of deregulation is often, but not always, accompanied by privatization.

There are two main advantages of deregulation:

- Increased competition and efficiency should lead to lower prices and greater choice for consumers. This was particularly noticeable with the entry of budget (low-cost) airlines and falling prices in the telecommunications industry.
- Regulation often involves excessive costs of bureaucracy. Such processes are also time consuming to administer.

However, deregulation does not guarantee more competition or efficiency as it can be difficult to create meaningful competition in an industry dominated by a natural monopoly (see Chapter 15) owing to extreme barriers to entry. Increased competition is not always beneficial to society in the case of wasteful competition. For example, the deregulation of public bus services has often led to a duplication of services and negative externalities of traffic congestion.

■ **Figure 26.4** Airline industries around the world have been deregulated to increase competition

Privatization

Privatization is a form of supply-side policy that involves the sale or transfer of public sector assets (such as state-owned enterprises or government housing stocks) to the private sector. The main purposes of privatization are to improve competition, efficiency and productivity in the economy, although selling off state-owned assets also helps to raise government revenue. It can encourage firms to become more efficient in order to survive and to earn a profit. In addition, selling off state-owned assets also cuts government expenditure on maintaining these assets or businesses, while raising one-off revenues from privatization proceeds for the government.

◆ **Privatization** is a form of supply-side policy that involves the sale or transfer of state-owned assets to the private sector, in order to improve competition, efficiency and productivity.

ATL ACTIVITY 26.1

In small groups, discuss whether countries should privatize their postal services and telecommunications services in the twenty-first digital century?

ATL ACTIVITY 26.2

Do you think that more schools and universities in your country should be privatized? Consider the economic and social impacts of such a decision on different stakeholders: consumers (students), workers (parents), firms (educational establishments), the government and the economy (society) as a whole.

The main advantages of privatization include the following arguments:

- Private sector businesses have the incentive to improve their operational efficiency because they need to remain competitive. This can also promote an entrepreneurship culture of risk-taking and innovation in the economy.
- The ability to earn one-off privatization proceeds from the sale of state-owned assets.
- Privatization helps to reduce government debt in three ways: the government no longer has to maintain the operation of state-owned enterprises, it earns revenue from the sale of the privatized businesses, and the private sector firms pay corporate taxes on an annual basis.
- It also reduces costs to taxpayers who no longer have to pay to finance the operations of the privatized businesses.

The main disadvantages of privatization include:

- The process can create a private sector monopolist, which is not always a positive outcome for customers who may face higher prices.
- To protect the public interest, privatization may still require government regulation and intervention, through methods such as price controls (see Chapter 10).
- There is a trade-off (opportunity cost) as state-owned enterprises can provide social benefits, which may no longer exist if the like of social housing, postal services and public transportation are privatized.

CASE STUDY

Privatization in the UK

Margaret Thatcher, the UK's first female prime minister, started the UK's privatization revolution, selling state-owned enterprises including: British Telecom (1984), Jaguar (1984), British Gas (1986), British Airways (1987), Rolls-Royce (1987) and British Steel (1988). Her privatization programme raised more than £29 billion ($38bn) from the sale of nationalized industries.

Despite the arguments in favour of privatization, the 2008 global financial crisis caused many banks to be bailed out by the government, making nationalized industries popular as a method to correct market failures and inefficiencies in the economy. Nationalization is the opposite of privatization, that is, the purchase of private sector assets by the government, bringing these into the ownership and control of the public sector. This can protect employment and promote economic stability in key industries. Industries that are commonly nationalized include public transport, telecommunication networks, energy, utilities and natural resources. However, there is a rather large opportunity cost to public sector ownership of such assets. For example, according to *The Guardian*, the UK spent around 20 per cent of its GDP on bank bailouts following the global financial crisis.

Trade liberalization

Trade liberalization occurs on an international scale, involving the reduction or removal of barriers to international trade, such as tariffs, quotas and subsidies. The purpose is to encourage more competition and greater efficiency in export and import markets. Improved trade relations and policies are needed to boost the level of international trade and foreign direct investment (FDI) in the economy. For example, it may involve allowing the free movement of capital flows between countries. This helps to create freer and fairer international trade between different nations by increasing competition, productivity and efficiency.

◆ **Trade liberalization** is the reduction or removal of trade barriers such as tariffs and quotas, thereby increasing competition, productivity and efficiency in international markets.

By contrast, protectionist policies (see Chapter 28) use trade barriers to safeguard domestic jobs from the threat of international competition. For example, Japan imposes up to 778 per cent import taxes on rice – the highest rate in the world – to protect agricultural jobs in its country. However, most economists agree that trade liberalization is generally better for long-term economic prosperity of the world economy. The reasons for trade liberalization are covered in more detail in Chapter 27.

Anti-monopoly regulation

Anti-monopoly regulation refers to laws that control or limit the restrictive practices and market power of dominant firms in an industry. Such regulations protect the interest of consumers against anti-competitive behaviours of monopolies. For example, firms with monopoly power have the ability to set higher prices than in competitive markets. Therefore, anti-monopoly regulations are used to promote or enforce competition in targeted industries to benefit consumers.

◆ **Anti-monopoly regulation** are laws that control or limit the restrictive practices and market power of dominant firms in an industry.

CASE STUDY

Anti-competitive supermarket pricing strategies

Large supermarket chains such as Walmart, Carrefour and Tesco use an extensive range of pricing and non-pricing strategies to compete. Market research published by *Which?* (a consumer campaigning organization in the UK) showed that most shoppers believe large supermarkets deliberately try to mislead them by using confusing pricing strategies. The *Which?* findings showed that supermarkets often use poor labelling and puzzling prices for their products. For example, without using a calculator, do you think a 600 g jar of mayonnaise priced at $3.49 is better value than a 400 g jar of the same brand but priced at $2.35? Which product offers better value? Similarly, pre-packed fruits and vegetables often have very different prices compared to loose varieties of the same produce.

Figure 26.5 Large supermarkets have the power to manipulate prices

Examples of anti-monopoly regulations include:

- Investigations by government agencies, such as the Competition Commission (see Chapter 14), into anti-competitive behaviour and unfair practices of firms with market power.
- Price controls such as a price ceiling or maximum price (see Chapter 10) can be used to limit the price that monopolies can charge. Alternatively, the government can use price caps to limit how much prices are allowed to increase by.
- Regulation of mergers and takeovers exist to prevent any potential growth of firms that create anti-competitive monopoly power. In many countries, if a merger or takeover creates a firm with market share of at least 25 per cent, it can be investigated by the authorities. For example, the UK's Competition and Markets Authority (CMA) blocked the proposed merger between Sainsbury's and Asda, two of the country's largest supermarkets, as this was deemed to be against the public's interest. The CMA also has the authority to break up monopolies if it is judged to be necessary.

Top tip!

You are expected to be able to use an AD–AS diagram and the LRAS curve to show the effects of supply-side policies. Refer to Figure 26.2 and reflect on how you can use this to show the impact of market-based supply-side policies on the economy.

PAPER 2 EXAM PRACTICE QUESTION 26.1

Explain why privatization of state-owned enterprises is a form of market-based supply-side policy. *[4 marks]*

Labour market policies

In addition to policies to encourage competition, market-based supply-side policies also include labour market policies. **Labour market policies** (or **labour market reforms**) are measures and reforms developed to create greater flexibility and efficiency in the labour market, such as reducing the power of labour union laws, reducing unemployment benefits to create incentives to work and abolishing minimum wages.

◆ **Labour market policies** are government policies designed to create greater flexibility and efficiency in the labour market.

Reducing the power of labour unions

Labour unions (also known as **trade unions**) are organizations that aim to protect the interests of their worker members. In particular, labour unions strive to safeguard and improve the terms and conditions of employment, including pay and benefits. They are organizations recognized as legal representatives of workers, thereby acting in the best interest of their members. They originated in the nineteenth century in the UK and the USA. People often worked in very poor conditions, so trade unions were created to bargain for better terms and working environments for their members. A worker becomes a member of a trade union by paying a subscription fee, usually on a yearly basis. The membership fees help to pay for the administrative and legal expenses of operating the trade union.

◆ A **labour union** is an organization that aims to protect the interests of its worker members, that is, the terms and conditions of employment, such as workers' pay and benefits.

However, labour unions are often criticized for causing labour market imperfections. There is often a conflict between the aims of employees and employers. For example, an increase in pay and benefits, along with pressures to reduce working hours, will improve the terms of employment for workers but increase costs for employers. Labour organizations also push for the

legal protection of employee rights, such as to impose a national minimum wage or to increase the statutory (legal) minimum wage. However, this can create inefficiencies in the labour market and cause domestic firms to be uncompetitive. Hence, labour market reforms by reducing the power of labour unions can help to make the economy more competitive.

ATL ACTIVITY 26.3

Investigate the number and types of unions that exist in your country or a country of your choice. Find a recent example of industrial action that has taken place and determine whether the action was successful or not in achieving its aim. What does this tell us about the degree of power of labour unions in the country?

Reducing unemployment benefits

Unemployment benefits are welfare transfer payments made by a government to support people who are willing to work but cannot find employment. The existence of generous welfare payments, such as unemployment benefits, can give some people an incentive or excuse not to work. Hence, by reducing or removing these benefits, the government can create incentives for the unemployed to seek employment. Alternatively, the government can make it more difficult for people to qualify or claim unemployment benefits, so that more people become proactive in searching for job opportunities. In any case, supply-side economists argue that reducing unemployment benefits helps to increase greater flexibility and efficiency in labour markets.

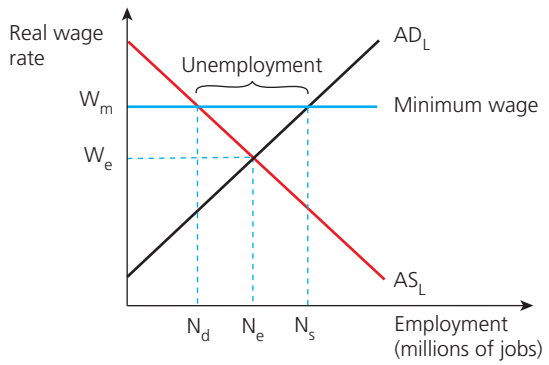
◆ **Unemployment benefits** are welfare transfer payments made by a government to support people who are willing to work but cannot find employment.

Abolishing minimum wages

A **minimum wage** is the lowest rate of pay that firms must pay to their employees per time period, as governed by the government. While a minimum wage helps low-income workers to enjoy a wage above the market equilibrium (see Figure 26.6), it causes higher costs for employers. In the long-term, this can have a negative impact on their profits and, hence, reduce the level of real national output and employment in the economy.

◆ A **minimum wage** is the lowest rate of pay that firms must pay to their employees per time period, as governed by the government.

In Figure 26.6, the market equilibrium wage rate is W_e with N_e millions of people in employment. The introduction or imposition of a minimum wage at W_m causes an increase in labour costs, so contracts the aggregate demand for labour (shown by the AD_L curve) from N_e to N_d. However, the higher wage rate attracts more people to the labour market, expanding the aggregate supply of labour (shown by the AS_L curve) from N_e to N_s. Hence, as $AS_L > AD_L$ at the minimum price, labour market disequilibrium is created leading to unemployment caused by higher labour costs. Therefore, supply-side economists argue that an underlying cause of imperfections in the labour market is high wage rates due to minimum wages. Hence, the reduction or removal of minimum wages can increase incentives for people to work. However, some economists believe that having minimum wages can increase real income and therefore increase AD, thus causing more jobs through higher derived demand for labour.



■ **Figure 26.6** Impact of a minimum wage on the labour market

Incentive-related policies

The third category of market-based supply-side policies is **incentive-related policies**, which aim to create greater incentives to work. They are driving forces that motivate people to seek employment opportunities, such as lower rates of personal income tax, business tax and capital gains tax. Cutting taxes on personal income creates incentives to work while lowering business taxes encourages more investment expenditure in the economy, in order to increase the economy's long run aggregate supply.

◆ **Incentive-related policies** are a form of market-based supply-side policies that create greater incentives to work, namely by cutting personal income tax and business taxes.

Personal income tax cuts

Personal income tax is a form of direct tax levied on the total earnings of an individual after certain tax allowances are considered. Reducing personal income tax for those earning low incomes in particular, can create incentives for people to work or to seek employment opportunities. Hence, such market-based supply-side policies are designed to reduce the level of unemployment, thereby increasing the economy's national output in the long run. For example, personal income tax rates in Hong Kong range from 2 per cent to 17 per cent, whereas they are as high as 57.2 per cent in Sweden. In theory, this means the lower personal income taxes in Hong Kong create greater incentives for people to work.

Cuts in business tax and capital gains tax (CGT)

Cuts in business taxes are used to create incentives for firms to invest by reducing the financial risks and returns from investments. For example, a cut in corporation taxes (imposed on the level of company profits) enables firms to retain a greater proportion of their profits. This helps to create more wealth, growth and employment in the economy. Similarly, cuts in the rate of capital gains taxes (imposed on the profit earned on a fixed asset once it is sold) can also encourage risk-taking and investments in the economy, such as the purchase of shares, land and commercial property. For example, Hong Kong's business tax rates range from 7 per cent to 16.5 per cent whereas it is 21.4 per cent in Sweden. There is no CGT in Hong Kong, whereas CGT is 30 per cent in Sweden. There is also no tax levied on dividends or capital gains from holding shares in a Hong Kong company. Hence, all things being equal, this makes Hong Kong relatively attractive to businesses.

Figure 26.7 Hong Kong is renowned for its low rates of direct taxes

Interventionist policies (AO2)

Interventionist supply-side policies are the deliberate attempts by a government to deal with market imperfections in the economy in order to influence aggregate supply and to increase the productive capacity of the economy. These policies include education and training, improving the quality, quantity and access to healthcare services, research and development (R&D), the provision of infrastructure and industrial policies. Interventionist supply-side policies rely on an efficiently functioning institutional framework – the established systems, structures and contexts that shape the economic behaviour in a country, including the legal system.

◆ **Interventionist supply-side policies** are the deliberate attempts by a government to influence aggregate supply and the productive capacity of the economy.

Education and training

Human capital refers to the stock of knowledge, skills, expertise and experiences of the labour force. An important interventionist supply-side policy is to increase human capital by spending more on education and training to raise the skills, mobility and productivity of the labour force.

Training can enhance the demand for and supply of labour by making the workforce more skilled and productive. Training also helps to improve labour mobility, so that workers can switch between jobs or occupations with relative ease. Efforts to improve communications in the job market can also help to reduce frictional (search) unemployment.

◆ **Human capital** is the accumulation of skills, knowledge and experience of the workforce, viewed in terms of their value or cost.

Figure 26.8 Investment in education is a form of interventionist supply-side policy

The government can also provide financial incentives to firms for training and hiring the long-term unemployed. This might include, for example, the government offering tax allowances and/or subsidies to reduce the costs of training and hiring workers for businesses participating in the scheme.

Investment in education and training also helps unemployed people to gain new skills so that they can find employment. For example, retraining people who are structurally unemployed in certain manufacturing industries enables people to find jobs in other sectors or industries.

Overall, investment in human capital increases the national income in the short run as higher aggregate demand raises national expenditure. In the long run, investment in education and training improves the productive capacity of the economy, shown by an increase in the country's long run aggregate supply (see Figure 26.2).

Improving quality, quantity and access to healthcare

A second form of interventionist supply-side policies is financing public healthcare services, such as hospitals and medical clinics. This helps to improve the quality and quantity of access to healthcare in the economy. Healthcare provision is vital for improving the supply side of the economy as businesses can face substantial costs from time lost to ill-health of their employees.

Healthcare spending is reported as a percentage of the nation's GDP or by the number of doctors and nurses per 1,000 people of the population (see Table 26.1). The greater these values are, the more likely it is that the nation's healthcare improves labour productivity and national output.

Table 26.1 Doctors and nurses per 1,000 people (selected countries), 2020

Country	Doctors (per 1,000 people)
Norway	5.53
Lithuania	4.85
Germany	4.64
Switzerland	4.37
Sweden	4.32
Canada	2.76
USA	2.74
Poland	2.58
Japan	2.50
Turkey	1.87

Country	Nurses (per 1,000 people)
Norway	20.83
Germany	16.70
Finland	15.84
Iceland	14.85
Australia	12.98
Israel	5.89
Slovakia	5.65
Poland	5.14
Latvia	4.59
Turkey	2.07

Sources: Trading Economics https://tradingeconomics.com/country-list/nurses, https://tradingeconomics.com/country-list/ medical-doctors

Improved healthcare can also come from discouraging unhealthy habits. For example, indirect taxes can be levied on the sale of cigarettes, alcohol and sugar in order to reduce the healthcare costs associated with second-hand smoking, drunkenness and obesity in society.

Research and development (R&D)

Research and development (R&D) refers to the process of business activities to improve, introduce and innovate products, processes and procedures. The R&D process creates new knowledge for commercial purposes. Spending on R&D can improve work processes, thereby enhancing efficiency and productivity. For example, the development of wireless broadband internet technologies has helped to improve operations in the workplace and at home. By contrast, countries that spend only a small proportion of their GDP on R&D usually find this to be a constraint on innovation so negatively affects their long-term global competitiveness.

◆ **Research and development** (R&D) is the process of business activities to improve, introduce and innovate products, processes and procedures.

Figure 26.9 R&D in the car manufacturing industry helps to keep companies and countries competitive

Policies that encourage investment in R&D have a short-term effect on aggregate demand because such expenditure will increase AD, ceteris paribus. These policies include low interest rates and tax rebates for domestic firms and can also encourage foreign direct investment in the economy. More importantly though, in the long run, investment in R&D can increase the productive capacity and productivity of the economy as a whole, shown by a rightwards shift of the country's LRAS curve (see Figure 26.2).

R&D can also generate new products for consumption and export, such as energy-saving light bulbs, wearable technologies, electric cars, smartphones and tablet computers. Therefore, investments in R&D and new technologies can be an important source of international competitive advantage.

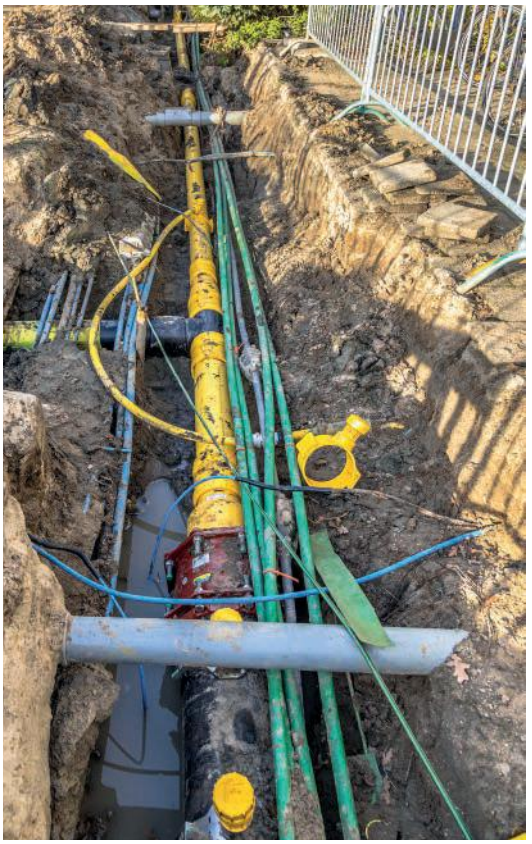


Figure 26.10 Cables, pipes and sewage systems are essential to an economy's infrastructure

Provision of infrastructure

Increasing investment in a country's provision of infrastructure helps to boost its productive capacity, which can be illustrated by an outwards shift of the country's production possibility curve (see Figure 26.3). **Infrastructure** refers to the physical and organizational structures and facilities of an economy necessary for the operations of society as a whole, including transportation networks, telecommunications networks, electricity grids and power supplies, waste disposal systems and sewage systems.

The provision and investment in infrastructure help to facilitate more efficient and productive output, in a more cost-efficient manner. Improved infrastructure also attracts more investment from both domestic firms and foreign multinational companies. In the short run, government expenditure on infrastructure will lead to a rise in aggregate demand thus boosting economic growth but may cause some inflation, ceteris paribus. This is because investment (I) is a key component of aggregate demand (recall that AD = C + I + G + X – M). In the long run, government expenditure on infrastructure will lead to greater efficiencies and national output. This is illustrated by shifting the LRAS curve to the right (see Figure 26.2).

However, government spending on improved infrastructure is highly expensive. Also, it can be very challenging to increase transportation or infrastructural capacity such as roads, railways and airports in busy city centres. Nevertheless, the government has key role in investing in the country's infrastructure such as constructing and maintaining roads, bridges and motorways. This also helps to create many jobs in the construction and related industries.

◆ **Infrastructure** is the physical and organizational structures and facilities of an economy necessary for its operations, such as transportation networks, telecommunications networks, power supplies and waste disposal systems.

Industrial policies

Industrial policies are those that target specific key industries to promote economic growth and employment. These are interventionist supply-side policies aimed at encouraging the development and growth of the manufacturing sector of an economy. For example, tax allowances (rebates) can be used to protect domestic industries from larger, well-established foreign rivals. A combination of tax breaks and lower interest rate charges on commercial loans can create incentives for firms to locate in less prosperous areas of a country, thus reducing unemployment and increasing the economy's long run aggregate supply. Subsidies can also be given to firms that hire youth workers, mature staff and discouraged workers (those suffering from long-term unemployment). Subsidies might also be offered on commercial loans to encourage business start-ups.

◆ **Industrial policies** are interventionist supply-side policies aimed at encouraging the development and growth of the manufacturing sector of an economy in order to promote growth and employment.

Another example is government assistance to help unemployed people to find jobs. For example, frictional unemployment (see Chapter 19) can be reduced by improved methods to match job seekers with suitable jobs in the economy. With advances in online technologies and the use of social media websites (such as LinkedIn and other job posting websites), it has become much quicker to match job seekers with appropriate job openings, thereby reducing the frictional unemployment rate. Government retraining programmes also help to improve the occupational mobility of the structurally unemployed. Hence, such interventionist supply-side policies involve direct government involvement or intervention in order to increase the economy's productive capacity.

Industrial policies, like all supply-side policies, can improve the economic well-being of society in terms of reduced unemployment, higher incomes and increased consumption expenditure for those working in these and related industries. Hence, these policies also help to improve the likelihood of achieving sustainable economic growth and development.

Common mistake

You may think that industrial policy is used primarily to increase competition in an economy. This is incorrect as industrial policy is an interventionist supply-side policy.

Table 26.2 Summary of market- and interventionist-based supply-side policies

Market-based	Interventionist-based
Government income tax reforms in order to improve incentives to work	Government finance for infrastructure programmes
Government reducing the power and influence of labour unions	Direct support via industrial policies to help targeted businesses and markets
Government subsidies to encourage firms to relocate to areas with high unemployment	Financing public healthcare services, such as hospitals and medical clinics
International trade deals and trade liberalization policies	Government funding for schools and universities
Privatization of government-owned assets and enterprises	Government provision of transportation programmes, including railroads and road networks
Reducing or removing minimum wages to reduce labour market rigidities (inflexibilities)	Government retraining programmes to improve the labour mobility of those in structural or cyclical unemployment
Tax allowances or tax discounts to encourage foreign direct investment	Tax incentives and/or subsidies to encourage R&D

Common mistake

You may be tempted to think that supply-side policies are facilitated or funded by the government. While this might be true for interventionist and labour market supply-side policies, this is not necessarily the case for all supply-side policies. Supply-side policies can be implemented by the private or public sector, such as in the case of investments in training and developing the labour force.

Demand-side effects of supply-side policies (AO2)

The IB Economics syllabus distinguishes between demand-management (fiscal and monetary policies) and supply-side policies. However, it is incorrect to assume that supply-side policies have no impact on aggregate demand in the economy, and vice versa. By definition, supply-side policies are long-term strategies used to increase the potential output of the economy. However, some of these policies will have demand-side effects on the level of aggregate demand, while some demand-side policies will have supply-side effects too (see Chapters 23 and 24).

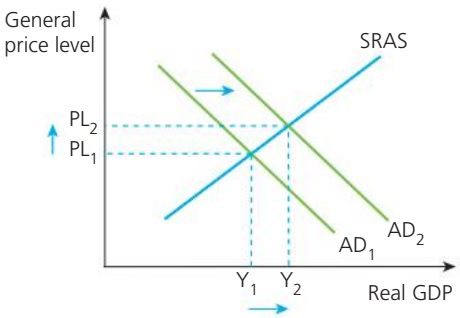


Figure 26.11 Demand-side effects of supply-side policies (short run)

While the benefits of supply-side policies, such as greater investment in education or R&D, may take many years to materialize, when they are implemented there are likely to be some demand-side effects. For example, tax cuts on personal income and business profits (being a form of interventionist supply-side policies) help to increase household disposable income and increase the level of retained profits for firms. Hence, these tax incentives will have an expansionary fiscal effect on both consumption (C) and investment (I), thereby boosting the level of aggregate demand. Hence, supply-side policies can be used to achieve economic growth and lower unemployment, even though this may come at the opportunity cost of increased inflation (see Figure 26.11).

In Figure 26.11, an increase in government spending (G) on education and healthcare increases aggregate demand (AD) from AD_1 to AD_2. This causes an expansion along the short run aggregate supply (SRAS) curve, raising real national output from Y_1 to Y_2. This comes at the cost of higher average prices in the economy, with the general price level rising from PL_1 to PL_2.

Note that all interventionist supply-side policies (education, training, healthcare, research and development, infrastructure and industrial policies) have a demand-side effect because they all involve taxation and/or government spending. Hence, the increased government spending (G) on the above measures represents an increase to aggregate demand.

Top tip!

Remember that demand-side policies can have a direct impact on the economy's long run aggregate supply. For example, expansionary monetary policy in the form of lower interest rates – while being classified as a demand-side policy – can influence aggregate supply by creating incentives for domestic firms to invest and attract foreign direct investments. In reality, aggregate demand and aggregate supply do not work in isolation.

Supply-side effects of fiscal policies (AO2)

While the implementation of supply-side policies can influence the demand side of the economy quite quickly, there is generally a lengthy time lag between their implementation and their effects on the economy's potential output or long run aggregate supply. While supply-side policies are long-term strategies aimed at increasing the productive capacity of the economy, expansionary fiscal policy measures such as an increase in government spending on education and healthcare can certainly increase consumption and investment expenditure, so therefore directly affect the level of aggregate demand (see Figure 26.12). Similarly, government spending on research and development will not only shift aggregate demand in the short run, but will also increase the quantity and quality of factors of production thereby increasing the economy's potential output in the long run.

In Figure 26.12, expansionary fiscal policy (a form of demand-side management) in the form of lowering personal income taxes and business taxes may increase the incentive for people to work harder and for firms to invest (as they become more profitable). Therefore, this increases the economy's productive capacity from Y_{f1} to Y_{f2} in the long run. This causes an expansion along the aggregate demand (AD) curve, while allowing the economy to also benefit from lower general prices in the economy, which fall from PL_1 to PL_2.

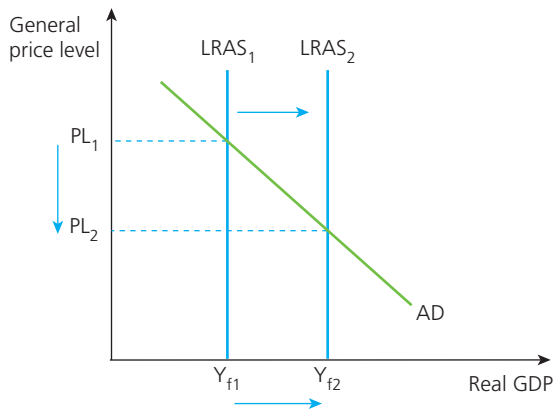


Figure 26.12 Supply-side effects of fiscal policies (long run)

Top tip!

The effects of supply-side policies can be remembered by the mnemonic EPIC:

- Supply-side policies improve economic **efficiency**.
- They increase labour and capital **productivity**.
- They provide **incentives** to work and invest.
- They make firms and the country as a whole more **competitive**.

PAPER 1 EXAM PRACTICE QUESTION 26.2

Explain how a cut in taxes may be considered as both a form of demand-side policy and a supply-side policy. *[10 marks]*

Effectiveness of supply-side policies (AO3)

The effectiveness of supply-side policies depends on the constraints and strengths of these policies in achieving the macroeconomic objectives of growth, low unemployment, and low and stable inflation.

Constraints on market-based supply-side policies

Equity issues

Market-based supply-side policies do not necessarily improve equity in the distribution of income in the economy because economic growth can create greater disparities and inequalities in income distribution (see Chapter 22). For example, a reduction in personal income tax rates is likely to benefit higher income earners more and those on lower incomes, thereby causing greater inequalities in income and wealth. Similarly, reducing business taxes is likely to benefit large multinational corporations in the form of larger profits, thus benefiting wealthy shareholders. This also leads to a larger income and wealth gap within the economy.

Time lags

A major criticism of all supply-side policies is the time that it takes for the economy to reap the benefits and to see any real impact on the country's potential output. For example, it is likely to take significant time to plan, execute and review deregulation of labour markets and industries. The same applies to the process of privatizing certain industries, policies to increase competition (such as anti-monopoly legislation) and trade liberalization. While supply-side policies tend to have more permanent impacts on employment, they take longer to accomplish compared to demand-side policies aimed at reducing unemployment in the economy.

Vested interests

A vested interest refers to an individual or group stakeholder with a personal reason for being involved in a particular policy, decision or situation. The extent to which certain supply-side policies are used will depend, at least to some extent, on the political and philosophical aims of the government in power. For example, governments may have a vested interest in certain privatized industries or infrastructural projects. The Hong Kong SAR government, for example, has a public–private partnership agreement with Hong Kong Disneyland, with the government owning 51 per cent of the business with the Walt Disney Company owning the remaining shares. Although the Hong Kong SAR government also privatized its rail services, it remains a major

Theory of Knowledge

Given all stakeholders have their own vested interests, how can we truly know whether the government acts in our best interest?

shareholder in the MTR Corporation, which has enjoyed monopoly power. This means there may be no incentive for the rail provider to reduce costs or prices.

Environmental impact

Supply-side policies strive to increase the potential output of the economy, but this can come at a huge opportunity cost to the natural environment (see Chapters 11 and 12). For example, deregulation and trade liberalization policies may have some negative effects on the environment if environmental regulations are relaxed or if multinational companies operate in countries where environmental protection laws are limited. Hence, the pursuit of growth and employment comes at the opportunity cost of the environment.

WISE ChoICES

Labour market reforms can lead to a reduction in living standards for low-income earners if there is a reduction in labour union powers, the removal of minimum wages, relaxed health and safety regulations, and zero-hours contract working. Considering these reforms to make the labour market operate with greater **efficiency**, to what extent can supply-side policies truly improve **economic well-being**?

Constraints on interventionist supply-side policies

Costs

A major constraint on interventionist supply-side policies is the huge costs of finance involved. For example, the government's budget is likely to be stretched and possibly going into deficit to fund better education and training (to improve labour productivity), improved transportation and infrastructure networks, and funding public healthcare services. The increased government expenditure may need to be funded through public sector borrowing, which increases the overall value of the national debt (see Chapter 21). Furthermore, there is an opportunity cost to

Figure 26.13 There are major time lags involved with interventionist supply-side policies

government spending. For example, by choosing to spend more money on improving the economy's infrastructure the government may need to relinquish expenditure on healthcare or education.

Time lags

Another major criticism or limitation of interventionist supply-side policies is the time that it takes to reap the benefits. These are more significant than the time lags for market-based supply-side policies. For example, it might take decades for a nation to enjoy the benefits of an improved education system or transportation and infrastructure networks in the country. So, during a recession, for example, such supply-side policies cannot tackle the fundamental problems related to the lack of aggregate demand in the economy.

Strengths of market based supply-side policies

Improved resource allocation

Supply-side policies can be used to achieve sustainable economic growth by increasing the potential capacity of the economy over time by making more efficient use of the economy's scarce resources, including labour resources. An increase in the economy's productive capacity (see Figure 26.3) will increase national output, thereby creating jobs in the economy in the long-term. Supply-side policies can also help to reduce both frictional and structural unemployment by improved labour market reforms. As supply-side policies increase in the productive potential of the economy, they also help to prevent the general price level from rising beyond control. This helps to improve the international competitiveness of the country, thereby improving resource allocation within the economy.

No burden on government budget

Market-based supply-side policies emphasize the importance of allowing markets to work more freely and efficiently. Hence, there is less importance placed on role of the government in the economy, especially in labour markets. As there is less government involvement in economic activity, this should improve resource allocation without any substantial costs to the government. For example, there are no significant costs of deregulation, privatization, policies to increase competition in markets, or trade liberalization policies. Hence, the economy can benefit from growth and employment in the long run without major burdens on the government budget or national debt. Note that this advantage does not apply to interventionist supply-side policies.

Strengths of interventionist supply-side policies

Direct support of sectors important for growth

Interventionist supply-side policies can be used to target specific sectors of the economy to facilitate their growth. For example, industrial policies such as tax credits (tax rebates) and/or subsidies help to improve the competitive nature of certain industries deemed to be strategically important for the economy in terms of economic growth and employment. Since supply-side policies can improve resource productivity and national output without pressures of the general price level, the international competitiveness of the country should improve, thus helping to increase export sales.

In addition, interventionist supply-side policies may be necessary to support sectors that provide merit goods and public goods (see Chapters 11 and 12) otherwise they could be underprovided by the market system. This includes, for example, investments for the provision

Top tip!

Although interventionist supply-side policies take time to implement and to reap tangible benefits for the economy, make sure you can explain *how* and *why* these policies can help the government to achieve its macroeconomic aims: economic growth, full employment (low unemployment) and stable prices (low inflation).

ATL ACTIVITY 26.4

In groups of 2–4 people, discuss the effectiveness of supply-side policies in achieving the main macroeconomic objectives. Ideally, if there are 4 people in the group, each student can be assigned one of the key macroeconomic objectives to discuss. Try to use real-world examples and/or real-world issues in your discussions.

of research and development, infrastructure and human capital (education and training). In the long run, the economy benefits from the potential welfare gains in terms of growth, low unemployment, and a low and stable rate of inflation.

In summary, successful supply-side policies have the effect of increasing the economy's productive capacity, shown by a shift in the economy's long run aggregate supply (LRAS) curve to the right (see Figure 26.2). Most economists believe that improved supply-side performance of the economy is vital to achieving sustained economic growth and low unemployment, without causing a rise in the general price level (inflation).

WISE ChoICES

To what extent does government **intervention** in the economy, by using demand-side and supply-side policies, enable individuals and households to make more informed **choices**?

PAPER 1 EXAM PRACTICE QUESTION 26.3

Using real-world examples, evaluate the extent to which interventionist supply-side policies can increase economic growth, attain low unemployment, and achieve a sustainable low rate of inflation. *[15 marks]*

PAPER 1 EXAM PRACTICE QUESTION 26.4

Using real-world examples, discuss whether market-based supply-side policies are effective in achieving a country's macroeconomic objectives. *[15 marks]*

Top tip!

When evaluating the effectiveness of supply-side policies, remember to consider why these are being used in the first place – in terms of the strengths and limitations of such policies in promoting economic growth, achieving low unemployment, and attaining a low and stable rate of inflation in the economy.

Theory of Knowledge

When evaluating the effectiveness of supply-side policies, how important are ethics and cultural differences?

Chapter summary

- Supply-side policies are long-term policy measures of the government used to increase the productive capacity (potential output) of the economy.
- Diagrammatically, the potential output of an economy is shown on its production possibility curve (PPC) or the full employment level of national output on the LRAS curve.
- The goals of supply-side policies are long-term growth (by increasing the economy's productive capacity), improving competition and efficiency, reducing labour costs and unemployment (through labour market flexibility), reducing inflation (to improve international competitiveness), and increasing firms' incentives to invest in innovation (by reducing costs).
- Market-based supply-side policies include policies to encourage competition, labour market policies (labour market reforms) and incentive-related policies.
- Interventionist policies include improving education, training, healthcare, research and development, provision of infrastructure and industrial policies.
- Both interventionist and market-based supply-side policies can be used to increase the productive capacity or long run aggregate supply (LRAS) of the economy, leading to a rise in its potential output.
- In the long run, supply-side policies help to shift the economy's LRAS curve to the right and improve its international competitiveness, thereby attracting more foreign direct investment. This can help the economy to achieve sustainable growth.
- In the short run, interventionist supply-side policies (such as investments in human capital through education and training, or investments in R&D) increase aggregate demand but in the long run will increase the economy's aggregate supply.
- Improved supply-side performance of the economy can result in sustained economic growth without causing a rise in inflation.
- Nevertheless, supply-side policies on their own are insufficient to achieve long-term economic growth. There must also be a high-enough level of aggregate demand so that the productive capacity of the economy materializes.

The global economy

27 Benefits of international trade

Real-world issue

Who are the winners and losers of the integration of the world's economies?

Conceptual understandings (WISE ChoICES)

- The increased **interdependence** of economies has benefits and costs.
- Increased economic integration may result in **efficiency**, welfare gains and improvements in **economic well-being** but the benefits may not result in **equity**.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- the benefits of international trade (AO2)
- a free trade diagram to illustrate exports when the world price is above the domestic price (AO4)
- a free trade diagram to illustrate imports when the world price is below the domestic price (AO4)
- a diagram, the quantity of exports, quantity of imports, import expenditure and export revenue (AO4) (HL only)
- absolute and comparative advantage, with reference to gains from trade, sources of comparative advantage and opportunity costs (AO2) (HL only)
- a linear PPC diagram to show differing opportunity costs and the potential gains from specialization and trade as a result of comparative advantage (AO4) (HL only)
- opportunity costs from a set of data in order to identify comparative advantage (AO4) (HL only)
- the limitations of the theory of comparative advantage (AO3) (HL only).

Benefits of international trade (AO2, AO4)

Economists in general believe that international trade is the single largest reason why various regional economic powers have prospered in the past. Historical records trace the existence of international trade back to 18BCE through the remains of Assyrian merchant colonies in eastern Turkey. **International trade** is the exchange of goods and services across national states. It involves the sale of **exports** (goods and services sold to overseas buyers) and **imports** (foreign goods and services bought by domestic households and firms).

Countries are compelled to trade internationally since scarce resources are unevenly distributed between countries, that is, different countries have different **factor endowments**. For example, Saudi Arabia and Kuwait are well endowed with oil supply, whereas Thailand and Vietnam are naturally well endowed with arable land and the climate for growing crops such as rice. Another example is South Africa, which produces almost two-thirds of the world's platinum.

◆ **International trade** is the transnational exchange of goods and services. It involves the sale of *exports* (goods and services sold to overseas buyers) and *imports* (foreign goods and services bought by domestic households and firms).

◆ **Factor endowment** refers to the quantity and quality of factors of production available in a country, such as natural resources and human capital. Countries well endowed with natural resources tend to have a cost advantage in the production of products using these resources.

Figure 27.1 Landscaped terraced rice fields in Mu Cang Chai, Vietnam

Most economists believe that the benefits of free international trade exceed the opportunity costs. Although not everyone gains, or gains equally, from international trade, the net benefits to society are positive. These positive impacts include:

- increased competition
- lower prices
- greater choice
- acquisition of resources
- more foreign exchange earnings
- access to larger markets
- economies of scale
- more efficient resource allocation
- more efficient production.

These interrelated benefits of international trade are outlined below.

Increased competition

International trade promotes competition among domestic and international firms. Domestic businesses encounter greater competition especially if foreign firms offer a greater variety of goods and services at cheaper prices, as well as products of better quality. Either way, as local producers and firms are exposed to increased competition from overseas, they are forced to be more efficient and innovative. This may be in the form of offering more competitive prices and/or adding superior features to their goods and services. Consumers benefit in such scenarios as more varieties of goods and services are offered, and often at a lower cost.

Lower prices

International trade makes it possible to reduce the price of goods and services. Because of intense competition among domestic and international businesses, firms strive for more efficient means of production. They can also benefit from economies of scale (see Chapter 14) by operating in larger markets, thus reducing their average costs of production. This can therefore result in a lower unit price of goods and services and thus consumers are able to access internationally traded goods

and services at a lower price. Domestic producers can also take advantage of international trade by buying cheaper raw materials and semi-finished goods from foreign suppliers. This also helps to lower prices owing to lower production costs.

Greater choice

International trade enables consumers and firms to access a larger variety of goods and services from different producers and suppliers from around the world. Consumers not only have access to domestically produced products, but also to the goods and services produced globally. For example, consumers in China can choose to buy blueberries, grapes and pears produced in Chile (one of the largest fruit exporters in the world), rather than being limited to locally produced fruits only. Another example is that consumers at newly opened IKEA stores in India can access lamps and other home furniture products designed in Sweden and made in Turkey.

Acquisition of resources

Owing to variances in factor endowments between countries, a nation may be rich in natural resources for production in primary industries but may have less suitable resources to support its manufacturing or services sectors. However, through international trade, countries can acquire resources such as natural resources and capital goods that are not available domestically to further their production. Hence, the acquisition of resources through international trade makes it possible for countries to grow and develop. For example, China's Lenovo is one of the largest producers of laptops. The company can access cobalt, a necessary raw material used to produce electronic goods, from the Democratic Republic of the Congo because China is not a cobalt-rich country. Singapore, a major service hub in South East Asia, acquires over 99 per cent of its primary resources from neighbouring countries.

Figure 27.2 Cobalt is a vital material for producing laptops

Foreign exchange earnings

When firms sell goods and services to overseas customers, the country earns export earnings in the form of foreign exchange (foreign currencies). This enables the exporting country to make payments to other countries for the purchase of foreign goods and services (import expenditure). For example, Iran sells crude oil, one of its major export items, which enables it to secure foreign currencies to be able to import products such as machinery, cereals, iron, steel and chemicals from different countries. By contrast, during the 2010s as a result of trade restrictions imposed by most major economies, Iran struggled to amass enough foreign currencies to pay for the country's demand for imported goods and services.

CASE STUDY

Hilti Corporation

Hilti is a multinational company from Liechtenstein, which is bordered by Switzerland and Austria and is Europe's fourth-smallest country, with an area of just over 160 square kilometres and a population of 38,749. Hilti develops, manufactures and markets products for the construction, building maintenance, energy and manufacturing industries. In 2019, Hilti had offices in 120 different countries, 29,000 employees and sales of over 5 billion Swiss francs ($5.28bn). Liechtenstein gains a substantial amount of foreign exchange earning from the export sales and profits of companies such as Hilti.

Common mistake

You may confuse the benefits of international trade with the benefits of free trade. However, access to larger markets through international trade is not necessarily the same as free trade. Gaining access to larger markets may come with restrictive trade policies (such as tariffs and quotas – covered in Chapter 28) whereas there are no such restrictions under free trade (see Chapter 29).

Access to larger markets

Having greater access to larger markets around the world enables countries to reach a larger customer base. By increasing the level of production of goods and services, firms are able to achieve greater economies of scale, which leads to higher levels of profit and the creation of jobs in the long-term. Hence, greater access to larger markets also helps to tackle the problems of poverty (see Chapter 23).

The growing integration of the world's economies through membership of trading blocs (see Chapter 30) offers countries the potential to access even more markets, enabling countries to achieve sustainable rates of economic growth, job creation and a reduction in extreme poverty (see Chapter 30).

Figure 27.3 Soft drinks are massed produced, enabling manufacturers to enjoy huge economies of scale

ATL ACTIVITY 27.1

Watch this video which shows poor cocoa producers in the Ivory Coast enjoy the taste of chocolate for the first time in their life. The Ivory Coast is the the largest producer of cocoa in the world. **https://youtu.be/zEN4hcZutO0**

Discuss the extent to which it is socially acceptable to deprive producers of primary resources of the final product the rest of the world enjoys.

Economies of scale

An increase in output enables firms to achieve lower average costs of production. By operating on a larger scale in international markets, international trade enables more firms to benefit from economies of scale, so that cost savings can be passed on to consumers in the form of lower prices. International trade provides domestic businesses the opportunity to increase output as firms supply a greater volume of goods and services to countries outside of the home country (in addition to supplying these to local customers). The larger scale of operations also helps domestic businesses to take advantage of division of labour and specialization and to invest in specialized capital and machinery, leading to further cuts in average costs of production.

Theory of Knowledge

Do trade liberalization policies always lead to international trade being fairer? How do economists know who actually gains or loses from trade being freer?

Efficient resource allocation

International trade benefits the economy as it encourages an efficient allocation of the world's scarce resources. Many countries eliminate trade barriers to make sure that international trade takes place within a less restrictive environment, which brings the aforementioned benefits. Usually neighbouring countries, or countries with strategic

economic relationships, move towards trade liberalization and reduce or remove trade protectionist measures for more efficient allocation of resources (see Chapter 30). Freer and fairer international trade forces domestic firms to focus on improving the quality of their output due to foreign competition, again improving resource allocation in the economy.

Efficient production

Similarly, international trade enables more efficient production of goods and services to take place. International trade incentivizes domestic and foreign firms to compete on price and non-price factors (such as quality). International trade helps domestic consumers to access more goods and services and at lower prices. Inefficient and unproductive firms become uncompetitive, especially when exposed to foreign competition. In order to stay relevant and marketable, domestic firms are forced to use resources more efficiently in the production process, by using minimal inputs to maximize output. Domestic firms become more responsive to the needs of the market, so need to be efficient and innovative. This may include cutting unnecessary costs and allocating resources more effectively to improve productive efficiency. For example, international trade helps domestic firms to source cheaper and/or better-quality materials and components for production.

ATL ACTIVITY 27.2

With a partner, identify up to **five** goods and services that you use but are not produced in your home country. Discuss why you have not chosen to buy the same products produced domestically but preferred to purchase imported rival products instead. Be prepared to share your responses with the rest of the class.

Theory of Knowledge

To what extent should policymakers limit their country's involvement in international trade in order to protect the language, traditions and culture of the country? What criteria should policymakers use to make such decisions?

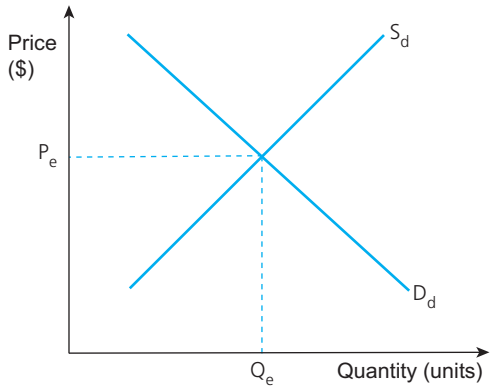


Figure 27.4 Domestic production before international trade

Free trade illustrating exports when world price is above domestic price (AO4)

Figure 27.4 illustrates the demand and supply of a given product in the domestic market. From the intersection of the domestic demand for the product (D_d) and the domestic supply of the product (S_d), an equilibrium is determined. The domestic price of the commodity is P_e and the quantity traded is Q_e.

With international trade, the world price of the good or service determines if the domestic country will import the product from other countries or export it to its trading partners. If

the export price of a product is higher than the world price (thus the import price) as other countries may have a natural advantage or specialize in the production of this product, it makes financial sense for the domestic country to import the good or service. By contrast, if export prices are lower than world prices, then the country will sell more of its products to overseas customers, ceteris paribus.

Figure 27.5 illustrates the situation when the world supply or world price (P_w) is higher than the domestic price (P_d). The higher world price causes demand to contract from Q_e to Q_d, and for supply to expand from Q_e to Q_s, creating a situation of excess supply (see Chapter 5). Hence, the country is able to export this excess, shown by the distance $Q_s - Q_d$ at the price of P_w to its trading partner countries. This scenario occurs when the domestic country has a comparative cost advantage in the production of this particular good or service.

In Figure 27.6, when the world supply curve or world price of a product (shown by P_w) is lower than the domestic price of the product (P_d), the quantity demanded will expand to Q_d but the domestic quantity supplied contracts to Q_s. To meet the excess domestic demand for the product, the country will have to import $Q_s - Q_d$ amounts of the good. This scenario occurs when overseas economies have a comparative cost advantage in the production of this particular good or service, so it makes economic sense for the domestic country to import this product.

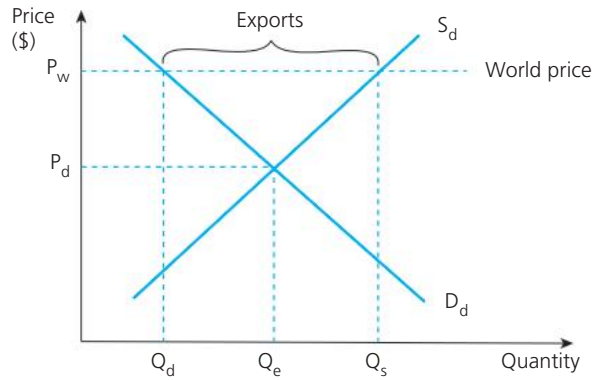


Figure 27.5 Impact on exports when the world price is above the domestic price

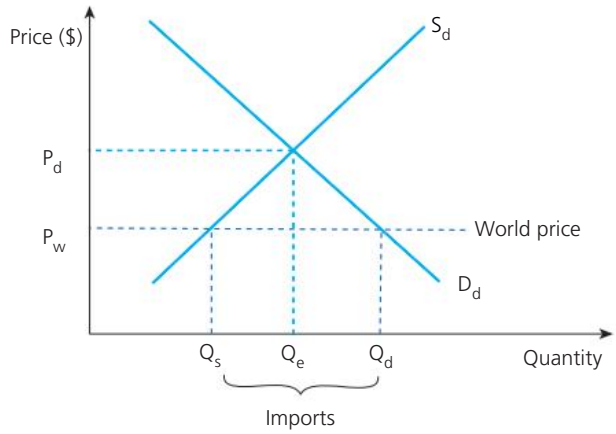


Figure 27.6 Impact on imports when the world price is below the domestic price

Top tip!

Make sure you can distinguish between the merits of *international trade* and the merits of *free trade* – they are not quite the same because not all international trade entails free trade. Thus, the merits of free trade are greater than the benefits of international trade (which may involve some trade barriers).

Common mistake

It is incorrect to assume that international trade improves the welfare of *all* members of society, as this is not necessarily the case. For example, some exporters might lose out owing to the fierce competition from foreign rivals.

PAPER 2 EXAM PRACTICE QUESTION 27.1

Bangladesh is one of the world's largest producers of rice and tropical fruits. In Brunei Darussalam, crude oil and natural gas account for around 90 per cent of the country's gross domestic product. This makes Brunei Darussalam one of the leading producers of oil in South East Asia.

a Explain **one** reasons why countries such as Bangladesh and Brunei Darussalam trade with each other. *[2 marks]*

b Outline **one** problem for Brunei Darussalam relying on oil exports. *[2 marks]*

c Explain how Bangladesh's export of rice and tropical fruits helps its farmers to achieve economies of scale. *[4 marks]*

PAPER 3 EXAM PRACTICE QUESTION 27.2 (HL ONLY)

Country Z is a developing country and one of the world's largest exporters of plum oil. The diagram below (Figure 27.7) illustrates the market for plum oil in Country Z. D_d and S_d represent domestic demand and supply per year, in thousands of litres, while P_w is the world price at $7 per litre.

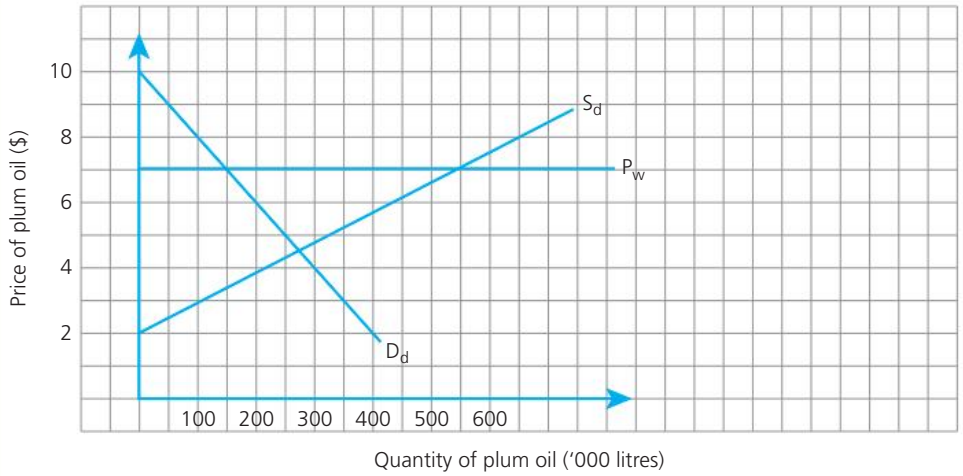


Figure 27.7 The market for plum oil

a Calculate the value of plum oil exports per year from Country Z. *[2 marks]*

b Calculate the social (community) surplus earned by stakeholders in the plum oil market in Country Z under conditions of free trade. *[3 marks]*

c Calculate the total value of domestic consumption of plum oil in Country Z under free trade. *[2 marks]*

Absolute and comparative advantage (AO2, AO4) (HL only)

The benefits of international trade outlined earlier in this chapter help to explain how international trade accounts for the prosperity of wealthy nations throughout history. The concept of absolute advantage was coined by Scottish economist Adam Smith (1723–90), extending his doctrine of specialization and the division of labour. Smith argued in his monumental book *An Inquiry into the Nature and Causes of the Wealth of Nations* (1776) that absolute advantage and free trade are the underlying reasons behind the wealth that the nations are able to accrue.

A country has an **absolute advantage** when it can produce more of a good or service than another country using the same amount of resources (or being able to produce the same amount of a good or service using fewer resources). If countries specialize in the output of what they are most efficient at producing, world output will subsequently increase – in the same way that if workers in a firm specialize then labour productivity will improve. Smith stressed the importance of international specialization to explain the theory of absolute advantage. An absolute advantage occurs mainly because different countries have different factor endowments. For example, Qatar is abundant in natural gas, while Madagascar has the natural climate suitable for producing vanilla. By contrast, some countries have an absolute disadvantage, so are inefficient and unproductive in the output of given goods or services relative to their trading

◆ **Absolute advantage** occurs when a country can produce more of a good or service than another country using the same amount of resources.

partners. For example, Scandinavian countries do not have the natural climate suitable for growing tropical fruits, so they import these products from countries with an absolute advantage in the production of these goods, such as Brazil, Costa Rica, Ecuador, India, the Philippines and Spain.

An absolute advantage in the production of a good or service gives a country a competitive edge because it is more productive and efficient. For example, if Country A spends $10m to make 40,000 units of output of a particular good while Country B can produce only 30,000 units with the same amount of money and labour input, then Country A has an absolute advantage in the output of that product. International specialization in the output of goods or services based on absolute advantage results in an increase in world production and consumption.

Figure 27.8 Denmark, Norway and Sweden import tropical fruits

The theory of absolute advantage is based on the following assumptions:

- Costs of production are based on the number of labour hours in the production process.
- For simplification, only two nations and two commodities are used.
- There is internal labour mobility, but domestic labourers are not expected to work outside of the country.
- Trade between two countries takes place only when the respective countries have a lower cost in the production of any one of the two commodities to be traded.

Absolute advantage: gains from trade (AO2, AO4)

Suppose Brazil and Australia, with the same amount of resources, produce timber and cattle. Also assume that both countries divide their resources equally between the production of timber and cattle. In Table 27.1, the first two columns show the amount of timber OR cattle that the countries can produce, based on the model of the production possibility curve (see Chapter 1). Australia can produce either 6 units of timber or 12 units of cattle. Brazil can produce either 10 units of timber or 5 units of cattle. Hence, it is clear that Australia has an absolute advantage in producing cattle and Brazil has the absolute advantage in producing timber.

Table 27.1 Absolute advantage

Country	Timber (units)		Cattle (units)
Australia	6		12
Brazil	10	OR	5
Total output	16		17

With this information, two production possibility curves (PPC) can be drawn, showing the different combinations of the two goods that both Brazil and Australia can produce (Figure 27.9). Recall that the production possibility curve (or production possibility frontier) model shows the maximum combinations of any two products that can be produced using the same factor endowments and technology, shown by any point on the country's PPC.

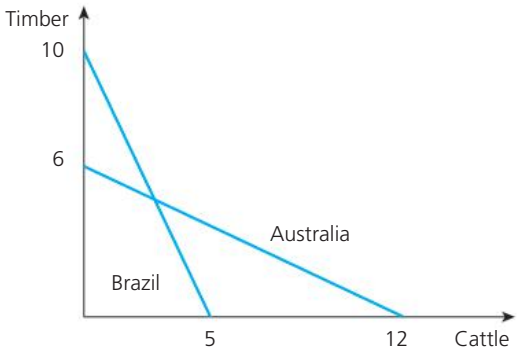


Figure 27.9 Absolute advantage using the PPC model

Brazil's production possibility curve shows that it can produce a maximum of 10 units of timber or 5 units of cattle. Australia's PPC indicates that

it can produce only 6 units of timber but 12 units of cattle. Thus, Brazil has an absolute advantage in the output of timber and Australia has this advantage in the output of cattle.

Using a linear PPC model (for ease of illustration), it can be seen in Figure 27.10 that Australia can produce 6 units of timber and no cattle, or no output of timber but 12 units of cattle, or any combination in between. This is based on the opportunity cost ratio of 1 unit of timber for 2 units of cattle. Similarly, Brazil could produce 10 units of timber and no units of cattle, or it could produce 5 units of cattle and no output of timber, or any combination of these two goods based on its opportunity cost of 2 units of timber for 1 unit of cattle.

If both the countries decide to be self-reliant and choose not to trade with each other, technically known as **Autarky**, they could produce the following combination of timber and cattle shown in Table 27.2. Without trade, the total production of timber and cattle is limited to 7 and 9 units respectively.

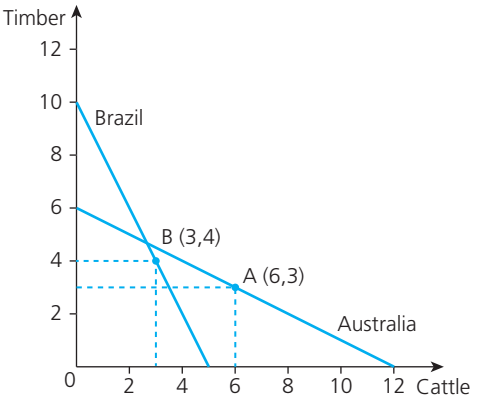


Figure 27.10 Absolute advantage without trade

Table 27.2 Absolute advantage without trade

Country	Timber (units)	Cattle (units)
Australia	3	6
Brazil	4	3
Total output	7	9

ATL ACTIVITY 27.3

In the past many countries have adopted such no trade (or limited trade) policies. For example, Japan adopted 'Sakoku' (meaning closed country) during 1639 to 1853 where any travel outside the country, international relations and international trade were strictly limited if not forbidden. From 1814 to 1840, after gaining independence, the president of Paraguay adopted the policy of isolation by closing all national borders to travel and trade with other countries.

With a partner, investigate the different reasons why some countries chose not to trade with other nations in the recent past. Share your main findings with the rest of the class.

◆ **Autarky** refers to self-sufficiency, that is a country that does not engage in international trade.

It is clear from Figure 27.9 that Brazil should specialize in the output of timber and that Australia should concentrate on the production of cattle. If both countries specialize and produce the good for which they have an absolute advantage, Australia would produce only 12 units of cattle and 0 units of timber, and Brazil would produce 10 units of timber with no cattle (see Table 27.3). With this arrangement, the total output of timber goes up by 3 units (previously 7 units) and the output of cattle increases by 3 units (previously 9 units).

Table 27.3 Absolute advantage with trade

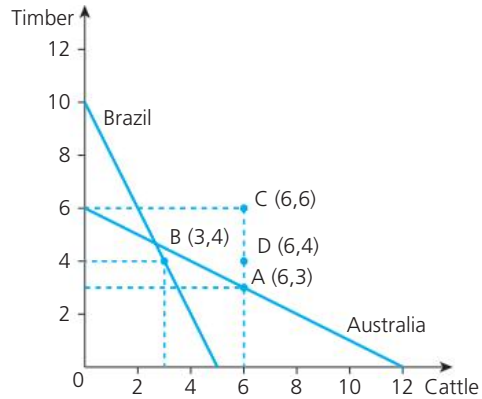
Country	Timber (units)	Cattle (units)
Australia	0	12
Brazil	10	0
Total output	10	12

As it is unrealistic for these nations to specialize and produce just one good, by engaging in trade, both countries can benefit. In this example, assuming 1 unit of timber is exchanged for 1 unit of cattle, Australia could trade 6 units of cattle with Brazil for 6 units of timber. This would leave Brazil with 4 units of timber and Australia with 6 units of cattle. Table 27.4 shows the production of timber and cattle for both Australia and Brazil before trade (see Table 27.2) and after trade. This can also be represented by using a PPC diagram, as shown in Figure 27.11.

Table 27.4 Absolute advantage with and without trade

	Without trade		With trade	
Country	Timber (units)	Cattle (units)	Timber (units)	Cattle (units)
Australia	3	6	6	6
Brazil	4	3	4	6
Total output	7	9	10	12

Figure 27.11 shows how both countries benefit after trading the goods in which they have an absolute advantage. Before trade, Australia was on point A, indicating a production of 3 units of timber and 6 units of cattle. With trade, Australia can now operate at point C, with 6 units of cattle but 6 units of timber, so the country enjoys more goods overall after trade. Note that point C is outside of Australia's PPC. Similarly, Brazil can enjoy 4 units of timber and 6 units of cattle (point D), enabling the country to also operate beyond its PPC. Total output of both countries (world output in this model) also moves up to 10 units of timber and 12 units of cattle, as shown in Table 27.4.



■ **Figure 27.11** Benefits of trade based on absolute advantage

Comparative advantage: gains from trade and opportunity cost (AO2, AO4)

The theory of comparative advantage was coined by British classical economist David Ricardo (1772–1823) who suggested that trade should not necessarily be based on what a country has an absolute advantage in the production of, but should be based on the comparative advantage it might have for the production of certain goods and services. He argued that countries should specialize in the output of goods and services based on what they are able to produce at a relatively lower opportunity cost (or relatively lower unit costs of production). A country has a **comparative advantage** if it is able to produce a given amount of a good or service at a lower opportunity cost than another country. This means the country gives up fewer resources than other nations in order to produce the good or service. The theory of comparative advantage suggests that countries would benefit from producing and trading products in which they have a comparatively lower opportunity cost, even if one of the trading partners has an absolute advantage in the output of both products.

◆ **Comparative advantage** exists when a country can produce a given amount of output at a lower opportunity cost than another country, that is, the country gives up fewer resources than other nations in order to produce the good or service.

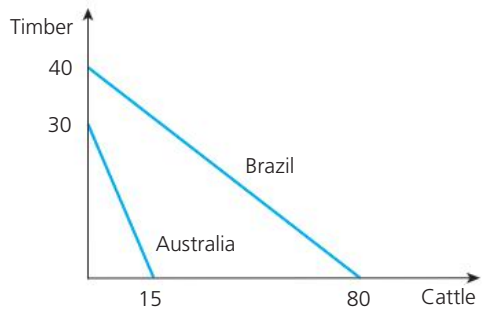
Continuing with our example used to illustrate the theory of absolute advantage, suppose Brazil and Australia divide their resources equally between the production of timber and cattle. Table 27.5 shows that Australia can produce either 30 units of timber or 15 units of cattle, while Brazil can produce either 40 units of timber or 80 units of cattle. Hence, Brazil has an absolute advantage in producing both cattle and timber because it can produce more timber and cattle compared to that of Australia using the same unit of resources.

■ **Table 27.5** Comparative advantage before trade

Country	Timber (units)	Cattle (units)
Australia	30	15
Brazil	40	80

This information can also be shown by constructing two production possibility curves, showing the different combinations of two goods (see Figure 27.12).

Figure 27.12 shows that Brazil has an absolute advantage in the output of both cattle and timber, because it can produce more of both goods compared to Australia. However, Ricardo pointed out that comparative advantages exist if there are different opportunity costs of producing these goods. Brazil's opportunity cost of producing 1 unit of timber is 2 units of cattle (40 units of timber for 80 units of cattle), whereas Australia's opportunity cost is only 0.5 units of cattle (30 units of timber for only 15 units of cattle). In this case, it would make more economic sense for Australia to give up the production of cattle and to specialize in producing timber. The opposite is true when considering which country should specialize in the output



■ **Figure 27.12** Comparative advantage before trade

of cattle. Australia's opportunity cost of producing 1 unit of cattle is 2 units of timber, whereas this is only 0.5 for Brazil. In other words, Brazil can gain 2 units of cattle for each unit of timber it gives up (Table 27.6).

Table 27.6 Comparative advantage and opportunity costs

Country	Output (PPC)		Opportunity costs		Lowest opportunity cost
	Timber (units)	Cattle (units)	Timber	Cattle	
Australia	30	15	0.5	2	Timber
Brazil	40	80	2	0.5	Cattle

Therefore, according to the theory of comparative advantage, Brazil should specialize in and export cattle, while Australia should specialize in and export timber. This information is summarized in Table 27.7 showing that Australia and Brazil produce 30 units of timber and 80 units of cattle respectively. Suppose that with trade, Australia exports 25 units of timber for 25 units of cattle from Brazil (an exchange rate of 1:1). Hence, Australian consumers enjoy 5 units of timber and 25 units of cattle. Brazil exports 25 units of cattle to Australia and buys 25 units of timber with its export earnings. Hence, Brazilian consumers can enjoy 55 units of cattle and 25 units of timber, after trade. This situation is also illustrated in Figure 27.13, showing that both countries are able to operate beyond their PPC before trade.

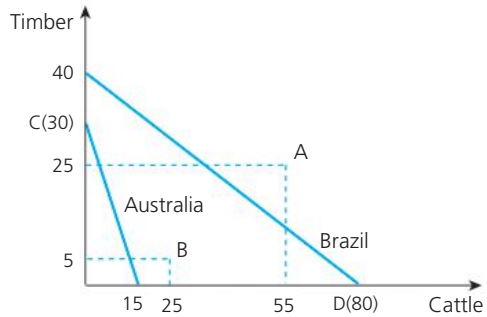


Figure 27.13 Benefits of comparative advantage with trade

Table 27.7 Comparative advantage without and with trade

Country	Without trade		With trade		Total output (units)
	Timber (units)	Cattle (units)	Timber (units)	Cattle (units)	
Australia	30	0	30 – 25 = 5	25	30
Brazil	0	80	25	80 – 25 = 55	80

Figure 27.13 illustrates how trade based on comparative advantage brings benefits to both countries, as domestic consumers in both nations get to enjoy more timber and cattle.

After specialization and trade, Australia can consume at point B (5 units of timber and 25 units of cattle), which is beyond its original production possibility frontier prior to trade. Similarly, Brazil can consume at point A (25 units of timber and 55 units of cattle), which is again outside Brazil's original PPF. Point D indicates maximum production of cattle for Brazil (with zero production of timber) and point C indicates maximum production of timber for Australia (with production of cattle).

PAPER 3 EXAM PRACTICE QUESTION 27.3 (HL ONLY)

Refer to the following production possibilities for two countries and answer the questions that follow.

	Fruits (units)	Vegetables (units)
Country K	8,000	0
	0	10,000
Country P	4,000	0
	0	8,000

a Identify which country has the absolute advantage in the production of vegetables. *[1 mark]*

b Outline which country should specialize in the production of fruits. *[2 marks]*

c Calculate the opportunity cost of producing 8 units of fruits, in terms of vegetables, for Country P. *[2 marks]*

Common mistake

Don't forget to acknowledge that comparative advantage is not static in nature. You should be able to demonstrate knowledge of changes in comparative advantage over time. Without remaining competitive, countries can lose any comparative advantage they have very quickly. Some of the reasons for this are highlighted in the next section.

Sources of comparative advantage (AO2)

Some countries can produce goods and services at lower opportunity costs compared to other countries for any of a combination of the following reasons:

Figure 27.14 Saudi Arabia possesses around 18 per cent of the world's petroleum reserves

- Factor endowments and the quality of factors of production – For example, major oil exporting countries including Saudi Arabia, Kuwait, Russia and the United Arab Emirates are well endowed in oil supply. These countries can increase production through specialization and international trade. Other countries do not have access to oil so need to specialize in producing other goods and services using factor resources that they are better endowed with.
- Levels of technology – For example, workers with access to the latest machinery and technologies will be far more productive than those using outdated techniques and substandard equipment. High-income countries have the financial capability to invest in physical capital and superior technologies, whereas this is not possible in low-income countries that resort to producing basic commodities that are exported for a relatively low price.
- Investment in research and development (R&D) – The ability to invest in R&D and to be innovative can give countries a competitive advantage. R&D expenditure can also generate new work procedures and processes that reduce the relative costs of production. Countries like Israel, South Korea and Finland spend almost 3 to 4 per cent of their respective gross national product (GNP) on R&D, whereas countries like Algeria, Indonesia and Macao spend less than 0.1 per cent on R&D. As comparative advantage is not static, supply-side economists (see Chapter 26) argue it is necessary for countries to invest in R&D in order to increase the economy's productive capacity.
- Price stability – The inability to stabilize prices (that is, unstable rates of inflation in an economy) can damage the comparative cost advantages of a country. This is because higher prices mean that foreign buyers are less willing and able to purchase the exports of domestic firms, even if these may be of higher quality (see Chapter 20). Venezuela, Zimbabwe and Sudan have extremely high rates of inflation, damaging or destroying any of the cost advantages of these countries.
- Exchange rate fluctuations – A change in the exchange rate (see Chapter 31) affects the relative prices of exports and imports. A long-term unfavourable change in the exchange rate (caused by a currency revaluation or an appreciation in the value of the currency can trigger a reduction in the demand from overseas customers due to prices being relatively higher.

WORKED EXAMPLE

Suppose two countries, with the same amount of resources, produce fruits and toys. Also assume that both countries divide their resources equally between the production of fruits and toys. The pre-trade situation is shown below:

Country	Fruits (units)	Toys (units)
Delta	3,000	1,500
Gamma	500	1,000
Total output	3,500	2,500

Delta has an absolute advantage in the production of both toys and fruits, that is, it is better at producing these goods than country Gamma. However, Gamma can produce toys at a lower opportunity cost:

Country	Fruits (units)	Toys (units)	Opportunity cost
Delta	3,000	1,500	2.0:1
Gamma	500	1,000	0.5:1

While Delta has to give up 2 units of fruits to gain 1 unit of toys, Gamma has to give up only 0.5 units of fruits to produce 1 unit of toys. Hence, Gamma should specialize in the output of toys. Alternatively, it can be seen that if Delta gives up 1 unit of toys, it can produce 2 units of fruits whereas Gamma can gain only 0.5 units of fruits by sacrificing 1 unit of toys. Hence, Delta should specialize in the output of fruits.

Suppose Gamma specializes entirely in toys and Delta decides to switch 500 units of toys to produce (1,000 extra units of) fruits instead. Hence, the total output increases by 500 units of both fruits and toys:

Country	Fruits (units)	Toys (units)
Delta	4,000	1,000
Gamma	0	2,000
Total output	4,000	3,000

If the countries trade their surpluses, say 700 units of each product, both countries will have more of fruits and toys compared to what they started with, when there was no international trade:

	Output with trade		Output without trade	
Country	Fruits (units)	Toys (units)	Fruits (units)	Toys (units)
Delta	3,300	1,700	3,000	1,500
Gamma	700	1,300	500	1,000
Total output	4,000	3,000	3,500	2,500

Top tip!

The causes or sources of comparative advantage can be easily remembered by using the acronym REFIT: **R**esearch and development, **E**xchange rates, **F**actor endowments, **I**nflation (price stability) and **T**echnology.

WORKED EXAMPLE

Use the production possibilities data below and an appropriate diagram to explain which country should specialize in the production of Aris and which country should specialize in the output of Walkers. *[4 marks]*

Country	Aris (units)	Walkers (units)
Lucland	1,500	600
Mayland	2,000	500

Solution

Mayland should specialize in (and export) Aris as it gives up 1 unit of Walkers to gain 4 units of Aris (whereas Lucland gains only 2.5 units of Aris if it gives up the same unit of Walkers).

Country	Aris (units)	Walkers (units)	Ratio
Lucland	1,500	600	2.5:1
Mayland	2,000	500	4.0:1

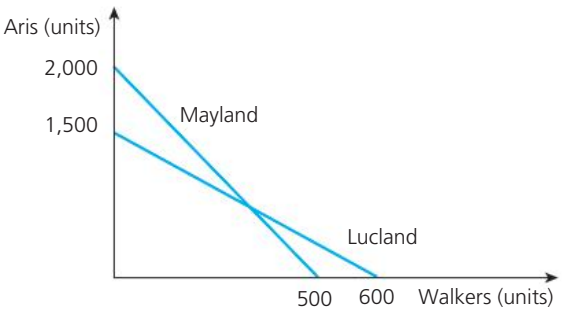


Figure 27.15 The PPC and comparative advantage

- Lucland has the absolute advantage in the output of Walkers.
- Mayland has the absolute advantage in the production of Aris.
- As can be seen in the diagram (Figure 27.15), the opportunity cost of Mayland giving up production of Aris is greater than the opportunity cost of Lucland giving up production of Walkers.

ATL ACTIVITY 27.4

Investigate the relationship between investment in research and development (R&D) and economic growth for a country of your choice. Discuss the extent to which R&D facilitates economic growth in your chosen country.

Limitations of the theory of comparative advantage (AO3) (HL only)

David Ricardo's theory of comparative advantage states that a country should specialize and export the goods and services that it is able to produce at lower opportunity costs than its trading partners. Comparative advantage theory is considered to be one of the key bases that have helped

many countries to develop their trade policies. The theory is regarded as an improvement to the theory of absolute advantage proposed by Adam Smith. However, the theory of comparative advantage has some limitations, including those listed below:

- The theory assumes that comparative advantage is fixed, whereas in reality it is not. For example, Hong Kong lost its comparative advantage in textiles and toy manufacturing many decades ago, while the UK no longer has a relative cost advantage in the output of steel and coal. Economies, with the evolutionary nature of technology, are constantly changing, and so do comparative advantages. Hence, it is a static model, which ignores the fact that comparative advantages can and do shift between countries over time.

CASE STUDY

British Steel

British Steel faced huge challenges during the 2010s due to declining demand for steel in Europe, intense competition from India and China, and the company's high level of debts. In late 2019, a deal was agreed with China's Jingye to buy out British Steel for around £50m ($65m), and effectively saving around 4,000 jobs. Jingye announced it would invest £1.2bn ($1.56bn) over the next decade to upgrade plants and machinery in Scunthorpe, North England. British Steel is the largest employer in Scunthorpe.

Figure 27.16 Comparative advantages can and do change over time

- Comparative theory assumes that there are no barriers to international trade, although this is not necessarily the case in the real world. As explained in Chapter 30, trade barriers such as tariffs and quotas are still being imposed by countries around the world, despite the theoretical benefits of freer and fairer international trade.
- While the theory explains how countries can gain from international trade, it ignores the disparities of the relative gains from trade. Some countries will gain more than others owing to their better bargaining position and terms of trade. Comparative advantage, therefore, also ignores the impact of trade on relative poverty of low-income countries.
- Transportation costs related to international trade are ignored in the calculation of comparative advantage. Hence, even if China is able to produce textbooks relatively cheaper than the UK, for example, transportation costs (and the associated time delays) to Europe could mean China has a comparative disadvantage.
- It is also assumed there is perfect occupational mobility of resources. Factors of production can be switched between different industries without any loss of efficiency, as shown in the PPC model (see Chapter 1). This contradicts the idea of increased gains from trade through specialization and rising opportunity costs of greater levels of production.
- Comparative advantage also assumes perfect knowledge of pricing information available to both consumers and producers. Owing to complications such as exchange rate fluctuations and relative inflation rates in different countries, it can be difficult for consumers and producers to have perfect knowledge of prices of exports and imports.

ATL ACTIVITY 27.5

Other than your country of origin, what is your favourite country? Carry out research on the goods and services this country spends most on importing.

Figure 27.17 The theory of comparative advantage ignores the impact of transportation costs

WISE ChoICES

To what extent does **interdependence** between countries always promote greater **economic well-being**?

ATL ACTIVITY 27.6

With regards to the importance of international trade, which IB learner profile do you think is the most relevant and important? Why do you believe this to be the case?

Chapter summary

- The benefits of free international trade tend to exceed the costs. Countries and societies benefit from international trade in terms of increased competition, lower prices of goods and services, greater choice for consumers, acquisition of resources, foreign exchange earnings, access to larger markets, economies of scale, efficient resource allocation and more efficient production.
- A country has an absolute advantage if it can produce more of a good or service than another country by using the same amount of resources (or being able to produce the same amount of a good or service by using fewer resources). (HL only)
- A comparative advantage exists when a country can produce a given amount of output at a lower opportunity cost than another country, that is, it gives up fewer resources than other countries in producing a certain good or service. (HL only)
- Reasons why a country may be able to produce goods and services at lower opportunity costs than its trading partners include factor endowments, levels of technology, investment in research and development (R&D), inflation and exchange rate fluctuations. (HL only)
- The theory of comparative advantage has several limitations including it being a static model, countries distorting comparative advantages by imposing trade barriers such as tariffs and quotas, not accounting for transportation costs, and being based on the unrealistic assumption of consumers and the producers having perfect knowledge of pricing information. (HL only)

Top tip!

When evaluating the real-world relevance and limitations of comparative advantage, ensure you do two things: 1) discuss the shortcomings of the assumptions of the model of comparative advantage; and 2) consider the arguments for and against free trade.

Theory of Knowledge

Free and fairer trade are not necessarily the same thing. Discuss how the use of subject-specific language is used to enhance in our knowledge of Economics.

WISE ChoICES

International trade brings about changes to economies. Discuss the social and economic **changes** brought about by an increase in international trade and whether these changes are **equitable** for all stakeholder groups.

28 Types of trade protection

Real-world issue

Who are the winners and losers of the integration of the world's economies?

Conceptual understandings (WISE ChoICES)

- The increased **interdependence** of economies has benefits and costs.
- Increased economic integration may result in **efficiency**, welfare gains and improvements in **economic well-being** but the benefits may not result in **equity**.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- the effects of **tariffs** on markets and stakeholders (AO3)
- a diagram to show the effects of a tariff on price, production, consumption, expenditures, revenues and social welfare (AO3, AO4)
- a tariff diagram the effects on stakeholders of tariffs (HL only)
- the effects of **quotas** on markets and stakeholders (AO3)
- a diagram to show the effects of a quota on price, production, consumption, expenditures, revenues and social welfare (AO3, AO4)
- a quota diagram the effects on stakeholders of tariffs (HL only)
- the effects of an export **subsidy** on markets and stakeholders (AO3)
- a diagram to show the effects of an export subsidy on price, production, consumption, expenditures, revenues and social welfare (AO3, AO4)
- an export subsidy diagram the effects on stakeholders of tariffs (HL only)
- the effects of **administrative barriers** (standards and regulations) as a type of trade protection (AO3).

Types of trade protection (AO3, AO4)

Trade protection (or **protectionism**) refers to the deliberate attempt by a government to safeguard domestic producers from foreign competition by use of **barriers to trade**. There are four main types of trade protection, which are covered in this chapter:

1. Tariffs
2. Quotas
3. Subsidy (export subsidy)
4. Administrative barriers

David Ricardo's theory of comparative advantage (see Chapter 27) suggests that economies should specialize in and export the goods and services with which they have a relatively lower opportunity cost in producing.

◆ **Trade protection** (or **protectionism**) is the use of barriers to trade to safeguard an economy from excessive international trade and foreign competition.

◆ **Barriers to trade** are obstacles to international trade, imposed by a government to safeguard national interests by reducing the competitiveness of foreign firms.

Figure 28.1 Trade protection is used to limit foreign competition

This leads to greater efficiency in the global economy and should enable economies to expand their productive capacity (as shown by their production possibility frontier).

A concern that many governments have is that trading freely with economies that possess a comparative advantage may, in the short run, lead to lower revenues for domestic firms and possibly involve shutting down domestic industries in the long run that are less efficient than their foreign counterparts. This can cause unemployment and lower economic growth in the domestic economy. Therefore, pressure is placed upon the government to act quickly to protect the population against foreign competition.

Governments can use a variety of measures to protect their domestic producers from foreign competition. Some of these measures seek to limit the supply of imports, while others aim to generate tax revenue for the government. Economies that are members of the World Trade Organization (see Chapter 30) are bound by certain agreements that they will not engage in trade protection, so these countries become creative in their approach to protecting their domestic firms.

Tariffs (AO3, AO4)

Tariffs are a specific tax on imported goods and services. Tariffs can be issued by a government unilaterally or as a requirement of membership of a trading bloc (see Chapter 30). The aim of a tariff is to increase the costs of production for a foreign firm, which will subsequently increase the price of a good or service and make domestically produced goods more competitive. Tariffs are the most commonly used form of trade protection.

◆ **Tariffs** are a specific tax on imported goods and services.

Figure 28.2 demonstrates the impact of a hypothetical tariff on the market for wine in the UK, showing the domestic demand (D) and supply (S) curves. It should be noted that the world supply of wine (S_{world}) is assumed to be so large that, realistically, no amount of demand for wine from the UK would be able to generate enough scarcity for the price of wine to increase in the UK. Therefore, the world supply of wine under free trade is perfectly price elastic. In other words, the world can supply all of the wine demanded in the UK at a price of P_w.

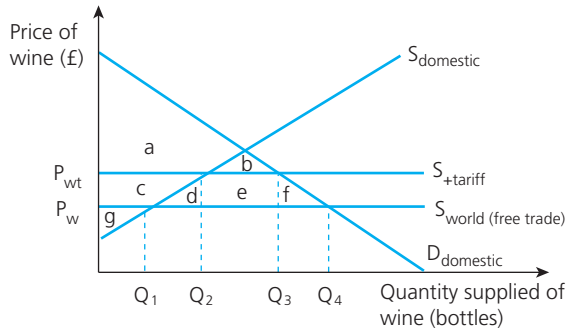


Figure 28.2 The effects of tariffs

It is assumed that foreign wine producers have a comparative advantage in wine production, which makes them able to supply wine to the UK at a lower opportunity cost than domestic wine producers. This leads to the world price (P_w) being lower than what would be the equilibrium price in this market (where the domestic demand and supply curves intersect).

UK wine producers are willing to supply only Q_1 amount of wine at a price of P_w. Therefore, there is only a small domestic producer surplus, shown by area g. At a price of P_w, the domestic demand for wine in the UK is Q_4. This means that the difference ($Q_4 - Q_1$) is provided by foreign wine producers. Hence, $Q_4 - Q_1$ represents the volume of imports prior to trade protection.

The opportunity cost of purchasing imported wine in the UK is the benefit gained from consuming UK wine. For producers, this means that every bottle of foreign wine purchased in the UK is a loss of revenue and less demand for labour in the domestic wine industry.

As a result of this, the UK government decides to impose a tariff on the importation of foreign wine. As a tariff is a specific tax, it will shift the world supply curve upwards by the amount of the tax. Depending on price elasticity of demand (PED) for wine, the price of wine in the UK (both domestic and imported wine) will increase from P_w to P_{wt}. As a result of this price increase, domestic producers have an incentive to produce more wine, indicated by an expansion in domestic supply from Q_1 to Q_2. Domestic consumers, due to the income and substitution effect of the higher world price, now have a lower quantity demanded of wine (shown by the contraction in domestic demand from Q_4 to Q_3). This results in a reduced amount of wine imports (from Q_4–Q_1 to Q_3–Q_2).

Effects of tariffs on different stakeholders

Consumers

Consumers will lose consumer surplus as a result of a tariff. There will be fewer benefits to consumers as they have to pay a higher price than free trade price, because of the imposition of the tariff on imported wine. In the long run, some foreign wine producers may choose to cease trading in the UK if the tariffs are too high, so this could reduce the variety of wine for domestic consumers. In Figure 28.2, at the world price of P_w, consumer surplus is shown by the area a+b+c+d+e+f but this shrinks to area a+b only after the imposition of the tariff. Essentially, domestic consumers lose out because they pay a higher price and have less choice as a result of this form of trade protection.

Producers

Domestic wine producers will benefit from the tariff in the short run as they are now able to compete more effectively, albeit artificially, with foreign wine producers despite their lack of comparative advantage. As a result of the tariff, domestic producers will see a considerable increase in their producer surplus, shown by the area g increasing to area c+g. This also means the higher price should lead to an increase in producer revenue because domestic demand expands from Q_1 to Q_2. The extent to which domestic producers will benefit depends on the magnitude of the tariff and the price elasticity of demand for wine.

Figure 28.3 Import taxes increase prices paid by consumers

However, foreign wine producers will lose revenue from the introduction of the tariff. They will likely have to compete on quality or leave the market if they are unable to compete on price. For example, increased advertising expenditure may be required as foreign firms try to build brand loyalty for their wines. Note that in the long run, domestic producers may experience retaliation from foreign governments and could therefore see a reduction in their export revenues due to reciprocal tariffs being imposed on wines from the UK.

Governments

The UK government will receive tax revenue from the tariff imposed on wines imported from foreign firms. This is represented by the area e in Figure 28.2. In the long run, it can be argued that the UK government will be protecting jobs in the domestic wine industry. This also means the government will not need to support those who would have lost their jobs in the wine industry through unemployment benefits, thus also benefiting taxpayers in the UK.

However, the UK government will be burdened by costs of enforcement of the tariff. There will also be administrative costs as each port of entry processes the tariffed goods. The government may also face action from the World Trade Organization (which promotes freer and fairer trade between nations). This could include fines or sanctions for infringing the terms and conditions of membership of the WTO.

Effects of tariffs on market efficiency and welfare

Overall, the impact of tariffs on a specific market can be demonstrated by the change in social surplus. Recall from Chapter 5 that **community surplus** (or **social surplus**) is the sum of producer and consumer surplus in a specific market. It is used to demonstrate the level of efficiency in a market.

◆ **Community surplus** (also known as **social surplus**) is the sum of consumer and producer surplus at a given market price and output, thereby maximizing exonomic welfare.

Before the introduction of a tariff, there was a large consumer surplus in the market for wine in the UK, shown by the area of a+b+c+d+e+f in Figure 28.2. Domestic producers enjoyed a relatively small producer surplus, represented by area g. This results in a social surplus or community surplus shown by the area a+b+c+d+e+f+g.

After the introduction of the tariff, the consumer surplus is lowered as the price of wine increases from P_w to P_{wt}. This is now represented by the area a+b. As a result of the increase in the price of wine, domestic producers see an increase in their producer surplus from g to an area of c+g. This results in a new (smaller) social surplus of area a+b+c+g.

It can therefore be said that efficiency in the market is lower after the introduction of the tariff, that is, there has been an overall welfare loss. Area d represents inefficient production by firms in the market or joining the market as a result of the tariff. It is not economically efficient for those units to be produced as they are being supported only by the tariff. Area e is the amount of government revenue generated by the tariff. Area f represents the loss to consumers who are now paying a higher price for less wine as a result of the tariff. As no stakeholder group gains area d or f after the imposition of the tariff, these represent the overall welfare loss to the economy. The welfare loss is shown the area d + f.

Theory of Knowledge

If economists claim to know that the benefits of trade outweigh the costs, how can high-income nations justify the imposition of trade protectionist measures on low-income countries?

PAPER 1 EXAM PRACTICE QUESTION 28.1

Explain the economic impacts following the imposition of a tariff. *[10 marks]*

PAPER 1 EXAM PRACTICE QUESTION 28.2

Using real-world examples, evaluate the effects of a tariff on different stakeholders. *[15 marks]*

Calculating the effects of tariffs (AO4) (HL only)

HL students need to be able to use a tariff diagram to calculate the effects of the tariff on different stakeholders. The following analysis makes reference to both Figures 28.2 and 28.4.

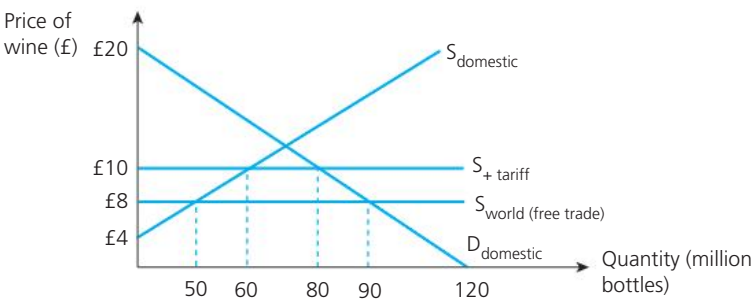


Figure 28.4 The effects of tariffs on stakeholders

Consumers

Consumers of wine in the UK will experience an increase in their wine expenditure and a loss of consumer surplus due to the higher price following the imposition of the tariff. Consumer expenditure on the product can be calculated as:

Consumer expenditure = Price × Quantity demanded

With reference to Figures 28.2 and 28.4, the change in consumer expenditure can be calculated as follows:

Pre-tariff consumer expenditure = $P_w \times Q_4$

Pre-tariff consumer expenditure = £8 × 90m

Pre-tariff consumer expenditure = £720m

Post-tariff consumer expenditure = $P_{wt} \times Q_3$

Post-tariff consumer expenditure = £10 × 80m

Post-tariff consumer expenditure = £800m

Hence, the change in consumer expenditure = £800m − £720m

Change in consumer expenditure = +£80m

Note that although consumers spend more on wine, this does not mean that producers receive this extra expenditure, as the change in consumption of wine has been caused by the imposition of the tariff.

Consumer surplus can be defined as the additional benefit to consumers willing and able to pay a price above the market price. It helps to remember that consumer surplus is the area below the demand curve and above the price (this is not the definition, but how it is calculated). Recall from Chapter 5 that as consumer surplus is usually shown as a triangular area, it can be calculated by using the formula:

$$\text{Area of a triangle of consumer surplus} = \frac{(\text{base} \times \text{height})}{2}$$

In Figure 28.4, the base of the consumer surplus triangle is the difference between the quantity demanded at the prevailing price and 0. In Figure 28.4, this is 90m bottles of wine at the unit price of £8 before the tariff. The height of the consumer surplus triangle is the difference between the price intercept of demand (the price at which demand is 0) and the prevailing price, that is, £20 – £8 = £12

$$\text{Consumer surplus} = \text{Quantity demanded} \times \frac{(\text{Price intercept of demand} - \text{World price})}{2}$$

Hence, the pre-tariff level of consumer surplus is calculated as:

$$\text{Pre-tariff consumer surplus} = \frac{90\text{m} \times (£20 - £8)}{2}$$

$$\text{Pre-tariff consumer surplus} = \frac{90\text{m} \times £12}{2}$$

$$\text{Pre-tariff consumer surplus} = \frac{£1{,}080\text{m}}{2}$$

Pre-tariff consumer surplus = £540m

After the introduction of the tariff, the price of a bottle of wine in the UK increases to £10, and as a result the quantity demanded falls to 80m bottles. After the introduction of the tariff, the change in consumer surplus is calculated as:

$$\text{Consumer surplus} = \text{Quantity demanded} \times \frac{(\text{Price intercept of demand} - \text{World price})}{2}$$

$$\text{Post-tariff consumer surplus} = \frac{80\text{m} \times (£20 - £10)}{2}$$

$$\text{Post-tariff consumer surplus} = \frac{80\text{m} \times £10}{2}$$

$$\text{Post-tariff consumer surplus} = \frac{£800\text{m}}{2}$$

$$\text{Post-tariff consumer surplus} = £400\text{m}$$

Therefore, the change in consumer surplus can be calculated as:

$$\text{Change in consumer surplus} = \text{New consumer surplus} - \text{Old consumer surplus}$$

$$\text{Change in consumer surplus} = £400\text{m} - £540\text{m}$$

$$\text{Change in consumer surplus} = -£140\text{m}$$

Producers

Domestic producers of wine in the UK will experience an increase in their revenue and a gain in producer surplus. Foreign wine producers exporting to the UK will see a fall in their revenue due to the tax. Producer revenue can be calculated by using the following formula:

$$\text{Producer revenue} = \text{Price} \times \text{Quantity}$$

The change in domestic producer revenue can be calculated as:

$$\text{Pre-tariff producer revenue} = P_w \times Q_1$$

$$\text{Pre-tariff producer revenue} = £8 \times 50\text{m}$$

$$\text{Pre-tariff producer revenue} = £400\text{m}$$

$$\text{Post-tariff producer revenue} = P_{wt} \times Q_2$$

$$\text{Post-tariff producer revenue} = £10 \times 60\text{m}$$

$$\text{Post-tariff producer revenue} = £600\text{m}$$

$$\text{Change in producer revenue} = \text{Post-tariff revenue} - \text{Pre-tariff revenue}$$

$$\text{Change in producer revenue} = £600\text{m} - £400\text{m}$$

$$\text{Change in producer revenue} = +£200\text{m}$$

The change in revenue for foreign producers can be calculated by using the formula:

$$\text{Foreign producer revenue} = \text{World price} \times \text{Quantity of imports}$$

The quantity of imports can be calculated by using the following equation:

$$\text{Quantity of imports} = \text{Quantity demanded} - \text{Domestic quantity supplied}$$

Therefore, foreign producer revenue can be calculated as follows:

$$\text{Pre-tariff foreign producer revenue} = £8 \times (90\text{m} - 50\text{m})$$

$$\text{Pre-tariff foreign producer revenue} = £8 \times 40\text{m}$$

$$\text{Pre-tariff foreign producer revenue} = £320\text{m}$$

Post-tariff foreign producer revenue = £10 × (80m − 60m)

Post-tariff foreign producer revenue = £10 × 20m

Post-tariff foreign producer revenue = £200m

Change in foreign producer revenue = Post-tariff revenue – Pre-tariff revenue

Change in foreign producer revenue = £200m − £320m

Change in foreign producer revenue = −£120m

Producer surplus can be defined as the additional benefit to producers willing and able to produce or supply a good or service at a price below the market price. It can help to remember producer surplus as the area above the supply curve and below the market price.

$$\text{Domestic producer surplus} = \text{Domestic quantity supplied} \times \frac{(\text{World price} - \text{Price intercept of supply})}{2}$$

$$\text{Pre-tariff producer surplus} = \frac{50\text{m} \times (£8 - £4)}{2}$$

$$\text{Pre-tariff producer surplus} = \frac{50\text{m} \times £4}{2}$$

$$\text{Pre-tariff producer surplus} = \frac{£200\text{m}}{2}$$

Pre-tariff producer surplus = £100m

After the introduction of the tariff, the price of a bottle of wine in the UK increases to £8, and as a result the domestic quantity supplied increases to 60m bottles. After the introduction of the tariff, the domestic producer surplus is calculated as:

$$\text{Domestic producer surplus} = \text{Domestic quantity supplied} \times \frac{(\text{World price} - \text{Price intercept of supply})}{2}$$

$$\text{Post-tariff producer surplus} = \frac{60\text{m} \times (£10 - £4)}{2}$$

$$\text{Post-tariff producer surplus} = \frac{60\text{m} \times £6}{2}$$

$$\text{Post-tariff producer surplus} = \frac{£360\text{m}}{2}$$

Post-tariff producer surplus = £180m

Therefore, the change in domestic producer surplus can be calculated as:

Change in domestic producer surplus = New producer surplus – Old producer surplus

Change in producer surplus = £180m − £100m

Change in producer surplus = +£80m

Government

The UK government will receive revenue from the tariff, which is collected from foreign firms selling their wines in the UK. The formula for calculating government tariff revenue is:

Government revenue = Tariff × Quantity of imports demanded

Therefore, government revenue can be calculated as:

Tariff = £10 − £8 = £2

Quantity of imports demanded = 80m – 60m = 20m

Government revenue = £2 × 20m

Government revenue = £40m

Welfare (society)

While there are winners and losers from the imposition of a tariff, overall, there is a social or welfare loss, as shown in Figure 28.5. Following the imposition of a tariff, note the following impact on various stakeholders:

- Consumer surplus falls by the area c+d+e+f.
- Domestic producer surplus increases by the area c.
- The government earns area e.
- Hence, area d+f represents a welfare loss to the society.

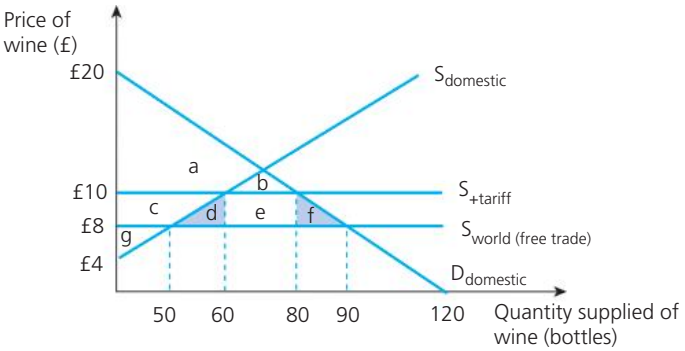


Figure 28.5 Welfare loss from the imposition of a tariff

PAPER 3 EXAM PRACTICE QUESTION 28.3 (HL ONLY)

Use the diagram (Figure 28.6) below to calculate the following:

a Consumer surplus before the imposition of the tariff. *[2 marks]*

b Producer surplus before the imposition of the tariff. *[2 marks]*

c Consumer surplus after the imposition of the tariff. *[2 marks]*

d Producer surplus after the imposition of the tariff. *[2 marks]*

e Revenue to the government after the imposition of the tariff. *[2 marks]*

f The welfare loss after the imposition of the tariff. *[2 marks]*

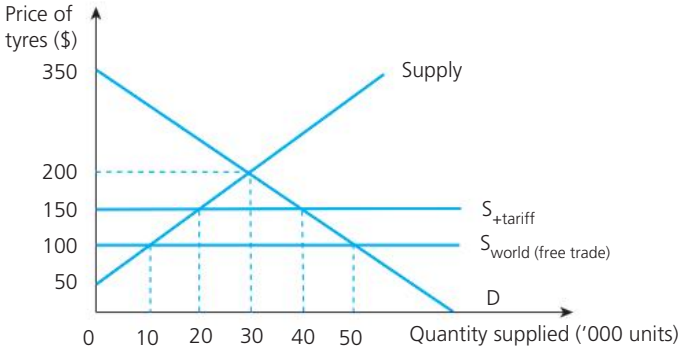


Figure 28.6 The imposition of a tariff

Quotas (AO3, AO4)

Tariffs are the main method of trade protection. All other forms of trade protection are collectively known as **non-tariff barriers**. These are restrictions to international trade that do not involve a tax or duty, such as quotas, export subsidies and administrative barriers to trade.

The main form of non-tariff barriers are **quotas**. These are quantitative limits on the importation of a good into a country. Quotas can be issued by a government unilaterally or as a requirement of membership of a trading bloc (see Chapter 30). The aim of a quota is to restrict the supply of a good or service at the expense of foreign firms. As a result of the imposition of a quota, the fall in supply (caused by the reduction in imports) will lead to an increase in the price of the product. The impacts of a quota are illustrated in Figure 28.7.

◆ **Non-tariff barriers** are restrictions to international trade that do not involve a tax or duty (tariff), such as quotas, export subsidies and administrative barriers.

◆ **Quotas** are quantitative limits on the importation of a good into a country.

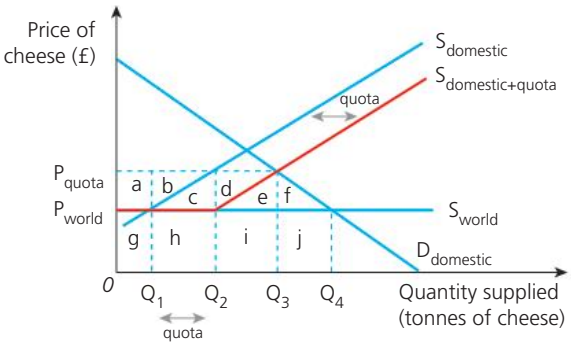


Figure 28.7 The effects of a British quota on cheese production

Suppose that the British government decides to protect its domestic cheese producers by imposing a quota on all cheese products entering its country. As can be seen in Figure 28.7, prior to the quota being imposed, the domestic quantity demanded of cheese is represented by $0Q_4$ at the world price of P_{world}, whereas the domestic production of cheese is only $0Q_1$ due to the lower price. This means the original volume of imports is $Q_4 - Q_1$. It should be noted that the world supply of cheese (S_{world}) is assumed to be so large that (due to competition from around the world) no amount of demand for cheese from British consumers would be high enough to generate sufficient scarcity for the price of cheese to increase. Therefore, the world supply of cheese is perfectly price elastic, that is, the world can supply all of the demand for cheese in Britain at a price of P_{world}.

The introduction of the quota from Q_1 to Q_2 creates greater scarcity in the market, thus causes an increase in the price of cheese from P_{world} to P_{quota}, where demand now equals supply at $0Q_3$ (the total quantity of cheese traded). Even though domestic supply increases, the reduction in imports means that it is no longer possible to import cheese products below the price of P_{quota}. The domestic supply curve shifts outwards by the amount of the quota, from $S_{domestic}$ to $S_{domestic+quota}$. Domestic supply of cheese by British producers increases from Q_1 to $Q_1 + (Q_3 - Q_2)$. Importers make up the rest of the supply, based on the quota, that is, $Q_2 - Q_1$.

Effects of quotas on different stakeholders

Consumers

Consumers in Britain will have to pay more for their cheese, as price rises from P_{world} to P_{quota}. This is because the quota will make cheese scarcer in the UK and allow British producers to charge higher prices. In addition, there will be a reduction in the importation of various cheese products from abroad, which is likely to lead to a reduction in choice for consumers.

In Figure 28.7, consumer expenditure on cheese in Britain will have changed from area g+h+i+j to a+b+c+d+e+g+h+i. Also, note that similar to the tariff diagram, the loss in consumer surplus due to the higher price is shown by the area a+b+c+d+e+f.

Producers

Domestic cheese producers will see an increase in their producer surplus. This is because the additional benefit to those producers willing to provide cheese at a price lower than equilibrium will increase. This will raise the revenue for domestic cheese producers, based on the higher spending on cheese by British consumers (that is, from area g+h+i+j to area a+g+d+e+i). Domestic producers may also expand their range of goods and services to provide domestic substitutes to the (restricted types of) imported cheese. The increased cheese production by domestic producers may lead to economies of scale, which could lead to an increase in their profits.

Foreign producers will usually see a reduction in their export revenues as their supply becomes restricted. While they will receive a higher price for all of the cheese that is within the quota ($Q_2 - Q_1$), the limit on their supply is likely to lead to lower revenues. In Figure 28.7, foreign producer revenue will change from areas h+i+j to b+c+h. Whether this constitutes an increase or a decrease in revenue will depend on price elasticity of demand and supply. However, in this particular case, it can be seen that the loss from fewer imports (area i+j) is greater than the gain from the higher price (area b+c). Therefore, it is likely that foreign producers may need to refocus their efforts on exporting to other countries or to concentrate on expanding in their own domestic cheese market.

Government

Technically, the British government will receive some revenue from the cheese quota as governments typically limit the supply of a good with a quota through the sale of import licences. This means that they are compensated with proceeds from the sale of these licences. It is likely that there are some enforcement costs associated with a quota, as governments have to administer the quota, sell the licences and deter smuggling operations. These financial costs and benefits of a quota are not illustrated in the quota diagram.

Welfare (society)

Although domestic producers are made better off from the imposition of a quota, this comes at the expense of efficiency in the market. Following the quota on cheese imports, there will be a welfare loss present in the market. This is due to inefficient production (caused by lower levels of competition and less choice for domestic consumers) and higher prices for the product. This is represented in Figure 28.8 by the triangular area e+f. The base of the triangle is the loss in imports of cheese, and the height of the triangle represents the higher price of cheese. Hence, in terms of welfare and efficiency in the British cheese market, society is made worse off.

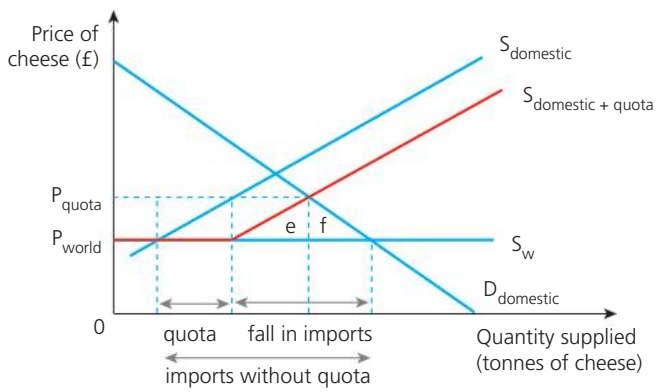


Figure 28.8 The effect of a British cheese quota on market efficiency

PAPER 1 EXAM PRACTICE QUESTION 28.4

Explain the economic impacts on different stakeholders following the imposition of a quota. *[10 marks]*

Calculating the effects of quotas (AO4) (HL only)

You need to be able to use a quota diagram to calculate the effects of the quota on different stakeholders. The following analysis refers to Figure 28.9.

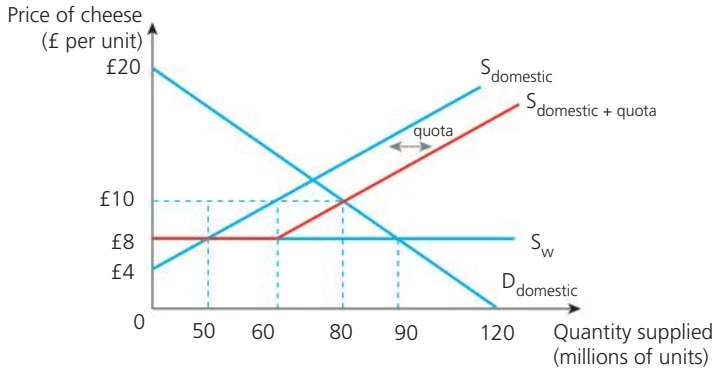


Figure 28.9 The effect of a 10-million-unit cheese quota on stakeholders

Consumers

As a result of the 10-million-unit quota on imported cheese, British consumers will experience an increase in price from £8 to £10 per unit. Owing to the income and substitution effect of the higher price, this will lead to a reduction in quantity demanded from 90 million units to 80 million units of cheese.

The impact of the quota on consumer expenditure can be calculated as:

Consumer expenditure = Quantity demanded × Price

Pre-quota consumer expenditure = 90m × £8

Pre-quota consumer expenditure = £720m

Post-quota consumer expenditure = 80m × £10

Post-quota consumer expenditure = £800m

Change in consumer expenditure = £800m − £720m

Change in consumer expenditure = +£80m

Producers

As a result of the cheese quota, domestic firms will be able to sell their cheese at a higher price on the market. The higher price also means that they will be willing and able to provide more cheese after the quota.

Producer revenue can be calculated by using the following formula:

Producer revenue = Price × Quantity

The change in domestic producer revenue is calculated as:

Pre-quota producer revenue = $P_{world} \times Q_1$

Pre-quota producer revenue = £8 × 50m

Pre-quota producer revenue = £400m

Post-quota producer revenue = $P_{quota} \times [(Q_1 + (Q_3 - Q_2)]$

Post-quota domestic supply = 50m + (80m – 60m)

Post-quota producer revenue = £10 × (50m + 20m)

Post-quota producer revenue = £10 × 70m

Post-quota producer revenue = £700m

Change in producer revenue = New revenue – Old revenue

Change in producer revenue = £700m − £400m

Change in producer revenue = +£300m

The change in foreign producer revenue is calculated as follows:

Producer revenue = Price × Quantity

Pre-quota foreign producer revenue = $P_{world} \times (Q_4 - Q_1)$

Pre-quota foreign producer supply = 90m – 50m = 40m

Pre-quota foreign producer revenue = £8 × 40m

Pre-quota foreign producer revenue = £320m

Post-quota foreign producer revenue = P_{quota} × Quota

Post-quota foreign producer supply = 60m − 50m = 10m

Post-quota foreign producer revenue = £10 × 10m

Post-quota foreign producer revenue = £100m

Change in foreign producer revenue = New revenue – Old revenue

Change in foreign producer revenue = £100m − £320m

Change in foreign producer revenue = −£220m

ATL ACTIVITY 28. 1

With a partner, research a list of common goods on which a country imposes tariffs and quotas. Reflect on why some of these have tariffs imposed and others have quotas.

Government

As mentioned earlier, technically, the British government will receive some revenue from the cheese quota through the sale of import licences. Therefore, the total government revenue from this will be calculated as:

Government revenue = Import licence fee × Number of import licences

Nevertheless, as this cannot be determined from the quota diagram, you will not be required to calculate the numerical impact of a quota on government revenues.

Subsidies (AO3, AO4)

Subsidies are a form of financial assistance to local firms to help them compete against foreign rivals by lowering the costs of production for domestic suppliers. There are two types of subsidies that can be used to protect domestic industries:

- **Production subsidies** are the most common form of trade subsidies, used to help reduce the production costs of domestic firms.
- **Export subsidies** are less wide-ranging as they are targeted at protecting specific export-orientated firms.

◆ **Subsidies** are a form of financial assistance to domestic firms by lowering their costs of production in order to help them compete against foreign imports.

The analysis below refers to production subsidies but can equally be applied to export subsidies.

Figure 28.10 represents a production subsidy to the producers of baby milk powder in the USA. As with the case of subsidies examined in microeconomics (see Chapter 10), the subsidy used to protect domestic producers will lower the costs of production for local firms producing baby milk powder. This will make them more competitive against foreign firms in the market, ceteris paribus.

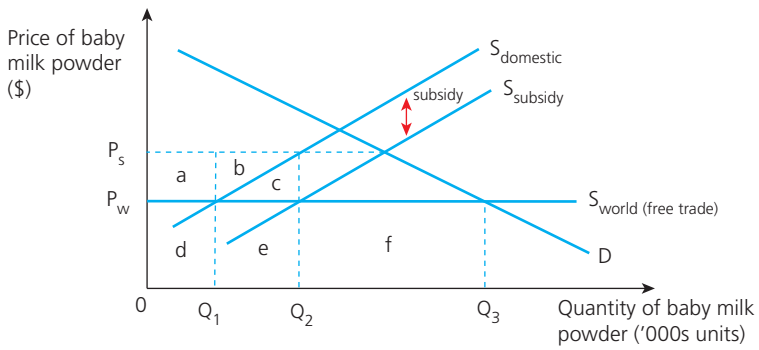


Figure 28.10 The effects of subsidies

Prior to the introduction of the subsidy, domestic firms were only willing and able to produce and supply a quantity of $0Q_1$ at a world price of P_w. The quantity demanded for baby milk powder is $0Q_3$. The introduction of the subsidy shifts the domestic supply curve downwards by the value of the subsidy. This intersects the world price (P_w) at a quantity of $0Q_2$. The domestic firms, while selling at the world price, are actually receiving a price equivalent to P_s (on the original supply curve) due to them receiving the subsidy. This will reduce the importation of baby milk powder from $Q_3 - Q_1$ to $Q_3 - Q_2$. It is important to remember that consumers continue to buy $0Q_3$ units of baby milk powder at a price of P_w before and after the subsidy (the subsidy made to domestic producers of baby milk powder does not change the world price).

Figure 28.11 Agriculture is the one of the most heavily subsidized sectors in the world

The effects of subsidies on stakeholders

The analysis below refers to Figure 28.10, which represents a production subsidy to domestic producers of baby milk powder in the USA.

Consumers

The total consumption of baby milk powder in the USA is not affected by the subsidy as the price remains at the world price, P_w, so $0Q_3$ is still the quantity traded. Consumers do, however, buy more domestic goods so whether they gain or lose depends on the relative quality of the foreign goods relative to the baby milk powder produced by US firms. Hence, consumer expenditure before and after the subsidy is represented by the areas d+e+f (that is, $P_w \times Q_3$). As there is no change in the price paid by consumers (P_w) or the amount they purchase (Q_3), there is no change in consumer surplus.

Producers

Domestic producers gain from the subsidy as they are able to supply more (Q_2 rather than Q_1). Note that domestic firms would still sell their products at the world price of P_w after receiving the subsidy. This means that the domestic producer revenue before the subsidy is represented by area d, whereas after the subsidy this increases to area a+b+c+d+e.

Foreign producers were able to sell $Q_3 - Q_1$ amount of imports prior to the US government subsidizing their local firms. However, the subsidy reduces the amount of imports to $Q_3 - Q_2$. Foreign producer revenue will change from area e+f to area e.

Government

There is a negative impact on the government's budget due to the expenditure on subsidies. The government expenditure on the subsidy is represented by area a+b+c, which is the difference between the price that domestic producers receive (P_s) and the world price that they sell at (P_w), for all of the units that the domestic producers supply (Q_2). Taxpayers are worse off as there is an opportunity cost of using this money for subsidizing US producers of baby milk powder.

Welfare (society)

After the approval of a subsidy for domestic firms, it is important to remember that there is no change in the price paid by consumers (P_w) or the amount of baby milk powder that they purchase (Q_3). Hence, there is no change in the welfare for consumers, unless the quality of US baby milk powder is inferior to that made by foreign firms. Producers gain additional revenue by being able to sell more (the gain in sales revenue is shown by area e), plus the value of the subsidy (shown by the area a+b+c). However, the area of the subsidy also represents the opportunity cost to US taxpayers.

In the case of market efficiency, the subsidy will attract new domestic producers to the baby milk powder industry or encourage existing producers to expand their production past what would comply with the law of comparative advantage (see Chapter 30). In other words, the subsidy encourages inefficient output from domestic firms. This generates a welfare loss represented by area c – the triangular area represents the subsidized units of inefficient production by domestic firms ($Q_2 - Q_1$) and the value of the subsidy ($P_s - P_w$).

Common mistake

When referring to welfare loss in your diagrams (be it for tariffs, subsidies to quotas) don't fall short with the explanation. Make sure you understand how protectionist measures distort market forces and lead to an inefficient allocation (or misallocation) of resources. Protecting inefficient or uncompetitive domestic firms in an artificial way tends to raise prices and reduce choice for consumers, leading to a net loss to society.

PAPER 1 EXAM PRACTICE QUESTION 28.5

Explain the economic impacts on different stakeholders following the provision of a producer subsidy to protect domestic producers. *[10 marks]*

Calculating the effects of subsidies (AO4) (HL only)

HL students need to be able to use a subsidy diagram to calculate the effects of the subsidy on different stakeholders. The following analysis refers to both Figures 28.10 and 28.12.

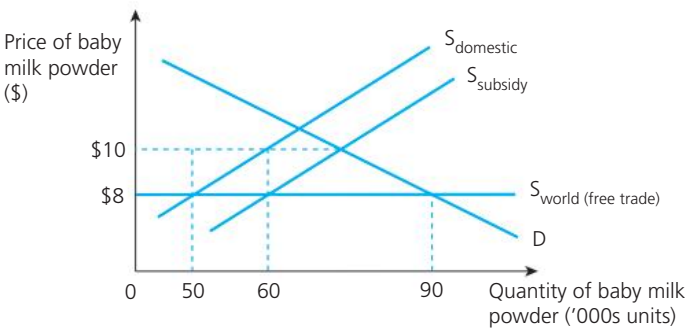


Figure 28.12 Calculating the effects of subsidies

Consumers

As a result of a $2 subsidy shown in Figure 28.12, domestic consumers are not impacted as the price that they were paying before the subsidy ($8) remains the same after the subsidy paid to local firms. Therefore, there is no change in consumer expenditure.

Consumer expenditure = Quantity demanded × Price

Pre-subsidy and post-subsidy consumer expenditure = 90,000 × $8

Pre-subsidy and post-subsidy consumer expenditure = $720,000

Producers

As a result of the $2 subsidy, domestic firms will effectively receive a higher price of $10 (that is, $8 + $2) for their baby milk powder. Domestic firms will also expand their production from 50,000 units to 60,000 units. Producer revenue can be calculated by using the following formula:

Producer revenue = Price × Quantity

The change in domestic producer revenue is calculated as:

Pre-subsidy producer revenue = $P_w \times Q_1$

Pre-subsidy producer revenue = $8 × 50,000

Pre-subsidy producer revenue = $400,000

Post-subsidy producer revenue = $P_s \times Q_2$

Post-subsidy producer revenue = $10 × 60,000

Post-subsidy producer revenue = $600,000

Change in producer revenue = New revenue − Old revenue

Change in producer revenue = New revenue = $600,000 − $400,000

Change in producer revenue = New revenue = +$200,000

The change in foreign producer revenue is calculated as follows:

Producer revenue = Price × Quantity

Pre-subsidy foreign producer revenue = $P_{world} \times (Q_3 - Q_1)$

Pre-subsidy foreign producer revenue = $8 × (90,000 − 50,000)

Pre-subsidy foreign producer revenue = $8 × 40,000

Pre-subsidy foreign producer revenue = $320,000

Post-subsidy foreign producer revenue = $P_{world} \times (Q_3 - Q_2)$

Post-subsidy foreign producer revenue = $8 × (90,000 − 60,000)

Post-subsidy foreign producer revenue = $8 × 30,000

Post-subsidy foreign producer revenue = $240,000

Change in foreign producer revenue = New revenue − Old revenue

Change in foreign producer revenue = $240,000 − $320,000

Change in foreign producer revenue = New revenue = −$80,000

Government

By introducing the $2 producer subsidy, the government will commit to the following expenditure:

Government expenditure = Subsidy × Post-subsidy domestic supply

Government expenditure = $2 × Q_2

Government expenditure = $2 × 60,000

Government expenditure = $120,000

Top tip!

When evaluating the use of subsidies as a type of trade protection, bear in mind that they can help to increase domestic employment (as local firms are able to increase their output) but the counter-argument is that there is a huge opportunity cost to subsidizing inefficient firms in the economy.

PAPER 3 EXAM PRACTICE QUESTION 28.6 (HL ONLY)

Use the diagram (Figure 28.13) below to calculate the following:

a The cost of the subsidy to the government. *[2 marks]*

b The amount of imports before government intervention. *[2 marks]*

c The amount of imports after government intervention. *[2 marks]*

d The total amount spent by domestic consumers under free trade. *[2 marks]*

e The total amount spent by domestic consumers after the imposition of the subsidy. *[2 marks]*

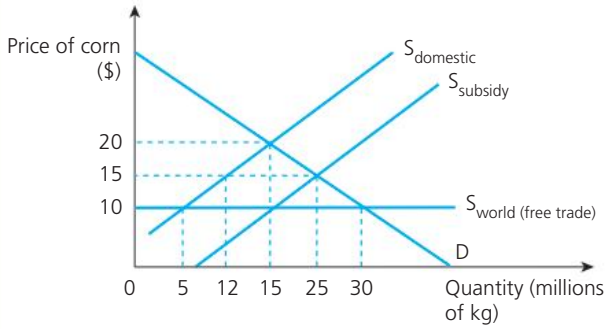


Figure 28.13 The imposition of a subsidy

Administrative barriers (AO3)

Administrative barriers are the application of standards and regulations (bureaucratic rules and policies) imposed on foreign firms. Essentially, they are a form of trade protection to safeguard domestic firms. Governments often use bureaucratic rules and regulations as a form of trade protection to also safeguard domestic consumers, such as strict rules regarding food safety, environmental standards and product quality.

◆ **Administrative barriers** are the application of bureaucratic standards and regulations imposed on foreign firms in order to protect domestic firms and consumers.

Complying with these administrative trade barriers increases the costs for foreign firms, thereby giving an advantage to domestic firms. This can be due to the costs of complying with rules and regulations (such as meeting strict health and safety standards). It can also be due to the additional time needed to get imports into the country, that is, administrative barriers slow the supply chain of foreign goods and services, which creates a shortage so enables the gap to be filled by domestic firms.

An example of administrative barriers is the Chinese government's protection of the domestic publishing industry. All textbooks (including this one!) are checked by customs control to ensure any written material or images that might be deemed to be offensive or inappropriate, from the government's perspective, are not allowed into the country. This bureaucratic process significantly reduces the importation of textbooks published overseas and protects the Chinese publishers from foreign competition.

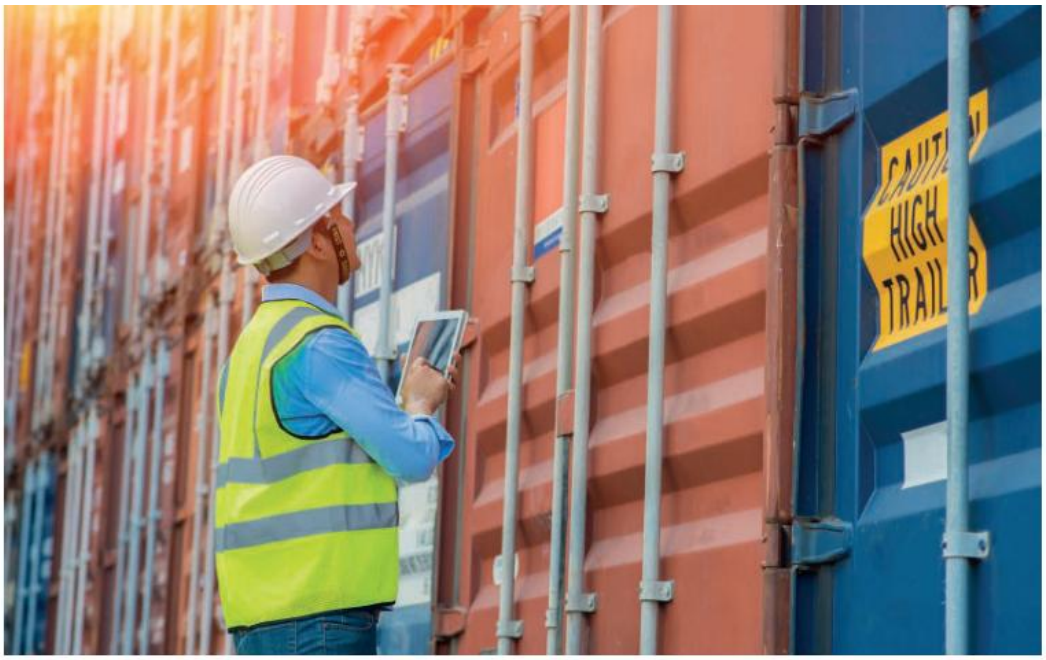


Figure 28.14 Import controls are an important, but often bureaucratic, process

Other examples of trade barriers include:

- **Embargoes** are bans on trade with a certain country, often due to political and/or economic disputes (although they can be used for health and safety reasons for specific products too). An embargo rarely benefits local consumers who suffer from a lack of choice and higher prices due to the lack of supply. It can also trigger retaliation from other countries.
- **Exchange controls** are trade restrictions on the quantity of foreign exchange that can be bought or sold by domestic residents, such as daily limits on the amount of foreign currency that tourists or investors can exchange. This limits the amount of foreign exchange available to importers, thereby restricting the volume of foreign trade.

◆ **Embargoes** are a form of administrative barrier that involves the use of bans on trade with a certain country, often due to political and/or economic disputes.

◆ **Exchange controls** are a form of administrative barrier involving restrictions on the quantity of foreign exchange that can be bought or sold by domestic residents.

CASE STUDY

Oktoberfest

Oktoberfest is a large festival held in Munich, Germany, each year (usually at the beginning of September). Typically, around 6 million people engage in the folk festival, which involves singing, dancing, traditional dress and the drinking of German beer.

Beer in Germany is brewed under a strict series of regulations called *Reinheitsgebot* meaning 'purity order', which specifies water, malted barley, hops and yeast as the only ingredients. The purity regulation has been in place since 1516! This has favoured domestic producers and requires foreign producers of beer to seek permission to import variations of beer into Germany. Some economists argue that this is a form of administrative barrier, giving domestic firms a competitive advantage.

Figure 28.15 Oktoberfest in Munich, Germany

WISE ChoICES

Given the growing **interdependence** of the world's economies, do trade protectionist measures through government intervention bring about greater levels of **economic well-being** for individuals and societies?

PAPER 1 EXAM PRACTICE QUESTION 28.7

Using real-world examples, evaluate the arguments for trade protection. *[15 marks]*

WISE ChoICES

Free trade is argued to lead to greater **efficiency**, yet there are both winners and losers of international trade. Protectionism requires **intervention**. How can economists determine whether efficiency or intervention is more of a priority?

ATL ACTIVITY 28.2

Investigate the use of administrative barriers as a potential form of trade protection for a product and countries of your choice. The example of the Volkswagen scandal is one example – but try to choose your own.

Chapter summary

- Protectionism (trade protection) refers to deliberate attempts by a government to safeguard domestic producers from foreign competition.
- Barriers to trade are obstacles to free trade, imposed by a government to protect national interests by reducing the competitiveness of foreign firms. These include tariffs, quotas, subsidies and administrative barriers.
- Tariffs are a specific tax on imported goods and services. There is a loss in consumer surplus as a result of a tariff. Domestic producers will benefit in the short run as they are now able to compete more effectively, albeit artificially. The domestic government receives revenue from the tariff.
- Quotas are quantitative limits on the importation of a good into a country. The aim is to restrict the supply of an imported product at the expense of foreign firms. Consumer expenditure will increase as a result of the quota. Domestic producers will see an increase in their producer surplus.
- Subsidies are a form of financial assistance to local firms by lowering the costs of production to help them compete against foreign firms. The consumption of the good is not affected as the price remains at the world price, but domestic firms gain from receiving the producer subsidy and are able to supply more. There is a negative impact on the government's budget due to the expenditure on subsidies.
- After the imposition of a tariff, quota or subsidy there will be a welfare loss in the market. This is due to protection of inefficient production of domestic firms and a loss to consumers.
- Administrative barriers are the application of standards and regulations (bureaucratic rules and policies) imposed on foreign firms in order to protect domestic industries. Examples include strict rules regarding food safety, environmental standards and product quality.
- Embargoes are bans on trade with a certain country, often due to political and/or economic disputes (although they are often used for health and safety reasons).
- Exchange controls are restrictions on the quantity of foreign exchange that can be bought or sold by domestic residents, thus limiting the amount of foreign exchange available to importers.

29 Arguments for and against trade control/protection

Real-world issue

Who are the winners and losers of the integration of the world's economies?

Conceptual understandings (WISE ChoICES)

- The increased **interdependence** of economies has benefits and costs.
- Increased economic integration may result in **efficiency**, welfare gains and improvements in **economic well-being** but the benefits may not result in **equity**.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:
- the arguments for trade protection, that is, the advantages of trade protection (AO2)
- the arguments against trade protection, that is, the disadvantages of trade protection (AO2)
- the argument of free trade versus trade protection (AO3).

Arguments for trade protection/ advantages of trade protection (AO2)

Trade protection is the use of barriers to international trade to safeguard an economy from excessive foreign trade and competition from overseas firms. The previous chapter explored the various ways in which governments attempt to protect their domestic industries from foreign competition (namely through tariffs, quotas, subsidies and administrative barriers). This chapter examines a range of arguments that economists put forward to justify or condemn those who practise trade protection. While governments feel a commitment to domestic firms that face foreign competition in their markets, the same governments may have also made commitments to various trading blocs or international bodies such as the World Trade Organization (see Chapter 30).

◆ **Trade protection** is the use of barriers to international trade, such as tariffs and quotas, to safeguard an economy from excessive foreign trade and competition from overseas firms.

The IB Economics syllabus specifies 10 interrelated arguments or reasons for trade protection: protection of infant (sunrise) industries, national security, health and safety, environmental standards, anti-dumping, unfair competition, balance of payments correction, government revenue, the protection of jobs, and economically least developed country (ELDC) diversification.

The following analysis contains varying degrees of economic theory, and the reader may consider that some of these arguments are more political in nature than economic. It should also be noted that the arguments put forward in favour of an idea are often just an advantage of that idea.

Protection of infant industries

Infant industries (or **sunrise industries**) are usually new or unestablished industries that have the potential to achieve a comparative advantage but are too underdeveloped to compete with foreign firms, so need government protection. For example, the Hong Kong SAR (Special Administrative Region of the People's Republic of China) government protects its movie industry by imposing quotas on the number of foreign films allowed into the SAR each month.

◆ **Infant industries** are often new or unestablished industries that have the potential to achieve a comparative advantage but are too underdeveloped to compete with foreign firms.

A common argument against the use of comparative advantage as a determinant of economic efficiency is that it is often a 'static' rather than 'dynamic' view. Some governments argue that economies that have already achieved a comparative advantage have an unfair advantage because free trade prevents any other economy competing with them. Governments can claim that their economy has the potential to produce a good or service at a lower opportunity cost than the current comparatively advantageous economy but are unable to achieve the economies of scale needed under free trade. Governments may see this as grounds to give assistance to the domestic firms with the hope of them achieving economies of scale and enjoying improved economic efficiency in the long run.

However, a major criticism is that some 'infant' industries are protected for very long periods of time – the Hong Kong movie industry has been protected for decades. Hence, the arguments for protecting an infant or sunrise industry are somewhat subjective and political, rather than economic.

Figure 29.1 The Hong Kong SAR protects its domestic film industry

National security

Some governments insist that it is in the national interest for some essential industries to remain 'self-sufficient' in case of times of major crises, such as global political conflict, trade wars or armed wars. Should a war arise, an economy may be cut off from international markets through embargoes, sanctions or more drastic measures. If this were to happen,

the domestic government would lose the conflict if they were not self-sufficient. Therefore, tariffs, quotas and/or administrative barriers may be placed on foreign firms to help protect domestic industries.

Naturally, there are some industries in an economy that contain trade secrets and confidentiality, which are valuable to the security of the nation. The Nuclear Weapons project in Iran would be an example of a closely guarded industry of national importance. While it might be more efficient for a foreign firm to run these industries, it would not be in the national interest for these secrets to be widely known.

Governments may also block trade with a particular economy owing to the prevailing political ideology in that economy. They may see this as an ethical responsibility and therefore in the national or even international interest. For example, the US government has used arms embargoes on the export of weapons and armaments to several countries, including Afghanistan, China, Indonesia, Iran and North Korea.

Health and safety

A government may have concerns about the safety of particular goods that are entering an economy from overseas. Different economies have different health and safety standards for a wide range of goods such as food, medicines and electrical goods. Should an imported good not meet the domestic government's health and safety standards then it may introduce trade protection methods. The USA, for example, has used such trade restrictions on Chinese products such as smartphones and home appliance goods. It should be noted here that only an embargo would address this particular issue, whereas tariffs and quotas would still allow the goods to enter the domestic market.

ATL ACTIVITY 29.1

Investigate the use of embargoes for a country of your choice. To what extent do you think the bans were justified on economic grounds?

CASE STUDY

The Volkswagen scandal

In September 2015, the US Environmental Protection Agency (EPA) issued a notice of violation of the Clean Air Act to Volkswagen (VW). The EPA found that the German carmaker had deliberately chipped its diesel engines to cheat laboratory emissions tests. The nitrogen oxides emitted from VW engines (which causes air pollution) were about 40 times higher than what is allowed in the USA. The EPA found that VW had been deceptive with the truth, which resulted in the company's share price plummeting and the CEO resigning over the scandal. It was found that around 11 million VW vehicles made between 2009 and 2015 were programmed to bypass environmental standards and regulations, with about 500,000 of these cars exported to the USA. VW stated it would spend $18.32bn to rectify the emissions scandal; a very costly price to pay for a crisis that could have been handled better. In April 2017, a US federal judge ordered VW to pay a $2.8bn criminal fine for 'rigging diesel-powered vehicles to cheat on government emissions tests'.

Figure 29.2 The Volkswagen diesel scandal cost the company tens of billions of dollars

Environmental standards

As environmental protection becomes more of a priority for governments across the globe, some domestic firms have started to relocate their production facilities to avoid stricter controls and fines. They then import their goods back into the economy. This makes it hard for the firms that stayed domestically (and adhered to cleaner production guidelines) to compete. In this case, a government might introduce protectionist measures such as tariffs and quotas on the imported (and more environmentally harmful) goods in order to protect the domestic firms and to reduce pollution. Essentially, protection is used to upload environmental standards of the country – see Volkswagen scandal case study in Chapter 29 (page 457).

Anti-dumping

Dumping is the sale of products by foreign firms at a price lower than the cost of production. Protection is therefore needed to prevent dumping from foreign firms. It also occurs when importers flood the domestic market with a product that is deliberately priced lower than the price charged by local firms, who are therefore not able to compete. As dumping typically involves foreign firms selling a substantial volume of the product in overseas markets, it often threatens the survival of local businesses. Hence, there is an argument to protect domestic firms against dumping from large foreign multinational companies. For example, in 2018, an anti-dumping investigation by the Chinese government found that Chlorobutyl rubber, used to make car tyres, was being dumped by the USA, the EU and Singapore at the expense of local firms in China.

◆ **Dumping** is the sale of goods and services by foreign firms at a price lower than the cost of production.

Figure 29.3 China has accused certain countries of dumping rubber used to make car tyres

There is an argument that when foreign governments subsidize a particular good or service, and then that good or service is exported, the competition becomes unfair. Subsidies are a form of financial assistance to local firms to help them compete against foreign imports by lowering the costs of production (see Chapter 28). Governments then feel justified in taxing that good (via the imposition of tariffs) as a way to correct the price. For example, should a foreign good be subsidized by $2 by the overseas government, the domestic government may wish to use a $2

Theory of Knowledge

To what extent can we know whether trade protectionist policies improve the economic well-being of low income individuals and households?

Theory of Knowledge

Given that 'fairness' is a matter of opinion, how do economists know what is and what is not fair in terms of international trade and exchange?

tariff to correct the price. Some economists may even go as far to argue that this then creates a situation of free trade.

The difficulty with this approach is determining how much the original subsidy was. This information may not be readily available. The impact of that subsidy on price will also depend on the price elasticity of demand and supply and the prevailing exchange rate. Therefore, it is almost impossible for a government to accurately use tariffs to adjust the price and quantity back to a position of free trade or competitive trade.

It is important to note that should evidence of dumping be obtained, it is possible for that government to register a complaint with the WTO and seek permission to introduce trade protection measures. These measures might then be done collectively by a group of governments, such as applying sanctions to the dumping nation.

Unfair competition

A common argument put forward by governments is 'not everyone is playing by the same rules'. Every economy is unique, and many have its own laws and business regulations, which can be seen as favouring domestic firms. For example, a lower corporate tax rate has a similar effect as a subsidy, in so far as firms will have lower costs of production. Should these firms then export to foreign markets, the impact is quite similar to dumping.

Therefore, a government may believe it is justified in taxing the imports of those overseas firms. However, this argument is not based on economic efficiency, as the foreign producers may have a natural comparative advantage. Domestic firms are only disadvantaged when a product is imported at a lower cost of production. Domestic consumers would tend to benefit from greater competition in international markets, at least in the short run.

Interestingly, it can be argued that the reasons for countries joining a trade bloc or customs union (see Chapter 30) creates further unfair competition for non-member states, as members impose trade protectionist measures of their own.

Balance of payments correction

The **balance of payments** is a financial record of a country's transactions with the rest of the world over the course of a year. This mainly consists of the country's trade in goods and services with other countries. The demand for a currency represents credits to the balance of payments (see Chapter 32), as the only reason that people would buy foreign currencies is to pay for goods or services from their respective economies. By contrast, the supply of a currency represents debits to the country's balance of payments, as the only reason one would sell that currency would be to buy another currency, thereby injecting money into a foreign economy. If the market for a currency is freely floating (see Chapter 31), it will be in equilibrium, which means that the debits and credits in the accounts of the balance of payments will be equal over time. This means that there will be a *balance of payments*.

An imbalance of the balance of payments is the result of a currency that is not freely floating – it is managed or being manipulated by others. If this currency is revalued or overvalued then the quantity supplied of that currency will be greater than the quantity demanded, meaning the debits to the balance of payments will exceed the credits, leading to a balance of payments deficit. Hence, there is an argument for government intervention via protectionist measures to correct this imbalance by reducing the demand for imported goods and services. This will ultimately improve the imbalance on the balance of payments.

The argument against this approach would suggest that the government engaging in trade protection is treating only a symptom of the problem and not the cause. If their currency is being

◆ The **balance of payments** is a financial record of a country's transactions with the rest of the world, usually over one year. It includes the country's trade in goods and services with other countries.

Top tip!

It is important to distinguish between an economy's *current account deficit* and an *imbalance of the balance of payments*. A current account deficit occurs when the total value of imported goods and services exceeds the total value of exported goods and services, per time period. If the economy has a freely floating currency, then this deficit will be met with a surplus in any of the other accounts of the country's balance of payments. This is why economists refer to a *balance of payments*.

manipulated, then they would be better off directing their efforts towards interventions in the currency market rather than in trade in general.

Table 29.1 Countries with highest tariffs in the world

Country	Weighted mean applied tariff
Palau	34.6%
Solomon Islands	30.3%
Bermuda	27.6%
St. Kitts and Nevis	21.1%
The Bahamas	17.1%
Fiji	16.3%
Benin	15.3%
Antigua and Barbuda	14.9%
Guinea Bissau	14.4%
Togo	12.8%

Source: World Bank, 2018 data

Generating government revenue

Governments in need of additional revenue may regard tariffs as an opportunity to gain this source of public sector income. Tariffs are an important source of government revenue for many countries. If a government is considering the use of tariffs for the sole purpose of generating revenue, then it will target goods and services that have a low price elasticity of demand (that is, relatively price inelastic products). This is because the increase in price (caused by the tariff) will lead to a less than proportionate change in quantity demanded, meaning that the government will be able to impose relatively large taxes without much impact on the availability of products to consumers.

For example, the government may choose to target luxury goods for this purpose, as they tend to be consumed by high-income individuals and therefore tend to be less price elastic, although they are generally price elastic for the general population (see Chapter 7). There is also an ethical reason for this, as it would give the government a bad reputation if it began to heavily tax imported goods that are regarded as necessities.

While this approach is likely to generate tax revenue, it will also generate significant inefficiencies in the various markets targeted by the government. Domestic firms will become protected from foreign firms and consumers will experience a loss of consumer surplus. The tariff may also evoke retaliatory trade protection measures from other countries and may lead to sanctions from the World Trade Organization (WTO), which is covered in Chapter 30. This may be a high price to pay in the long run for the pursuit of an increase in government revenue.

Protection of jobs

Whenever domestic consumers buy imported goods, they are creating a derived demand for labour in another economy rather than in their own domestic country. Eventually, this lack of demand for domestic goods and services will lead to job losses in domestic industries. The loss of jobs caused by foreign competition can have significant impacts on the standards of living of households in an economy. Therefore, some governments believe that they are justified in using trade protectionist measures to protect jobs and therefore safeguarding the livelihoods of their citizens.

However, many economists disagree with this argument. While they concede that in the short run these job losses are harmful, in the long run the economy will be more efficient because of international trade. The failure or collapse of domestic businesses due to foreign competition could be seen as the market correcting itself and gaining greater efficiency. In the long run, it would be better if those resources were allocated to more competitive sectors of the economy. For example, the sale of Jaguar (to India's Tata Motors) and British Steel (to China's Jingye) have ensured jobs are protected in the UK as production has continued following a change in ownership of these former British companies.

The economic argument here is based on the difference between an absolute advantage and a comparative advantage (see Chapter 27). Even if an economy fails to obtain a comparative advantage in any of its most consumed products, consumers will still benefit from the lower

WISE ChoICES

Economically developed countries use protectionist measures to safeguard domestic jobs. How might **intervention** result in greater **efficiency** and improvements in **economic well-being** but not necessarily improve **equity?**

prices gained from importation. From an economics perspective, it would be inefficient for a domestic firm to produce a product and incur a greater opportunity cost compared with a foreign producer. This will lead to higher prices for domestic consumers and ultimately a loss of consumer surplus.

Economically least developed country (ELDC) diversification

An **economically least developed country** (ELDC) is a low-income country that faces severe structural barriers to sustainable economic development. ELDCs have low levels of human capital and are highly vulnerable to economic and environmental shocks. The United Nations has identified 47 countries as ELDCs, many of which are highly specialized in the production of a small range of goods and services. Owing to a lack of physical and human capital investment in their economies, these goods tend to be primary sector products such as cocoa, coffee, sugar and rice. Primary sector goods tend to have a low price elasticity of demand and high price volatility, causing financial problems for ELDC economies. Hence, trade protection is used to help ELDCs diversify their production.

◆ **Economically least developed countries** (ELDCs) are low-income countries facing severe structural barriers to sustainable economic development. They have low levels of human capital and are highly vulnerable to economic and environmental shocks.

As a strategy to achieve economic development, an ELDC may wish to diversify its production away from the primary goods it has traditionally specialized in producing, choosing instead to establish its manufacturing (secondary) industry. In other words, trade protection is used to avoid the risks of over-specialization. To achieve economies of scale in the secondary industry, ELDC governments may feel the need to use tariffs or quotas on imported manufactured goods to generate a larger demand for domestically produced goods.

Figure 29.4 Trade protection can help ELDCs to avoid overspecialization of agricultural output

Top tip!

The most common form of trade protection is the use of **tariffs**. All other types of trade protection are collectively known as **non-tariff barriers**.

However, this strategy has considerable weaknesses. The main drawback is that it contradicts the law of comparative advantage, so causes a loss of efficiency in the protected markets. Consumers in ELDCs (who are likely to have low levels of incomes) will have to pay more for goods that are needed for their subsistence. In the short run, this could lead to a lower standard of living for the populations of ELDCs, which will eventually lower economic growth and development in the long run.

CASE STUDY

Openness to trade in sub-Saharan Africa

Between the mid-1990s and 2008, sub-Saharan Africa's openness to trade increased from 16 per cent to 81 per cent as measured by the sum of international trade relative to the size of the region's GDP. However, falling commodity prices and slowing external demand, particularly from China and the euro area (the region's two largest trading partners), caused this ratio to fall to 74 per cent of GDP by 2017. Greater manufacturing sector participation has been constrained by the region's lack of international competitiveness, partly due to relatively low labour productivity and imposition of a range of non-tariff barriers. Hence, openness to trade in sub-Saharan Africa has brought winners and losers.

Figure 29.5 Food and agricultural products are major exports for sub-Saharan Africa

Source: https://www.worldbank.org/en/publication/global-economic-prospects#firstLink51618

PAPER 1 EXAM PRACTICE QUESTION 29.1

a Explain how trade protection supports economically least developed countries (ELDCs). *[10 marks]*

b Using real-world examples, discuss the possible implications of trade protection on the balance of payments. *[15 marks]*

Arguments against trade protection/ disadvantages of trade protection (AO2)

The IB Economics syllabus refers to seven interrelated arguments against trade protection, that is, the disadvantages of trade protection: misallocation of resources, retaliation, increased costs, higher prices, less choice, domestic firms lacking incentives to become more efficient, and reduced export competitiveness. These can be remembered by the acronym **PC CRIME** (Figure 29.6).

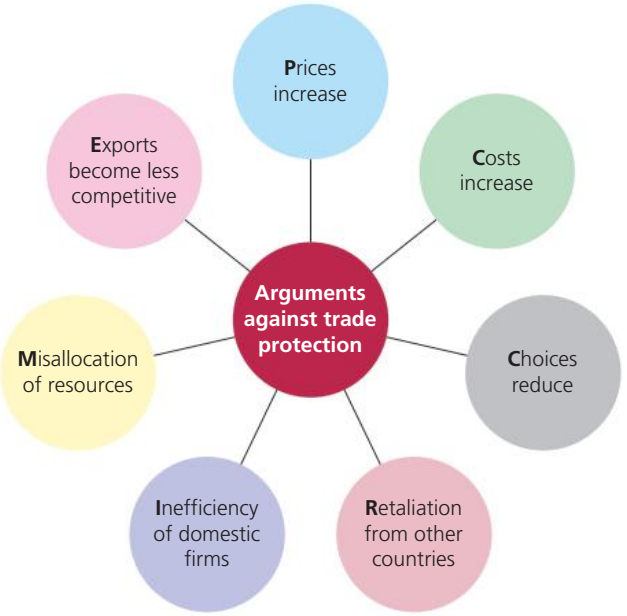


Figure 29.6 The disadvantages of trade protection

Misallocation of resources

When examining the various forms of trade protection, many instances cause a loss in consumer surplus and a gain in domestic producer surplus. These diagrams (see Chapter 28) demonstrate that an introduction of tariffs, quotas and subsidies will result in a net loss in social surplus. This means that the sum of consumer surplus and producer surplus is lower *after* the imposition of trade protection. Ultimately, trade protection distorts market forces and comparative advantages. It can therefore be argued that trade protection creates a less efficient market to the detriment

of the country's economic well-being. This is made worse in the case of trade protection on goods and services that are deemed to be socially desirable, that is, interventions by the government are considered to make markets worse-off.

Retaliation

Governments that engage in trade protection may also find themselves the victim of retaliation. **Retaliation** (or simply **trade retaliation**) means a country imposes trade barriers in response to trade restrictions being imposed on them from other countries. In the worst-case scenario, retaliation can lead to **trade wars**. These are economic conflicts between two or more countries that result from extreme and escalating reciprocal measures of trade protection imposed on each other.

This retaliation could be delivered unilaterally or multilaterally. *Unilateral retaliation* is the result of the domestic economy imposing trade restrictions on products from other countries in response to their existing trade sanctions. *Multilateral retaliation* has the potential to be significantly worse, where a group of economies retaliate at the same time. This could be as a result of them being members of a trading bloc and therefore having common external policies (see Chapter 30). In each of these circumstances, what may have started out as a measure to protect domestic firms in one industry may bring about far worse consequences if firms in a wide range of industries see significant losses in their export revenues.

◆ **Retaliation** (or **trade retaliation**) refers to the actions taken by a country in response to trade restrictions being imposed on it from other countries.

◆ A **trade war** is an economic conflict that results from extreme measures of trade protection between two or more countries, by imposing reciprocal restrictions to international trade.

ATL ACTIVITY 29.2

Investigate the trade war between China and the USA that started in 2018. Prepare a report on the economic costs of this bilateral trade war between the world's two largest national economies. Discuss the winners and losers from this prolonged trade war.

Figure 29.7 The USA and China embarked on a major trade war in 2018

Increased costs

Economies such as Hong Kong and Singapore import a large amount of their factors of production. This means that the importation of raw materials (such as crude oil), capital and even labour are vital to the firms operating in these economies. In this case, the customers of these goods and services are actually firms. If the Singaporean government engages in trade protection measures (such as imposing tariffs and quotas on raw materials and components), there is likely to be an increase in the price of these imported products and therefore an increase in the costs of production for Singaporean producers. This can reduce the profits of domestic firms in the economy and could lead to firms moving their production overseas (a process known as offshoring). This will have detrimental impacts in the long run on the economy's growth, development and jobs.

Figure 29.8 Singapore is one of the most expensive countries in the world to live in

Higher prices

When a government engages in trade protection measures such as tariffs, quotas and administrative barriers, there is a resulting increase in price in the market. Even with producer subsidies, these cost savings just enable domestic producers to operate at a larger level of output, rather than being used to reduce the price for consumers. As a result of trade protection in the form of tariffs, quotas and/or administrative barriers, domestic consumers face higher prices, which leads to a loss of consumer surplus.

The lower level of consumption for domestic consumers can represent a decline in their standard of living as fewer consumers are able to maximize their benefit. In the case of lower income consumers, the impact of tariffs can be described to be regressive (see Chapter 22) because the tax on imported goods represents a higher proportion of their income than it does for higher income consumers. This inequality is compounded by the fact that those on higher incomes might continue to consume the imported product owing to their ability to pay, whereas those of a lower income have to make do without or settle for an inferior substitute product.

Less choice

It was examined in Chapter 28 that each method of trade protection will result in lower revenue for foreign producers. It is rational behaviour for a firm to sell its goods and services in the market where it will receive the highest revenue. This means that trade protection measures can lead to foreign firms exiting the local market altogether. This results in less competition in the market and leaves domestic consumers with less choice or only domestically produced goods to satisfy their needs and wants. The lack of competition may also provide less incentive for domestic firms to be competitive, productive or innovative. This will lead to less choice for domestic consumers, representing a fall in their overall standard of living.

Figure 29.9 Trade increases choice for consumers

Domestic firms lack incentive to become more efficient

When a government introduces trade protection, domestic firms will either enter the protected market (attracted by the higher prices and less foreign competition) or produce more in that market. However, this increase in domestic production is inefficient, because it exists only as the result of domestic government intervention. In the long-term, the lack of competition for domestic firms could lead to inefficiency and lower productivity due to rising average costs of production associated with complacency. This is referred to as **X-inefficiency**, which is the difference between the behaviour of rational firms in competitive markets and those in the absence of competition.

◆ **X-inefficiency** occurs when firms lack the incentive to control costs of production. It represents the difference between the behaviour of rational firms in competitive markets and those in the absence of competition.

Reduced export competitiveness

Export competitiveness refers to the ability of domestic firms to export products successfully to foreign markets owing to their ability to compete in overseas markets. While trade protection may increase the revenue of firms domestically, the subsequent removal of competition from the market can lead to higher costs, less innovation and ultimately higher prices for consumers. When firms in these protected markets become complacent yet try to export their goods and services to foreign markets, they may find that their average costs are higher than foreign firms that operate more efficiently and productively.

◆ **Export competitiveness** refers to the ability of domestic firms to export products successfully to foreign markets owing to their ability to compete in overseas markets.

It should be noted that for most economies, the demand for a firm's products from abroad is significantly larger than the domestic demand for the firm's products. Therefore, gains in domestic demand as the result of trade protection methods may be far outweighed by losses in export revenues from a lack of export competitiveness in overseas markets.

Free trade versus trade protection (AO3)

The validity of the arguments for and against trade protection is very much anchored to the economic and political beliefs of the reader or policymaker. It is important to

consider different stakeholder perspectives when evaluating the relative merits of free trade versus protections. It is also useful for students to consider the short-term and long-term implications of free trade and trade protection.

It is undeniable that trade protection benefits domestic producers rather than domestic consumers of regulated goods and services. In the case of tariffs, quotas and administrative barriers, consumers experience an increase in price and a loss of consumer surplus. They are unable to maximize their benefit and can experience a loss in their standard of living.

The extent to which these protectionist methods will impact a consumer will depend on a variety of factors. Firstly, it is important to establish the specific good or service that is being restricted. A blanket tariff on all imported goods will have a much greater negative impact on domestic consumers and firms than specific and targeted tariffs on socially undesirable goods such as tobacco products, junk food, high-caffeine energy drinks and violent video games. While tariffs and quotas will reduce efficiency in the domestic market, this may be socially desirable if the good or service generates considerable negative externalities. Therefore, each instance of trade protection must be evaluated on its own merits.

Figure 29.10 Regulating the sale of high-caffeine energy drinks can benefit society

Another limitation is that economic analysis sometimes ignores the link between consumers and producers. These stakeholders are often treated separately, but it can be seen from the circular flow of income that households own factors of production, which are employed by firms (see Chapter 1). The benefits of production of domestic producers and the consumption of domestic consumers are therefore linked. Should a domestic firm see an improvement in its income, it can be argued that this may filter down to their employees and therefore consumers in the long run. Therefore, the extent to which trade protection is beneficial will depend on how many people are employed in the protected industry. An evaluation of the extent to which the increase in revenue to domestic firms will trickle down to their employees is also necessary when judging the arguments for or against trade protection.

A factor that needs to be considered is the reliance that all free trade arguments have on the law of comparative advantage (see Chapter 27). The argument for protecting an industry that has a potential comparative advantage contains asymmetrical information and a degree of subjectivity. Nobody can be certain that a targeted domestic industry will grow into a highly efficient and internationally competitive one. Such predictions are based on economic models, which themselves are based on assumptions including the rational behaviour of consumers and firms. As economists, we acknowledge that these are assumptions and that the future may be very different from what is predicted now.

Arguments around unfair competition and anti-dumping as justification for trade protection may also have their limitations. The notion that 'nobody plays by the rules' is not itself an argument that is based on economic efficiency. David Ricardo's law of comparative advantage assumes a position of free trade. This could lead to the argument that if economies are not engaging in free trade then it will be damaging to continue exposing domestic firms to unfair competition. This argument does not dismiss the view that the global economy will be more efficient under free trade.

Theory of Knowledge

If economists remove the *ceteris paribus* assumption from their hypotheses, models and theories, can they really claim to know anything?

The global economy is continually challenged by the eagerness of governments to place their own economies first and the resulting scepticism of trading blocs and other globalized forums. It is not the responsibility of the economist to make value-judgements or political arguments in favour of free trade or protectionist measures. On a national basis, one may find merits in the arguments of trade protection, but until Ricardo's law of comparative advantage is repudiated, economists will continue to believe that free trade will generally lead to improvements in the efficiency of global markets. Perhaps one could go further and say that Ricardo's theory encompassed the philosophy of international mindedness, which is a key feature of the International Baccalaureate Programmes.

Top tip!

While trade protection can provide short-term benefits for a country, most economists believe that it is detrimental to economic growth and development in the long-term. Indeed, such a view is supported by the World Trade Organization.

Theory of Knowledge

In the context of the arguments for and against trade protection and control, is it possible to determine whether one country knows what is right or best for another country when they have markedly different cultures or political systems?

WISE ChoICES

In the context of international trade and protectionism, does **interdependence** or **intervention** have a greater impact on **economic well-being**?

PAPER 2 EXAM PRACTICE QUESTION 29.2

In early 2020, Diageo announced that thousands of jobs could be lost due to an ongoing US–EU trade dispute. The British manufacturer stated the 25 per cent tariff on imports of Scotch Whisky meant that the company's profits were being hit. Diageo also said more than 10,000 people are directly employed in the Scotch whisky sector in Scotland. Another 40,000 jobs in the UK are supported by the Scotch whisky industry, including distillers and distributors. This figure includes around 7,000 jobs in rural areas of Scotland.

Using information from the text/data and your knowledge of economics, discuss the possible economic outcomes for the United States and Scotland from the ongoing trade dispute and introduction of the 25% tariff. *[15 marks]*

Chapter summary

- Trade protection is the use of tariff and non-tariff barriers to trade in order to safeguard an economy from excessive international trade and foreign competition.
- The arguments used to justify trade protection can often be more political than economic.
- The key arguments for trade protection (advantages of trade protection) include:
  - protection of infant (sunrise) industries
  - national security
  - health and safety
  - environmental standards
  - anti-dumping
  - unfair competition
  - balance of payments correction
  - source of government revenue
  - protection of domestic jobs
  - diversification for economically least developed countries (ELDCs).
- Key arguments against trade protection (disadvantages of trade protection) include:
  - misallocation of resources
  - retaliation from trading nations
  - increased costs of production
  - higher prices for domestic consumers
  - less choice for consumers
  - domestic firms lack incentives to become more efficient if protected
  - reduced export competitiveness due to inefficiencies.
- While trade protection methods may increase the revenue of firms domestically, the subsequent reduction or removal of competition from the domestic market can lead to higher costs, less innovation and ultimately higher prices and less choice in the long run. Ultimately, this is likely to limit economic growth and development over the long-term.

30 Economic integration

Real-world issue

Who are the winners and losers of the integration of the world's economies?

Conceptual understandings (WISE ChoICES)

- The increased **interdependence** of economies has benefits and costs.
- Increased economic integration may result in **efficiency**, welfare gains and improvements in **economic well-being** but the benefits may not result in **equity**.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:
- preferential trade agreements, in terms of bilateral, regional and multilateral trade agreements (AO1)
- various forms of trading blocs: free trade areas/agreements, customs unions and common markets (AO2)
- the advantages and disadvantages of trading blocs (AO3)
- trade creation and trade diversion (AO3) (HL only)
- monetary union (AO2)
- the advantages and disadvantages of monetary union (AO3) (HL only)
- the objectives and functions of the World Trade Organization (AO2)
- the factors affecting the influence of the WTO (AO2).

Preferential trade agreements (AO1)

Economic integration refers to the process of countries becoming more interdependent and economically unified. It can be achieved by preferential trade agreements between member governments to remove trade barriers. Methods of economic integration include trade agreements (preferential, bilateral, regional and multilateral), trading blocs (such as free trade areas, customs unions and common markets) and monetary union. Economic integration intensifies competition for producers within a trading bloc, which should lead to greater efficiency. It should also offer consumers access to a broader range of goods and services at more competitive prices, and of higher quality. By contrast, trade protectionism (see Chapter 28) is generally detrimental to consumers in the long run because the prices of the imported products are relatively higher and choice is limited.

◆ **Economic integration** is the process of countries becoming more interdependent and economically unified.

A **preferential trade agreement** (PTA) is a form of economic integration that involves member states or governments in a trade treaty (or trade deal) consenting to special or favourable terms and conditions of trade. PTAs give certain countries special and easier access to specific products in a market owing to advantages such as the reduction or removal of tariffs and non-tariff barriers to international trade. They comprise bilateral trade agreements, regional trade agreements and multilateral trade agreements through the World Trade Organization (WTO) in order to promote freer and fairer terms and conditions of trade.

◆ A **preferential trade agreement** (PTA) is a trade treaty between two or more countries, giving special or favourable terms and conditions of trade to member countries.

Bilateral trade agreements

A **bilateral trade agreement** is the simplest form of preferential trade agreement as it involves a legally binding trade contract between two countries. An example is the closer economic partnership arrangement (CEPA) between China and Hong Kong, signed in 2004 to create improved cross-border trade and investment between the two member states. The main purposes of a bilateral trade agreement are to increase economic and political co-operation between the two member states and to facilitate international trade between the two countries, usually by mutual agreement to reduce or remove tariffs and non-tariff barriers on products traded between the two member states. Bilateral trade agreements have far greater flexibility than multilateral trade agreements because they involve only two countries, so they are far simpler.

Note that unlike traditional trade deals between countries, bilateral trade agreements are *not* bound by the rules of the World Trade Organization (WTO) especially as they do not always focus solely on trade-related issues.

◆ A **bilateral trade agreement** is a preferential trade agreement between two countries, usually by mutual agreement to reduce or remove barriers to trade.

Figure 30.1 Hong Kong and China have numerous bilateral trade agreements

Regional trade agreements

According to the World Bank, a regional trade agreement (RTA) is a reciprocal trade agreement between two or more countries, usually but not always belonging to the same geographical region. Examples of RTAs include:

- the 27 members of the European Union (EU)
- the 21 member economies of Asia-Pacific Economic Co-operation (APEC)
- the five member countries of Mercosur (Argentina, Brazil, Paraguay, Uruguay and Venezuela)
- the Central American-Dominican Republic Free Trade Agreement (CAFTA-DR), which is the first free trade agreement between the USA and a group of Central American countries (Costa Rica, El Salvador, Guatemala, Honduras and Nicaragua), plus the Dominican Republic.

ATL ACTIVITY 30.1

Investigate the economic reasons why a country of your choice chooses to establish a bilateral trade agreement. Discuss the winners and losers of this agreement.

Regional trade agreements account for the vast majority of world trade, with more than 300 RTAs in place. Word Bank data show that the EU has the world's largest number of regional trade agreements, although Chile, Singapore, Mexico and Turkey are the largest individual users of RTAs.

RTAs involve member countries that formally agree to the free movement of goods and services across the borders of their members. As with all forms of preferential trade agreements, RTAs encourage the reduction or removal of barriers to trade. RTA deals come with internal rules that member states follow among themselves. This means there are external rules in place when dealing with non-member states.

Multilateral (the World Trade Organization) trade agreements

A **multilateral trade agreement** is a legally binding preferential trade agreement between more than two countries and/or trade blocs. These preferential trade agreements are made within the guidelines of the World Trade Organization (WTO), with the agreement and intention to reduce or remove international trade barriers between member countries. Many regional trade agreements are also classified as multilateral trade agreements, as long as they fall under the rules and regulations of the WTO. They are numerous in scale and scope, such as the East African Community (EAC) comprising six countries: Burundi, Kenya, Rwanda, South Sudan, Tanzania and Uganda.

◆ A **multilateral trade agreement** is a legally binding preferential trade agreement between more than two countries and/or trade blocs, under the guidelines of the WTO.

Note that bilateral trade agreements have far greater flexibility than multilateral trade agreements because the latter involves far more countries, possibly in different regions, so they are far more complex and trade negotiations can be challenging. Nevertheless, the multilateral trade agreements reduce tariffs and non-tariff barriers for member states, thereby making it easier for businesses to import and export.

Trading blocs (AO2)

A **trading bloc** is a group of countries that agree to economic integration and freer international trade by reducing or removing trade barriers with one other. The group will, however, impose barriers to international trade for non-member states, such as the imposition of tariffs, to increase the price of imported products (see Chapter 28). Economic integration within the trading bloc will also intensify the degree of competition for producers within the member countries. However, firms can gain from access to larger markets without trade barriers, thereby benefit from economies of scale. Trading blocs, with their varying degrees of economic integration, can be categorized as free trade areas (or free trade agreements), customs unions and common markets.

◆ A **trading bloc** is a group of countries that agree to economic integration and freer international trade by reducing or removing trade barriers with one other.

Free trade areas/agreements

A **free trade area** (FTA) or **free trade agreement** is the least economically integrated type of trading bloc, where member countries agree to remove trade barriers with one another, but impose *separate* trade barriers with non-member countries. Examples of FTAs include:

- The United States-Mexico-Canada Agreement (USMCA), which replaced the North American Free Trade Agreement (NAFTA) in March 2020. NAFTA was established in January 1994 with the agreement for the region to become tariff-free. USMCA is one of the world's largest trade blocs as measured by gross domestic product.
- The European Free Trade Association (EFTA), a regional trade bloc between four European member states: Iceland, Liechtenstein, Norway and Switzerland.

◆ A **free trade area** (FTA) is a type of trading bloc between member states that agree to trade freely with each other but can impose separate trade restrictions with non-member countries.

- The South Asian Free Trade Area (SAFTA), established in January 2004, between Afghanistan, Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan and Sri Lanka to reduce customs duties of all traded goods between the member states. The SAFTA is one of the largest FTA as measured by population size, accounting for more than 1.6 billion people.

Customs unions

A **customs union** is a type of trading bloc consisting of member countries that agree to reduce and remove tariff and non-tariff barriers to trade, thereby ensuring the free flow of products between the member states of the union. The purpose of a customs union is to make it easier for member countries to trade freely with each other by promoting economic co-operation and reducing administrative and financial burdens associated with international trade. While member states engage in free trade with each other, the union imposes a **common external tariff** for all non-member countries. This means that all members of the customs union impose the *same* (common) trade barriers to non-member countries.

A customs union differs from a free trade area, where each member nation maintains the political and economic freedom to determine its own trade barriers with countries outside of the FTA. The countries in a customs union negotiate as a united trading bloc when discussing and negotiating trade deals with non-member states. Revenues from import tariffs are combined for all member states of the customs union. The European Union (EU) is by far the world's largest customs union in terms of the value of national output of its 27 member countries. Another example is the Southern African Customs Union (SACU), established in 1910 and comprises five nations: Botswana, Eswatini (formerly Swaziland), Lesotho, Namibia and South Africa.

◆ A **customs union** consists of member countries in a trading bloc that engage in free trade with each other but impose a common external tariff when trading with non-member states.

◆ A **common external tariff** means that all members of a customs union impose the same (common) trade barriers to non-member countries.

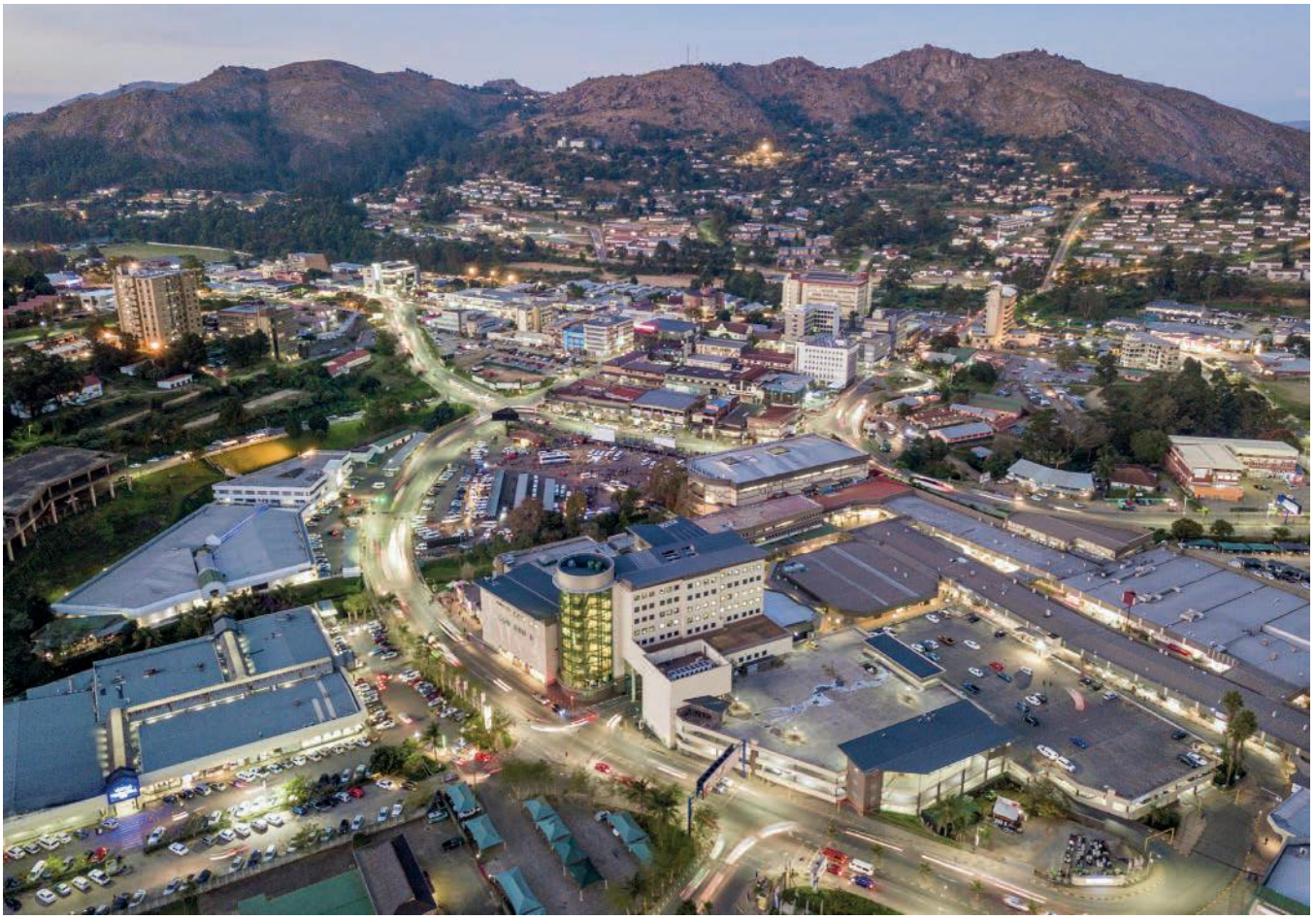


Figure 30.2 The Kingdom of Eswatini is a member state of the Southern African Customs Union

ATL ACTIVITY 30.2

The UK officially left the European Union in January 2020, having been part of the world's largest customs union since 1973. Investigate the reasons for and against the decision to 'Brexit' (Britain exiting from the EU). As an economics student, do you think the long-term benefits of Brexit outweigh the long-term costs of doing so? Justify your reasons.

Figure 30.3 The UK was a member of the EU from 1973 to 2020

Top tip!

Make sure you can distinguish between a **free trade area** (or free trade agreement) and a customs union. In an FTA, each member state continues to be able to set its own distinct external tariffs on products imported from the rest of the world. For a **customs union**, there are common external tariffs imposed on products imported from non-member states.

◆ A **common market** is the most integrated trading bloc, consisting of a customs union that allows the free movement of factors of production between member countries.

PAPER 1 EXAM PRACTICE QUESTION 30.1

Explain the differences between a free trade area (FTA) and a customs union. *[10 marks]*

Common markets

A **common market** (or **single market**) is a third type of trading bloc, similar to a customs union (in that goods and services are traded without tariffs and there is a CET) but, in addition, there is factor mobility within the bloc. This means that the four factors of production (land, labour, capital and enterprise) are permitted to flow freely between member states. This means there is a formal agreement that allows member countries to move resources among themselves freely and helps to improve the allocation of resources within and between member states. This makes common markets the most integrated of the three categories of trading bloc.

For example, Volkswagen, Germany's and the world's largest carmaker by sales volume, has production plants and facilities in several EU countries including Spain and Italy. There are no restrictions placed on Volkswagen on the flow of natural, physical and human capital resources and entrepreneurial talent.

The world's largest common market is the European Economic Area (EEA), which consists of the 27 members of the European Union (EU) plus three of the four members states from the European Free

Common mistake

Owing to the vast number of key terms in economics, and the overlapping nature of economic integration, students often confuse the key terms covered in this unit, such as *preferential trade agreements* (PTAs), *free trade areas* (FTAs), *customs unions*, *common markets* and *monetary union*. Therefore, make sure you review this chapter and ensure you can distinguish between these five key terms.

Trade Association (EFTA): Iceland, Liechtenstein and Norway. Essentially, the EEA grants extension to these three non-EU members that are generally governed by the same basic rules on trade and the free movement of labour, goods, services and capital within the common market. However, the EFTA members do not partake in all aspects of EU policies, especially regarding agriculture and fisheries.

Advantages of trading blocs (AO3)

The advantages and disadvantages of trading blocs will depend on the degree of economic integration, that is, whether the country is part of a free trade area, customs union, common market or monetary union. The generic advantages and disadvantages of being a member of trading blocs are listed below.

Advantages of membership of a trading bloc include: trade creation (HL only), economies of scale from having greater access to markets, greater employment opportunities, stronger bargaining power in multilateral negotiations, and greater political stability and co-operation.

Trade creation (HL only)

Trade creation occurs when economic integration causes trade deals to move away from higher-cost producers from outside the trading bloc to lower-cost producers within the trading bloc, due to the removal of trade barriers. Recall that members of a customs union agree to eliminate tariff and non-tariff barriers between themselves, but impose a common external tariff on imports from non-member states. The effect of this is that consumers increase their demand for these lower priced goods from within the customs unions, creating new trade for members of the trading bloc.

◆ **Trade creation** occurs when trade shifts from higher-cost producers outside of a trading bloc to lower-cost producers within the bloc, due to the removal of trade barriers.

Access to larger markets and the potential for economies of scale

Greater access to markets in a trading bloc offers member states the potential to gain from economies of scale. Membership of a trading bloc enables multinational corporations (MNCs) to locate in overseas member state countries and to export their products without trade restrictions. Thus, MNCs can gain easier access to larger markets and operate on a much larger scale, thereby allowing them to benefit from potential economies of scale (see Chapter 14). Consumers also benefit from this because of the increased choice, lower prices and higher quality under the free trade agreements of trading blocs.

Greater employment opportunities

In general, a greater degree of economic integration facilitates economic growth and therefore creates more job opportunities. Furthermore, with greater freedom of movement of labour within a common market, there are greater employment opportunities in the trading bloc. Most economists believe that trade protection is harmful to households and firms in the long-term. This is because greater economic integration leads to dynamic efficiency gains from freer and fairer international trade, thereby lowering unemployment in member countries.

Stronger bargaining power in multilateral negotiations

Membership of a trading bloc may allow for stronger bargaining power in multilateral negotiations. In particular, low-income countries can benefit from the expertise and support of larger, more economically empowered nations during multilateral trade negotiations. For example, the Dominican Republic-Central America Free Trade Agreement (CAFTA-DR) is a free trade agreement between Costa Rica, El Salvador, Guatemala, Honduras and Nicaragua, giving these members greater bargaining power in multilateral trade negotiations with the USA.

Greater political stability and co-operation

Through greater trade and employment opportunities arising from membership of trading blocs, countries are able to benefit from higher levels of output, growth and development. This interdependence between member states encourages countries to collaborate and co-operate in a harmonious way to improve their respective economies. Hence, membership of a trading bloc helps to promote greater political stability and co-operation.

Disadvantages of trading blocs (AO3)

The disadvantages of being a member of a trading bloc include trade diversion (HL only), the loss of sovereignty and challenges to multilateral trading negotiations.

Trade diversion (HL only)

Trade diversion occurs when economic integration through trading blocs shifts trade deals away from lower-cost producers outside the bloc to higher-cost producers within the bloc. Hence, trade diversion is considered undesirable and inefficient as it concentrates production in member countries, namely those in a customs union, with a higher opportunity cost and lower comparative advantage. Therefore, consumers tend to lose out from trade diversion as they pay higher prices from less efficient countries.

Another linked drawback of joining a trading bloc is therefore that trade diversion can lead to more unemployment, at least in the short-term. This is due to the increased intensity of competition facing domestic firms in the trading bloc.

◆ **Trade diversion** occurs when trade shifts from lower-cost producers outside of a trading bloc to higher-cost producers within the bloc, under the terms and conditions of the trade bloc agreements.

Top tip! (HL only)

HL students need to be able to distinguish between *trade diversion* and *trade creation*. Remember that a disadvantage of joining a trading bloc is trade diversion, whereas a key benefit of joining is trade creation with other member countries.

ATL ACTIVITY 30.3

'The formation of a trading bloc always results in greater political stability and co-operation.' Using appropriate real-world examples debate this claim with a partner.

Loss of sovereignty

Some economists argue that economic integration, especially through the establishment of common markets, causes a loss of national economic sovereignty, that is, the lost opportunity to enjoy economic independence. For instance, trading blocs might impose stricter environmental legislation and labour laws. Under monetary union, the common or single central bank sets interest rates across all member states. However, this means national governments lose the ability to use monetary policy (see Chapter 23) to directly influence the level of aggregate demand in their respective countries. At the same time however, detrimental changes in the economic conditions of any member country (such as inflationary pressures or an economic recession) are also likely to affect the economic performance and prosperity of all member countries in the trade bloc without any individual government having the autonomy or sovereign right to adjust its own macroeconomic policies accordingly.

Challenges to multilateral trading negotiations

Multilateral trade agreements can be very complex and rather inflexible especially when there are many countries involved, possibly in different regions. In general, the larger the number of countries involved in a multilateral trade negotiation, the more challenges are present. Differences in national or regional cultures, political systems and time zones can make trade negotiations far more complicated than with bilateral trade deals. Hence, there can be economic and political challenges to multilateral trading negotiations.

Figure 30.4 Differences between countries can create challenges for multilateral trade negotiations

Top tip! (HL only)

The formation of a customs union as a category of trading blocs is likely to cause both trade creation and trade diversion. Economic integration improves the welfare of the economy if it leads to more trade creation than trade diversion.

Theory of Knowledge

To what extent does possession of economic power by a country carry with it an ethical responsibility to others?

WISE ChoICES

To what extent can membership of trading blocs improve **equity** and **economic well-being**?

Monetary union (AO2)

Monetary union refers to a highly economically integrated trading bloc in a common market that requires the convergence of monetary policy, such as interest rate policies. To achieve monetary union, members of a common market must first agree to permanently fix their exchange rates, thereby effectively choosing to use a common currency that is governed by a common central bank. A well-known example of monetary union is the 19 member states (collectively called the 'eurozone') of the European Union that use the euro as their common currency, which is overseen by the European Central Bank (ECB). Monetary union requires convergence of interest rates within the single market, so member states do not have flexibility in exercising their own monetary policy. The ECB is responsible for managing the EU's monetary policy. For full monetary union, a single (common) currency is used by all member countries. This is not the case within the European Union member states.

◆ **Monetary union** refers to the monetary system in a common market that requires the convergence of monetary policy that is governed by a common central bank.

Top tip!

Like much of economics, there is a lot of jargon and terminology to learn in this topic. Make sure you can define and distinguish between the key terms for this unit. For example, note that all free trade areas, customs unions, common markets and monetary unions are also trading blocs. See Table 30.1 for further clarification.

Table 30.1 Summary of the different stages of economic integration

Category	No internal trade barriers	Common external tariff	Factor mobility	Common currency
Free trade area	✓			
Customs union	✓	✓		
Common market	✓	✓	✓	
Monetary union	✓	✓	✓	✓

Top tip!

Make sure you can use real-world examples in the external examinations. This means you should be able to give examples of the various categories of trade blocs, such as:

- Free trade areas (FTAs) such as the United States-Mexico-Canada Agreement (USMCA), a trilateral free trade agreement between the USA, Canada and Mexico to promote trade and investment.
- Customs unions such as the European Union (EU), or the five member countries of the Southern African Customs Union (SACU) – Botswana, Eswatini, Lesotho, Namibia and South Africa.
- Common markets such as the Caribbean Community (CARICOM), which promotes economic integration and co-operation among its 15 full member countries and five associate members.

Also, make sure you can use and integrate these examples to support your arguments about the merits and limitations of economic integration.

PAPER 1 EXAM PRACTICE QUESTION 30.2

Explain the differences between a common market and a monetary union. *[10 marks]*

Advantages of monetary union (AO3) (HL only)

The advantages of monetary union include the following arguments:

- **Exchange rate certainty** – There is exchange rate certainty in a monetary union as a common currency is used. This eliminates the risks associated with international trade due to exchange rate fluctuations and uncertainties. Hence, the stability of using a single currency helps to facilitate more trade between countries within the trading bloc.
- **Increased cross-border investments** – Similarly, there is likely to be more foreign direct investment between members of the monetary union. The use of a common currency is also likely to attract inward investments from non-member countries due to the removed risks associated with exchange rate fluctuations. Inward investment in the monetary union will have a positive impact on economic growth and employment.
- **Increased trade** – A greater degree of trade will occur between members of the monetary union due to the preferential trade agreements and the confidence in the use of a common

currency. Hence, membership of a monetary union should be good for economic growth and job creation in the long-term.

- **Lower transactions costs** – The use of a common currency eliminates transactions costs between trading nations because there is no need to exchange foreign currencies between members of the monetary union.
- **Price transparency** – There is greater price transparency as a single currency in the monetary union makes price comparisons far easier. Households, firms and tourists can easily compare the prices of goods and services across different member states with currency conversions or dealing with fluctuations in the exchange rate.

Disadvantages of monetary union (AO3) (HL only)

The disadvantages of monetary union to member countries include the following arguments:

- **Loss of economic sovereignty** – Being part of a monetary union means giving up the freedom and flexibility to adjust macroeconomic policies because member countries are bound by agreements with other member states. This means they lose the autonomy to deal with their own specific economic problems. For example, a country with relatively high unemployment cannot reduce interest rates to stimulate its economy as this is governed by the monetary union's central bank.
- **Loss of exchange rate flexibility** – Members of a monetary union are also deprived of exercising their own exchange rate policy. With independence, outside of a monetary union, the country can depreciate its currency during a recession or to combat a current account deficit. This autonomy is not possible in a monetary union, so the country cannot adjust its own exchange rate to gain a competitive advantage.

Figure 30.5 The European Central Bank is located in Frankfurt, Germany

- **Asymmetric impacts on member states** – Similarly, the actions taken by the common central bank in a monetary union will have an asymmetric impact on different countries owing to their varying circumstances. For example, contrasting rates of inflation and unemployment in Greece and Germany mean that a common monetary policy as determined by the European Central Bank might not work for either country.
- **Changeover costs** – There are setup and conversion costs involved in establishing a monetary union. For example, when members of the EU's eurozone permanently converted to using the euro as the single currency in the monetary union from 1 January 2002, there were costs associated with printing the new currency, getting rid of the old currencies, updating price lists and menus, adjusting software for bank accounts and stock market investments, converting vending machines and car parking meters, and so forth. The costs of convergence were part of the reason why some countries such as Denmark, Hungary and Sweden were not able or willing to join the eurozone. There are also setup costs and running costs associated with establishing the European Central Bank.

Theory of Knowledge

Do economists have a moral responsibility towards the losers in a free trade agreement or in a monetary union?

ATL ACTIVITY 30.4

Investigate the economic costs for a country in the European Union if it joins the eurozone in order to adopt the euro as its official currency. To what extent do the benefits of monetary union outweigh the drawbacks?

PAPER 1 EXAM PRACTICE QUESTION 30.3

Explain the advantages and disadvantages of being a member country of a monetary union. *[10 marks]*

The World Trade Organization (WTO) (AO2)

The **World Trade Organization** (WTO) was established in 1995 by 123 member countries to promote trade liberalization, to oversee multilateral trade agreements (free trade agreements between multiple countries) and to resolve trade disputes between member states. Everything the WTO does is through negotiations to promote freer and fairer international trade, and dealing with the global rules of trade between nations. The WTO originated from trade negotiations that followed the Second World War when the General Agreement on Tariffs and Trade (GATT) was established in order to rebuild economies and to promote economic co-operation and political harmony. The GATT was eventually replaced by the WTO. Today, it is the largest international economic organization in the world and accounts for around 98 per cent of world trade. The WTO consists of 164 member states and employs more than 640 full-time staff, including lawyers, economists, statisticians and communications experts. There are more than 20 countries seeking to join the WTO.

◆ The **World Trade Organization** (WTO) is a global organization that exists to promote trade liberalization, to oversee multilateral trade agreements and to resolve trade disputes between member states.

All signatory countries of the WTO are required to grant 'most favoured nation' status to other members. This means that any trade concessions granted by a member state to another member country must also be granted to all other WTO members. This is in line with the WTO's principle of promoting freer and fairer trade. Essentially then, the WTO operates as a system for international trade rules, which member countries must abide by.

Note that in some special circumstances the WTO will support maintaining trade barriers rather than promoting free trade. It provides exceptions for environmental protection, national security, and other important goals such as to protect consumers from the spread of diseases, such as prevention of trade of poultry and pork products during outbreaks of bird flu and swine flu.

The objectives of the WTO

There are six interrelated **objectives** of the WTO, as stated on its website, which serve as the core principles of an effective multilateral trading system:

1 **Non-discrimination**

Non-discrimination refers to the removal of discriminatory treatment in trade relations between member nations. It requires all members of the WTO to refrain from discriminating between the goods and services or nationals of their trading partners (except for those in trading blocs that have their own special set of agreed trade rules).

◆ **Non-discrimination** means a WTO country cannot discriminate between its trading partners in terms of their goods, services or people.

2 **More open**

A second objective of the WTO is to encourage free international trade by being more open as an economy. Member states of the WTO are obliged to reduce and remove artificial trade barriers, such as subsidies (used to lower the production costs of domestic firms, thereby gaining an unfair advantage). Other barriers to trade (see Chapter 28) include customs duties (or tariffs), quotas (that restrict the quantity of imports) and embargoes (import bans on selected products). By being more open, this helps producers of goods and services (both exporters and importers) to conduct their businesses in a more open, fair and transparent manner.

3 **Predictable and transparent**

By being more open and non-discriminant, trade becomes more predictable and transparent. This helps to ensure that multinational companies, investors and governments become more confident in international trade as trade barriers should not be used arbitrarily or indiscriminately. By promoting and encouraging stability and predictability, trade and investments are encouraged. Hence, predictability and transparency help to facilitate economic growth and development, thereby creating jobs in the long-term.

4 **More competitive**

The WTO discourages unfair trading practices and protectionist measures, such as export subsidies or dumping of products at below cost value in order to gain market share. However, the issues are complex, as the concepts of fairness or unfairness are somewhat subjective and open to interpretation and debate. Nevertheless, by promoting freer and fairer trade, more competition is created, enabling consumers around the world to enjoy the benefits of greater choice and lower prices. Essentially, more competition from free trade cuts the cost of living for all.

Figure 30.6 The entrance to the WTO, in Geneva, Switzerland

5 **More beneficial for economically less developed countries**

The WTO aims to provide trade opportunities for economically developing countries (low-income nations). This means that WTO member countries agree to engage in fairer and freer trade in order to enhance the economic growth and development prospects of low-income countries. WTO rules are also more relaxed to give low-income countries more time to adjust to the rules, regulations and requirements of WTO membership. This is an important objective, given that more than three-quarters of WTO members are developing countries or countries in transition to becoming market economies.

6 **Protect the environment**

While trade can increase economic prosperity for most people, the final objective of the WTO is environmental protection. WTO agreements allow member countries to take necessary measures to protect their natural environments and ecosystems, including aspects of public health, animal health and plant health. However, these protectionist measures must be applied in accordance with the principle of non-discrimination. So, any restrictions or sanction imposed on foreign businesses must also apply to domestic firms. This means that member states are not allowed to use environmental protection measures as a hidden form of trade protectionist policy.

The functions of the WTO

The WTO states that its overall purpose is to ensure that the trade between countries flows as smoothly, predictably and freely as possible. To this effect, it has five main **functions** as stated on its website:

1 **Trade negotiations**

The WTO is run by its member governments, with all major decisions made by the membership as a whole during 'Rounds' of trade negotiations. Hence, the WTO provides an international platform for member countries to negotiate trade deals, rules and agreements, which become legally binding contracts for governments. This helps to keep trade policies within the agreed parameters set by the WTO based on the principles of liberalization and fairer trade practices. It involves members committing to lower tariffs and non-tariff barriers to trade, and to open up their services markets. These agreements are not static but evolve over time. WTO trade negotiations cover goods, services and intellectual property.

2 **Implementation and monitoring**

A second function of the WTO is administering WTO trade agreements between member countries. Membership of the WTO requires the agreement of governments to commit to making their trade policies transparent. Officials in Geneva then monitor these policies to ensure that they are implemented in line with the agreements and requirements of the WTO. All WTO member states go through this periodical process, with the WTO scrutinizing their trade policies and practices. This function serves to encourage accountability and transparency on trade policies and practices on a multilateral level.

Figure 30.7 Dispute settlements between member states is a key function of the WTO

3 Dispute settlement

Handling trade disputes is a core function of the WTO as dispute settlement is vital for enforcing the rules and regulations of its membership. This helps to ensure that international trade flows smoothly and efficiently. Governments can raise their concerns with the WTO if they believe their rights under WTO membership have been infringed. The WTO acts as an independent arbitrator in trade-related disputes between member states, using its systematic trade disputes resolution process.

4 Building trade capacity

The WTO grants special provisions for developing countries to facilitate their growth and development. This includes allowing these member countries longer periods of time to implement trade agreements and commitments. It also includes providing trade-related technical assistance and training for low-income countries (such as strategies to handle trade disputes and to implement and meet technical standards) to help them to better understand WTO obligations and agreements and to build their capacity to trade. For example, the WTO's *Aid for Trade* training programme provides training for government officials of low-income countries to develop the skills and infrastructure needed to build and expand trade capacity.

5 Outreach

Outreach activities function to help increase global awareness of the WTO's objectives and activities. The WTO works and co-operates with non-governmental organizations and charitable foundations, other international organizations (such as the World Bank, International Monetary Fund and the OECD – the Organization for Economic Co-operation and Development), and the media in order to reach out to more stakeholders in pursuit of aspects of the WTO's goals. By collaborating with other international organizations and parties, the WTO is able to promote economic co-operation and growth between countries (be they WTO members or not).

Factors affecting the influence of the WTO

Membership of the WTO can generate many economic and political benefits through its trade liberalization policies, but many argue that the WTO has had limited influence and success in certain areas of trade. Since the WTO is the only global organization devoted to the enlargement of international trade, it can be challenging to determine the effectiveness of its influence on trade as there is no basis for comparison. As the only global organization that deals with the rules of

trade between nations, it can be difficult to measure the effectiveness of the WTO's policies that trade flows as smoothly, predictably and freely as possible. However, it is possible to consider the factors that affect the influence of the WTO, which include difficulties of reaching an agreement and the unequal bargaining power of members.

Difficulties of reaching an agreement on services/primary products

Despite the agreements with the WTO, high-income members such as the USA and the EU continue to use agricultural subsidies for their domestic producers and impose tariffs on primary products from low-income nations. This limits market accessibility for economically developing countries as their export competitiveness declines. The Doha Round of trade talks, which started in 2001, collapsed largely because rich member countries continued to protect themselves from primary goods produced by the poorer nations. For example, in 2005, average agricultural tariffs imposed by the USA and EU were 60 per cent, whereas tariffs imposed on manufactured products from these regions averaged only 5 per cent. The use of high tariffs on agricultural imports clearly harms farmers and agricultural sector workers in low-income economies. At the same time, these economies are unable to compete in the manufacturing or services industries. Hence, despite the intentions of the WTO, free trade is not conducted equally or fairly across the world.

Unequal bargaining power of members

Although the vast majority of WTO member states are of low-income status, the reality is that the countries with the largest real GDP tend to have much stronger bargaining power. Hence, trade deals tend to be unfair for low-income countries. For example, they are unable to use protectionist measures such as tariffs and quotas to support their infant industries against the market power of large multinational companies from high-income nations such as the USA and the EU nations. Hence, critics argue that the WTO is undemocratic in favour of large multinational companies and the most economically powerful countries.

Although the WTO rules state that members must adhere to the principle of *non-discrimination* (meaning a country cannot discriminate against other WTO member countries by imposing higher trade barriers on one country while reducing the trade barriers on imports from another country), there are exceptions. This also limits the influence of the WTO in trade negotiations and disputes. The two exceptions to the principle of non-discrimination are bilateral trade agreements and regional trading blocs (RTBs) such as the European Union. In both cases, non-member countries can have higher trade barriers imposed.

Top tip!

The two exemptions to the principle of non-discrimination under the WTO regulations are *bilateral trade agreements* and *regional trading blocs* (RTBs) such as the European Union. In both cases, non-member countries can have higher trade barriers imposed.

Theory of Knowledge

What evidence or forms of knowledge could be used to argue that greater levels of economic integration are undesirable?

WISE ChoICES

When does economic and political **interdependence** not improve **choice** for households and firms in a trading bloc?

WISE ChoICES

Economic integration requires greater levels of **integration** and **interdependence** of the world's economies. Using real-world examples, discuss whether economic integration improves the **economic well-being** of individuals and societies.

Chapter summary

- Economic integration is the process of countries becoming more interdependent and economically unified. Forms of economic integration include:
  - preferential bilateral, regional and multilateral trade agreements
  - trading blocs, such as free trade areas, customs unions and common markets
  - monetary union.
- While a trading bloc promotes free trade between its member countries, the group imposes barriers to trade for non-member states.
- A customs union is a form of trading bloc in which member countries agree to remove trade barriers among themselves and adopt common external trade barriers.
- A monetary union exists when member states of a common market adopt a single currency and, hence, a common central bank that oversees monetary policy.
- Advantages of being a member of a trading bloc include trade creation (HL only), greater access to markets/potential for economies of scale, greater employment opportunities, stronger bargaining power in multilateral negotiations, and greater political stability and co-operation.
- Disadvantages of being a member of a trading bloc include trade diversion (HL only), the loss of sovereignty, and challenges to multilateral trading negotiations.
- The formation of a customs union is likely to cause both trade creation and trade diversion. Economic integration improves the welfare of the economy if it leads to more trade creation than trade diversion. (HL only)
- Trade creation occurs when economic integration causes trade deals away from higher-cost producers from outside the trading bloc to lower-cost producers within the trading bloc, due to the removal of trade barriers. (HL only)
- Trade diversion occurs when economic integration shifts trade deals away from lower-cost producers outside the trading bloc to higher-cost producers within the trading bloc, due to trade agreements of the customs union. (HL only)
- Monetary union is an extreme form of trading bloc involving member countries converging their currencies, establishing a common central bank and using a common monetary policy.
- The advantages of monetary union include exchange rate certainty, increased cross-border investments, increased trade, lower transactions costs and price transparency. (HL only)
- The disadvantages of monetary union to member countries include a loss of economic sovereignty, loss of exchange rate flexibility, asymmetric impacts on member states and changeover (or conversion) costs. (HL only)
- Trade-related disputes between member states of the WTO are resolved by independent judges using the WTO's trade disputes settlement process.
- The objectives of the WTO are non-discrimination, more open, predictable and transparent, more competitive, more beneficial for less developed countries and protection of the environment.
- The functions of the WTO are trade negotiations, implementation and monitoring, dispute settlement, building trade capacity and outreach.
- Factors that affect the influence of the WTO include difficulties in reaching an agreement and the unequal bargaining power of members states.

31 Exchange rates

Real-world issue

Who are the winners and losers of the integration of the world's economies?

Conceptual understandings (WISE ChoICES)

- The increased **interdependence** of economies has benefits and costs.
- Increased economic integration may result in **efficiency**, welfare gains and improvements in **economic well-being** but the benefits may not result in **equity**.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- floating exchange rates, including the depreciation and appreciation of a currency (AO2)
- an exchange rate diagram to show the determination and changes in equilibrium in a floating exchange rate system (AO4)
- the price of a good in different currencies by using exchange rates (AO4)
- the factors that cause a change in the demand for and supply of a currency (AO2)
- changes in the value of a currency from a set of data (AO4)
- the consequences of changes in the exchange rate on key economic indicators (AO3)
- how an AD–AS diagram shows potential consequences of changes in the exchange rate on the economy (AO4)
- fixed exchange rates, including devaluation and revaluation of a currency, and how fixed exchange rates are maintained (AO2)
- an exchange rate diagram to show how a fixed exchange rate is maintained (AO4)
- managed exchange rates, including overvalued currencies and undervalued currencies (AO2)
- an exchange rate diagram to show exchange rate determination in a managed exchange rate system and changes in the equilibrium exchange rate (AO4)
- the merits and drawbacks of fixed versus floating exchange rate systems (AO3) (HL only).

Floating exchange rates (AO2, AO4)

An **exchange rate** refers to the value of one currency in terms of another. All economies use money as a means of exchange for goods and services within their borders. The money that is used within an economy can be referred to as legal tender or currency. Owing to different levels of scarcity, currencies of different economies have different values, therefore they are not exchanged at parity (that is, a one-for-one basis). The rate at which they are exchanged is referred to as that currency's exchange rate. It is important to note that the value of a currency is always measured in terms of another currency or basket of other currencies.

◆ **Exchange rates** are the value of one currency expressed in terms of another currency.

■ **Figure 31.1** Exchange rates are used to express the value of foreign currencies

It is important to note that the value of a currency is always measured in terms of another currency. One dollar will always be worth one dollar within the domestic economy, but it will command different values in relation to other currencies. This chapter examines the factors that determine the value of a currency and the different exchange rate systems that can be found in economies across the globe.

A **floating exchange rate** is where the value of a currency is determined by the demand for and supply of the currency in the foreign exchange market. In a floating exchange rate system, there should be no deliberate attempts by a government to influence the value of its currency. Naturally, a government will influence the demand and supply of a currency when it purchases other currencies for administrative reasons as part of its own expenditure, but the purpose of these purchases must not be to influence the value of its own currency.

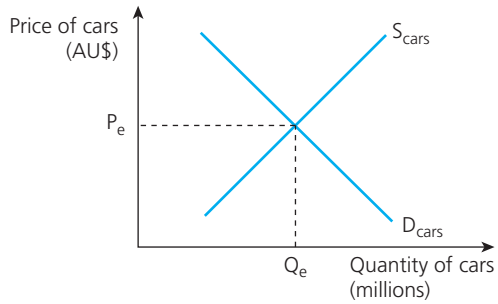
◆ A **floating exchange rate** is where the value of a currency is determined by the demand for and supply of the currency in the foreign exchange market.

As the value of currencies in a floating exchange rate system are determined by the market forces of demand and supply, economists treat currencies like any other good or service in microeconomics. The value of a currency increasing in a floating exchange rate system is referred to as an **appreciation** of the currency. The value of a currency decreasing in a floating exchange rate system is referred to as a **depreciation** of the currency.

◆ **Appreciation** is a sustained increase in the value of one currency in terms of another under a floating exchange rate system.

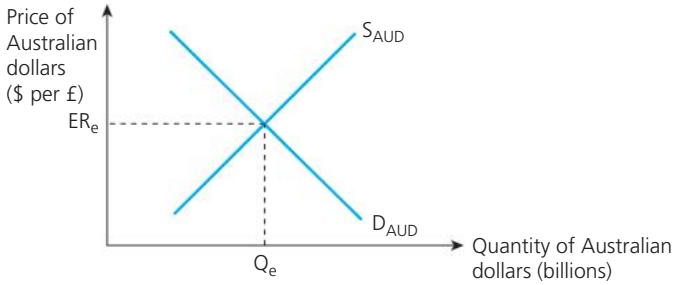
◆ **Depreciation** is a sustained decrease in the value of one currency in terms of another under a floating exchange rate system.

In microeconomics, if we were to use a diagram to illustrate the market for a car, we would draw the demand and supply curves, with the *y*-axis labelled *Price of cars* measured in specific currency and the *x*-axis labelled *Quantity of cars* using a suitable unit of measurement (see Chapter 5). In Figure 31.2, the market for cars in Australia is represented. The supply curve shows the supply of cars in Australia at different price levels and the demand curve represents the demand for cars in that market. An equilibrium price is established at P_e, and the equilibrium quantity traded is at Q_e.



■ **Figure 31.2** Market for cars in Australia

When analysing the market for foreign currencies, it is best to treat them in a similar way. When examining the exchange rate, economists measure the value of a currency in terms of another currency, for instance, the value of the Australian dollar (AUD or $) in terms of British pounds (GBP or £).



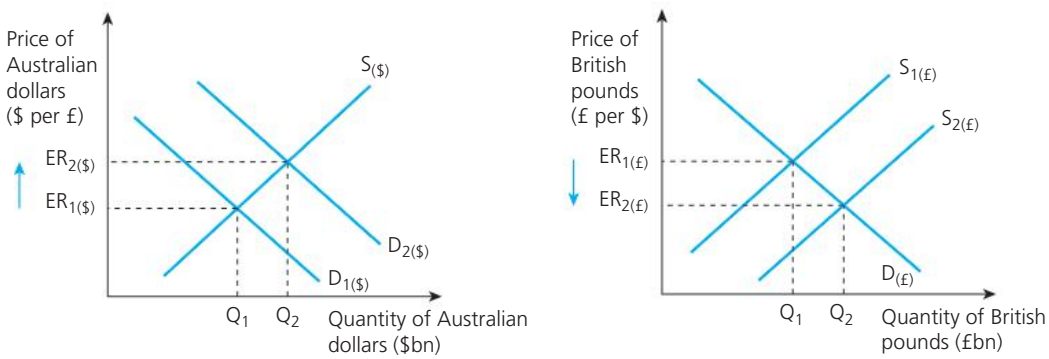
■ **Figure 31.3** Market for Australian dollars in British pounds

In Figure 31.3, the *y*-axis is labelled the *Price of Australian dollars ($ per £)*, as we are measuring the price of the Australian dollar expressed in terms of another currency (£ in this case). The *x*-axis is labelled *Quantity of Australian dollars* and in this case the units are likely to be in terms of billions (as billions of dollars are exchanged between these countries). The supply curve represents the supply of Australian dollars in the foreign exchange market, which comes from holders of Australian dollars willing and able to sell at a range of exchange rates. The demand for Australian dollars on the foreign exchange market comes from those wishing to purchase Australian dollars at a range of exchange rates. The market forces of demand for a currency and the supply of the currency determine the equilibrium exchange rate (ER_e) and quantity (Q_e).

In Figure 31.4, assume there is an increase in the demand for Australian dollars from $D_{1(\$)}$ to $D_{2(\$)}$ caused by higher demand in the UK for Australian goods. This is matched by an increase in the supply of British pounds from $S_{1(£)}$ to $S_{2(£)}$ by British buyers who exchange their pounds for dollars to complete their transactions. This means that an increase in the value of the Australian dollar against the British pound from $ER_{1(\$)}$ to $ER_{2(\$)}$ is matched by a fall in the value of the British pound against the Australian dollar from $ER_{1(£)}$ to $ER_{1(£)}$. In other words, an appreciation of the Australian dollar against the British pound must result in the pound depreciating in value against the dollar.

When changes in demand and supply occur in a currency market, the changes in the value of a currency are determined by the price mechanism, that is, buyers and sellers on the foreign exchange market. This works in the same way that the price and quantity of any other good or service is determined in a perfectly competitive market without government intervention.

In the case of an appreciation of the Australian dollar in Figure 31.4, the increase in demand for the AUD results in excess demand for Australian dollars from Q_2 to Q_1. This causes the rationing process to begin and signals to holders of Australian dollars that the price of the AUD is about to increase. This generates an incentive for holders of Australian dollars to increase the quantity supplied of AUD in anticipation of a better price. The rationing process causes those wanting



■ **Figure 31.4** Appreciation of the Australian dollar

Top tip!

It is wise to label each of the demand and supply curves on your diagram with the currency that is being examined (such as $S_\$$ and $D_\$$). If the current exchange rate is known, then this can be stated on the *y*-axis. If this is not known, or it is hypothetical, then label this as *ER* for the exchange rate.

to purchase AUD to bid-up what they are willing to exchange for them. This will lead to a new equilibrium exchange rate of $ER_{2(\$)}$ and Q_2 billions of AUD traded.

In the case of a currency depreciation, as shown in Figure 31.5, holders of Australian dollars wish to purchase British pounds. To do so, they must first exchange (sell) their Australian dollars, which causes an increase in the supply of Australian dollars from $S_{1(\$)}$ to $S_{2(\$)}$. This causes a fall in the value of the Australian dollar as the exchange rate drops from ER_1 to ER_2. In response to this, holders of Australian dollars will accept lower bids for their dollars, as they are less scarce. This signals to the market that the price of Australian dollars is likely to fall. This results in an expansion in the quantity demanded for Australian dollars, leading to the new equilibrium of exchange rate of ER_2 and Q_2 quantity traded.

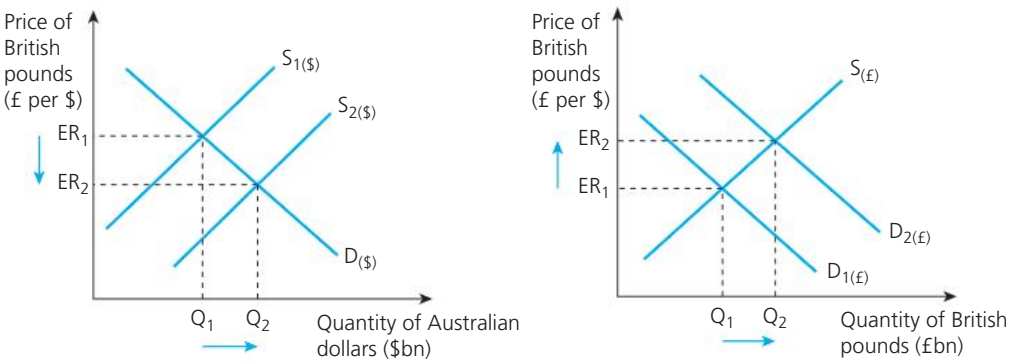


Figure 31.5 Depreciation of the Australian dollar

Top tip!

The demand for a currency can be seen as an inflow of money into an economy. The only reason why consumers would purchase a foreign currency is to spend it on foreign goods or services in overseas markets. The supply of a currency can be seen as an outflow of money from an economy, as the only reason consumers sell their own currency is to demand another.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 31.1

Suppose the exchange rate between the British pound (£) and the US dollar ($) is £1 = $1.35. Calculate the price for customers in Britain buying US cars priced at $35,500. *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 31.2

a If the exchange rate of the US dollar to pounds sterling is $1 = £0.65, calculate the price paid in pounds sterling by a British tourist spending $65 on a theme park ticket in Florida, USA. *[2 marks]*

b If the US dollar $ depreciates against the pound sterling to $1 = £0.60, calculate the new amount that British tourists would have to pay in pounds sterling for the same entrance ticket. *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 31.3

a Suppose the exchange rate between the Australian dollar ($) and the Chinese yuan renminbi (¥) is $1 = ¥6.5. Calculate the price for customers in China of buying textbooks priced at $65 from Australia. *[2 marks]*

b Suppose that the exchange rate between the Canadian dollar ($) and the British pound (£) is $1 = £0.65 and the euro (€) is $1 = €0.75. Calculate the exchange rate of the British pound against the euro. *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 31.4

With reference to the diagram (Figure 31.6) below, calculate the exchange rate in terms of US dollars. *[2 marks]*

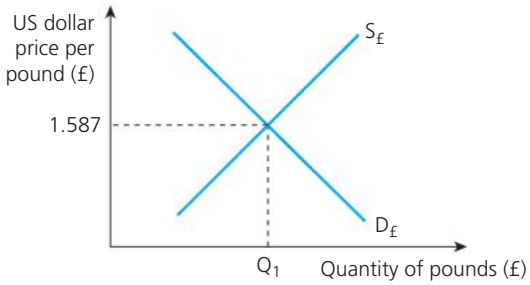


Figure 31.6 The market for pounds sterling (£)

Changes in demand and supply for a currency (AO2)

The value of a currency is determined by changes in demand and supply in the currency market (or the foreign exchange market). These changes result in a change in the value of a currency through the price mechanism. This is the same way that the price and quantity of any other good or service is determined in a perfectly competitive market without government intervention.

It is important to note here that any factor that generates demand for a currency will cause an increase in the supply of another currency. After all, the definition of exchange rates is the value of one currency in terms of another. Buyers of a currency must first exchange their own currency in order to buy the foreign currency. This feature of exchange rates has great importance for the balance of payments, which will be discussed in Chapter 32.

Another factor to consider is that of *ceteris paribus*. When analysing each of the factors impacting demand and supply of currencies, we must assume that all other factors remain equal. An increase in the demand for an economy's currency will cause an appreciation of that currency only if it is not matched by a factor increasing the supply of that currency.

The IB Economics syllabus states 10 factors that cause a change in the demand for and supply of a currency: foreign demand for exports, domestic demand for imports, inward versus outward foreign direct investment, inward versus outward portfolio investment, remittances, speculation, relative inflation rates, relative interest rates, relative growth rates and central bank intervention.

Foreign demand for exports

When foreign consumers wish to purchase goods or services from a foreign economy, they must first purchase the currency of that economy to make the transaction (via a foreign currency market). This is because domestic producers are highly unlikely to accept purchases made in foreign currencies as the money is not legal tender in the domestic economy.

Therefore, it can be said that whenever a foreign consumer purchases a good or service exported from a foreign economy, they will have to demand that economy's currency and supply their own. This leads to an appreciation in the currency of the foreign economy and a depreciation of their own currency, ceteris paribus. Likewise, should there be a fall in the demand for an economy's exports, then there will be a fall in the demand for that economy's currency and ultimately a depreciation will occur.

CASE STUDY

Mansudae Overseas Projects

Mansudae Overseas Projects is a construction company based in Jongphyong-dong, Pyongyang, North Korea. In the decade up to 2020 it earned an estimated $160 million, building Soviet style monuments and memorials in countries like Angola, Namibia, Benin, Botswana, Ethiopia and Democratic Republic of the Congo. The North Korean government said it would use the foreign exchange earned by the company to bolster its foreign exchange reserve and to pay for imports from countries like China.

Domestic demand for imports

It is logical that an economy's exports are another economy's imports. Therefore, when a good is imported into the domestic economy, this creates an increase in the supply of that economy's currency and a demand for the overseas economy's currency. Naturally, it would be an odd experience for consumers going to the supermarket in New York, London or Tokyo and paying in euros for Greek grapes and in lira for Turkish tomatoes. However, the suppliers of these imported goods will have been paid in their domestic currencies by the purchaser, such as distributors, supermarkets or other suppliers in the supply chain. Therefore, higher demand for imports will increase the supply of the domestic currency and lead to a currency depreciation, whereas it will increase the demand for the foreign currency and cause that to appreciate in value, ceteris paribus.

Inward/outward foreign direct investment (FDI)

Foreign direct investment (FDI) is the spending by multinational corporations (MNCs). **Inward FDI** refers to foreign MNCs expanding their operations in the domestic economy, such as investments in production facilities. For example, Manulife is a Canadian financial and insurance company with operations in 15 countries, serving more than 26 million customers worldwide. Therefore, its operations in overseas markets including Malaysia, Singapore and Hong Kong represent inward FDI for these economies. The result of this would be an increase in the supply of the Canadian dollar to demand the Malaysian ringgit, the Singaporean dollar and Hong Kong dollars. **Outward FDI** refers to domestic MNCs expanding their operations in overseas markets. In the above example, the spending by Manulife in overseas markets represents outward FDI for Canada.

◆ **Inward foreign direct investment** refers to foreign multinational companies expanding their operations in the domestic economy.

◆ **Outward foreign direct investment** refers to multinational companies from the domestic economy expanding their operations in overseas markets.

Figure 31.7 Finance companies require foreign exchange for their FDI

Therefore, it can be said that outward FDI for one country creates inward FDI for another. Inward FDI creates higher demand for a currency leading to its appreciation, ceteris paribus. Outward FDI leads to the supply of the domestic currency causing it to depreciate, ceteris paribus.

Inward/outward portfolio investment

Portfolio investment is the purchase of financial investments abroad, such as stocks, shares and bonds of overseas companies and governments. When investments are made in a foreign economy, the investors or their representatives will need to purchase the currency of that economy to complete the investment. For example, if investors from Japan wish to buy shares of US companies listed on the New York Stock Exchange, those investors will need to sell Japanese yen in order to purchase US dollars to complete their investment.

◆ **Portfolio investment** is the purchase of financial investments abroad, such as the purchase of stocks, shares and bonds of overseas firms and governments.

Inward portfolio investment represents spending in the domestic economy by foreign investors. Those investors must supply (sell) their own currencies on the foreign exchange market and demand (buy) the currency of the economy that they are investing in. This will cause a depreciation in the currency held by those investors and an appreciation of the currency in the economy that these investors are investing in.

Outward portfolio investment represents spending by an economy's investors in overseas markets. All things being equal, this will result in an increase in the supply of the domestic economy's currency and an increase in the demand for the currency of the other economy that investors are investing in. This will lead to a depreciation of the domestic currency and an appreciation of the currency in the other economy, ceteris paribus.

Remittances

Remittances refer to the movement of money when nationals working abroad send money back to their home country, thereby impacting the exchange rate. In an increasingly globalized world, it is not uncommon for citizens of one economy to work temporarily in another economy. These citizens, known as expatriates, may be attracted by higher salaries or favourable working conditions in overseas countries. For these citizens to be considered expatriates (or expats), there must be an assumption that they will return 'home' after a period of time. An example is teachers working overseas in international schools. A significant number of IB World Schools hire expatriate teachers.

◆ **Remittances** refer to the movement of money when nationals working abroad send money back to their home country.

While expats are working abroad, they may send some of their income back to their home country, either to their own bank accounts or to family members. For example, there are over 400,000 foreign domestic helpers working in Hong Kong, representing more than 5 per cent of the population of Hong Kong. Most of these foreign workers come from the Philippines, Indonesia and Thailand. Typically, these foreign domestic helpers send money home to their families on a monthly basis. To do so, they need to supply the Hong Kong dollar (the currency they are paid in) and demand their home currency (such as the Philippine peso, Indonesian rupiah and Thai bhat). This will result in a depreciation of the currency of the economy that they are working in and an appreciation of the currency of their home economy, ceteris paribus.

CASE STUDY

The Philippine's personal remittances

Personal remittances from Overseas Filipinos reached a record high of $33.5 billion in 2019, 3.9 per cent higher than the $32.2 billion recorded in 2018. Personal remittances are very important for the Philippines, which help to boost household income and consumption. In 2019, personal remittances accounted for 9.3 per cent and 7.8 per cent of the country's gross domestic product (GDP) and gross national income (GNI), respectively.

Figure 31.8 Manila, Philippines

Source: Bangko Sentral ng Pilipinas (2020) **http://www.bsp.gov.ph/publications/media.asp?id=5290**

Speculation

Speculation occurs when a financial asset is purchased in the hope or anticipation that the resale value will be higher. It is said that the vast majority of currency trades are speculative, which means that currency traders are purchasing different currencies in the hope that their value will increase, and they can sell these currencies at a favourable exchange rate. Similarly, if speculators expect a currency to fall in value, they will tend to sell (supply) this in favour of other stronger or more stable currencies or other assets such as gold.

Foreign exchange traders and investment companies move money around the world to take advantage of variations in exchange rates to earn a profit. As huge sums of money are involved owing to currency speculation, this can itself cause exchange rate fluctuations.

◆ **Speculation** occurs when a financial asset, such as the domestic currency or a foreign currency, is purchased in the hope or anticipation that the resale value will be higher.

Relative inflation rates

Inflation is a sustained increase in the general price level of an economy (see Chapter 20). An increase in the price of goods and services caused by domestic inflation will tend to decrease the demand for exports. According to the law of demand (see Chapter 3), this will lead to a reduction in the demand for a currency and therefore a depreciation of that currency. Speculators holding that currency may choose to sell it and purchase another more stable currency. This will lead to an increase in the supply of the domestic currency, further depreciating the value of that currency.

◆ **Inflation** is the sustained (continual) increase in the general (average) price level, over time. It is commonly measured by using a consumer price index (CPI).

Relative interest rates

Interest rates are the costs of borrowing money and the rewards for saving money (see Chapter 23). Investors may wish to save in an economy that has higher interest rates than in their own country. To do this, investors will need to purchase the currency of the foreign country that they will be saving their money in. This will cause an increase in the supply of the domestic currency of the investors but lead to an increase in the demand for the foreign currency.

By contrast, a cut in interest rates will tend to reduce incentives for investors, so they sell that currency in search of investments with better returns, ceteris paribus. This will therefore reduce the exchange rate of that currency.

◆ The **interest rate** is the price of money. Expressed as a percentage measure, interest rates represent the cost of borrowing money or the return for savers.

Relative growth rates

Economic growth is an increase in a country's real national income over time (see Chapter 18). Higher levels of economic growth in a foreign economy, represented by increases in real GDP, are indicative of rising wages and rising domestic consumer expenditure leading to an increase in aggregate demand (AD). The increase in AD can cause demand-pull inflation and will increase the likelihood of contractionary or tight monetary policy (see Chapter 23). The resulting higher interest rates will generate an increase in the demand for the currency of that economy, which is likely to lead to an appreciation in the value of that currency. Hence, economic growth tends to cause the central bank to raise interest rates thereby causing the exchange rate to also appreciate, ceteris paribus.

◆ **Economic growth** is a key macroeconomic objective that measures the increase in the level of national output over a specified period of time. It can be measured by the annual percentage change in the economy's real gross domestic product (GDP).

Central bank intervention

The central bank is the monetary authority in a country (see Chapter 23). In addition to the setting of interest rates, a central bank may also introduce measures to restrict the supply or sale of its currency in order to control its money supply. A good example of this is the current policies of the People's Bank of China, where the Chinese yuan (or renminbi) can be legally traded only from approved outlets. This prevents governments or large groups of investors from manipulating China's currency and therefore gaining a competitive edge in export markets. Currencies that are subject to these controls are likely to face lower levels of demand as they become less desirable to speculators.

Figure 31.9 The national currency of China is the yuan (also known as the renminbi)

Top tip!

The demand for a currency can be seen as an inflow of money into an economy. The only reason why consumers would purchase a foreign currency is to spend it on foreign goods or services in overseas markets. The supply of a currency can be seen as an outflow of money from an economy, as the only reason consumers sell their own currency is to demand another.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 31.5

With reference to the diagram (Figure 31.10) below, outline **two** possible reasons for the change in the exchange rate of the Kuwaiti dinar (KWD) against the Nigerian naira (NGN). *[4 marks]*

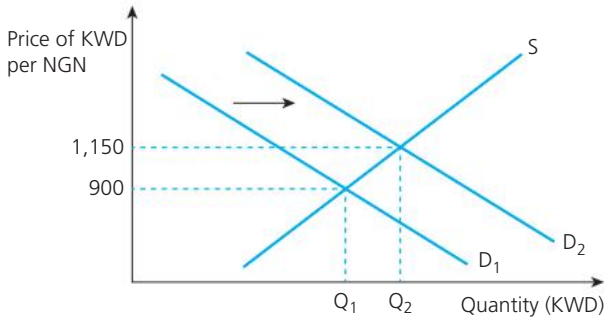


Figure 31.10 The market for Kuwaiti dinars (KWD)

◆ **Net exports,** the value of exports minus the value of imports, are part of the aggregate demand equation (and GDP).

Consequences of changes in the exchange rate on economic indicators (AO3, AO4)

Fluctuations in exchange rates cause changes in the price of imports and exports, and therefore the demand for imports and exports. Recall from Chapter 17 that **net exports**, the value of exports minus the value of imports, are part of the aggregate demand equation (and GDP). Therefore, changes in net exports will have macroeconomic consequences. In particular, changes in the exchange rate impact on economic indicators, such as the inflation rate (price stability), economic growth, unemployment, the current account balance and living standards. The consequences of changes in the exchange rate on economic indicators can be remembered by using the acronym **CUPEL** (Figure 31.11).

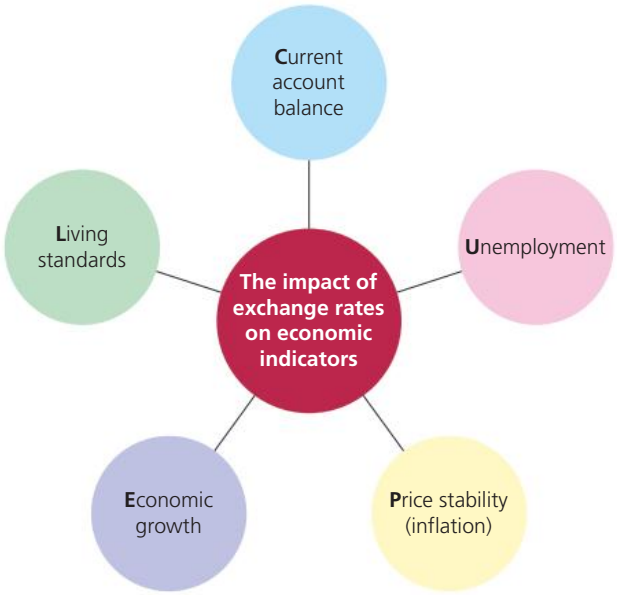


Figure 31.11 Consequences of changes in the exchange rate on economic indicators

The inflation rate

Fluctuations in the value of an economy's currency can lead to various forms of inflation, disinflation and even deflation in extreme circumstances. Changes in the exchange rate can impact both cost-push and demand-pull inflation. **Demand-pull inflation** is a sustained increase in the general price level due to an increase in the components of factors of production (increased consumption, investment, government spending and net export expenditure). **Cost-push inflation** is a sustained increase in the general price level due to an increase in the costs of production in an economy. These major causes of inflation are examined in Chapter 20.

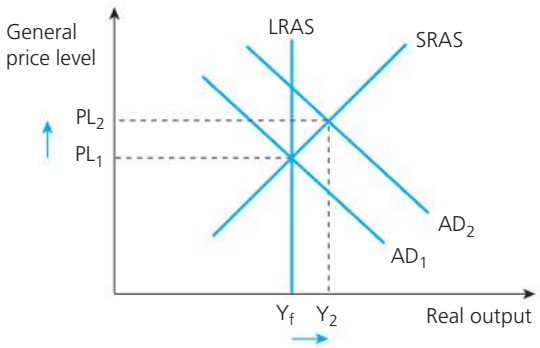


Figure 31.12 Currency depreciation and demand-pull inflation in the new classical model

As illustrated in Figure 31.12 and Figure 31.13, a currency depreciation will generally lead to a reduction in the quantity demanded of imports and an increase in the quantity demanded of exports. This is because the relative price of exports becomes cheaper, whereas there is a relative increase in the price of imports. This would increase the value of net exports (X – M). As a component of aggregate demand, this will increase AD from AD_1 to AD_2 and causes real output to increase from Y_f to Y_2 and from Y_1 to Y_f respectively, ceteris paribus. However, it also results in an increase in the general price level of PL_1 to PL_2.

In the case of an appreciation of the domestic currency, the price of exported goods will increase, and the price of imported goods will decrease. This is likely to lead to a reduction in net exports (X – M). Hence, aggregate demand is likely to fall from AD_1 to AD_2 as shown in Figure 31.14 and Figure 31.15. All things being equal, the fall in net exports causes the level of real output to decline from Y_f to Y_2. This will lower the general price level in the economy from PL_1 to PL_2.

A depreciation in the value of a currency can also be a cause of cost-push inflation. As firms may need to rely on the import of raw materials and components, for example, a lower value of the currency will lead to an increase in the price of imports and therefore an increase in the costs of production for these firms. This will cause a reduction in short run aggregate supply, causing a shift in the SRAS curve from $SRAS_1$ to $SRAS_2$ in Figure 31.16. This will result in scarcity and therefore an increase in the general price level from PL_1 to PL_2. This is representative of cost-push inflation. In this scenario, the higher production costs caused by the currency depreciation leads to lower economic growth as real GDP contracts from Y_f to Y_2.

Note that the change in the exchange rate, such as depreciation, can have a positive or negative impact on the level of real national output. This will largely depend on the extent to which the country relies on imports for production.

Economic growth

Economic growth is an increase in the value of real national income over time. An appreciation of the currency will lead to an increase in the price of exports and a decrease in the price of imports. The result of this may be a decrease in net exports, although this will depend on the value of price elasticity of demand for the country's exports. Net exports (X – M) are a component of aggregate demand. So, the decrease in the price of imports caused by the currency appreciation could lead to a higher value of import expenditure if the PED for imports is price elastic. If this materializes, there will be a net leakage from the circular flow of income (see Chapter 1), which reduces the value

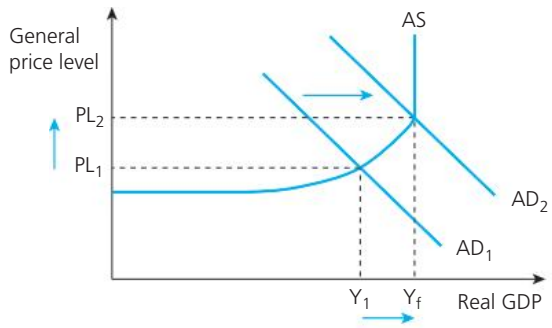


Figure 31.13 Currency depreciation and demand-pull inflation in the Keynesian model

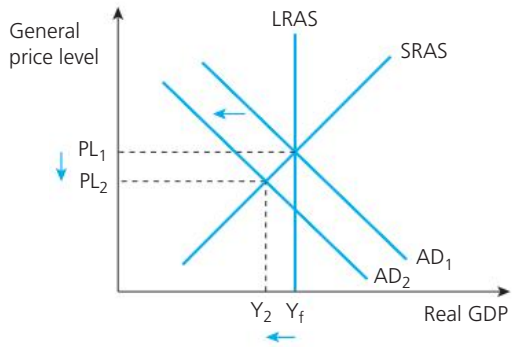


Figure 31.14 Currency appreciation and lower demand-pull inflation in the new classical model

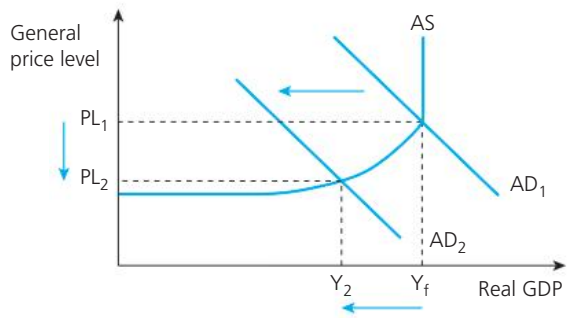


Figure 31.15 Currency appreciation and lower demand-pull inflation in the Keynesian model

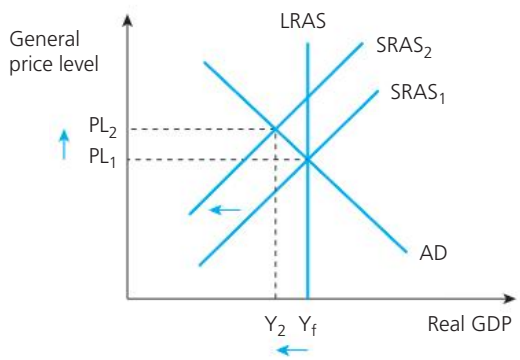


Figure 31.16 Currency depreciation and cost-push inflation in the new classical model

of aggregate demand in the economy. This could lead to economic decline, or even a recession. This would be represented by a reduction in real GDP in an AD–AS diagram. Economic growth is covered in more detail in Chapter 18.

Unemployment

An appreciation of a currency may lead to a significant reduction in the demand for exports, as they are now more expensive in foreign economies. This reduction in export revenues may eventually result in firms reducing their demand for labour (due to derived demand). This will generate some unemployment in the economy. This could be classified as cyclical due to a lack of aggregate demand caused by the fall in net exports, but some economists may argue that it is structural in nature if it occurs in the long run. For example, the tourism industry is particularly vulnerable to fluctuations in the exchange rate; tourists are more likely to visit a foreign country if their own currency is relatively higher in value due to a currency appreciation, ceteris paribus.

The above analysis suggests that a fall in the value of a currency tends to be beneficial for economy. This is because net export sales are likely to rise, leading to greater derived demand for those working in export-driven industries. Hence, a currency depreciation is likely to cause less unemployment in the economy, ceteris paribus.

The current account balance

The **current account** on the balance of payments (see Chapter 32) is a record of all exports and imports of goods and services, plus net investment income from overseas assets plus the net balance of transfers made between countries by individuals and governments. Given the current account balance (or value) is dependent on the international flow of money in and out of a country, fluctuations in the exchange rate will have a direct impact on the account.

To assume that an appreciation of a currency will lower export revenue and increase import expenditures, economists assume that the price elasticity of demand for exports and imports is greater than one, that is, price elastic. This ensures that the reduction in the demand for exports as a result of an increase in their price will lead to a greater than proportionate decrease in the quantity demanded of exports. Similarly, for the currency appreciation to lead to an increase in import expenditure, the reduction in the relative price of imports must lead to a greater than proportionate increase in the quantity demanded of imports.

Should these conditions exist (the PED for exports and imports being price elastic), the result of a currency appreciation will be a reduction in the value of the current account. This means a worsening in the current account balance, or a movement towards a deficit.

The opposite is true in the case of a currency depreciation. A fall in the value of a currency will, assuming an elastic PED value for imports and exports, lead to an increase in the value of the current account balance, that is, an improvement or a move towards a surplus. This is because imports will be relatively more expensive for domestic firms and consumers, while exports will be relatively cheaper for overseas buyers.

However, it is important to note that currency fluctuations can also be a result of current account imbalances. Should a current account be in deficit, this would mean that the value of imports exceeds the value of exports, which should mean that the supply of the economy's currency will exceed its demand, thus causing a depreciation. In the case of a current account surplus, where the value of exports exceeds the value of imports, the demand for that economy's currency will exceed the supply of it, meaning that there should be an appreciation. Remember that all of this analysis assumes that the economy has a freely floating exchange system.

Top tip!

Whether a currency appreciation will benefit economic growth depends on the PED for the country's imports and exports. For example, oil-rich nations such as Saudi Arabia and Kuwait may benefit from a higher exchange rate as the demand for oil is highly price inelastic, meaning foreign countries will spend more on oil if the price increases. This generates higher export sales revenue to these oil-exporting nations.

◆ The **current account** is a record of all trade flows (exports and imports of goods and services), income flows and income transfers between countries by individuals, firms and governments.

Top tip!

Remember to consider the interconnectedness of the content, contexts and concepts studied in economics. For example, higher unemployment caused by a **change** in the currency would lead to a change in spending, which would therefore also lead to a change in the rate of economic growth. This illustrates the **interdependence** of macroeconomic objectives and policies.

Living standards

The value of a country's currency can have a significant impact on the standard of living in that economy. A strong currency will bring benefits to consumers of imported goods and services as they will need to exchange less of their own currency or money to purchase the imported items. In more economically developed economies, these imported items are more likely to be wants (desires) and luxury goods and services. Hence, it could be argued that a strong currency can be a cause of high living standards rather than just an effect.

Figure 31.17 The falling Iranian rial over the last decade has left a negative impact on the standard of living of Iranian people

Alternatively, it could be argued that a strong currency will reduce the competitiveness of the exports of that economy owing to relatively higher export prices. This could lower export revenues and reduce the demand for labour from local firms. This could eventually lead to an increase in unemployment in the economy, reduce consumption expenditure, lower economic growth and cause a reduction in standards of living.

However, a relatively weak currency could bring cost-push inflation to the economy as domestic firms and consumers have to pay more for essential imported goods. In the case of some economically less developed countries (ELDCs), this could make it difficult for them to access essential goods that are typically imported, along with imports of physical capital needed to diversify their production away from low-value agricultural products.

By contrast, a weaker currency could make an economy's exports more competitive and therefore increase export revenues. This has a positive impact on the demand for labour, therefore increases employment and incomes in the economy. This is likely to then lead to increased economic growth and an improvement of living standards.

Unfortunately, in the case of some ELDCs, their overreliance on agriculture for export revenues means that the depreciation of their currency could lower export revenues as many primary products have a low (inelastic) price elasticity of demand. This means that the percentage increase in the quantity demanded for their exports will be less than the percentage reduction in price, thus causing a reduction in export revenues. This will then cause a drop in living standards in these ELDCs.

PAPER 1 EXAM PRACTICE QUESTION 31.6

a Explain what is meant by an 'appreciation' in the value of a currency. *[10 marks]*

b Using real-world examples, discuss the macroeconomic consequences of a fall in a country's exchange rate. *[15 marks]*

Fixed exchange rate (AO2, AO4)

A **fixed exchange rate system** exists when the central bank (or central monetary authority) of a country buys and sells foreign currencies to ensure the value of its currency stays at a single fixed or predetermined rate against another currency or basket of other currencies. For example, since 15 October 1983, the Hong Kong dollar has been fixed against the US dollar at a rate of HK$7.80 to US$1. The government can maintain this single rate by buying (demanding) or selling (supplying) foreign currency reserves as needed. **Foreign currency reserves** (or **reserve assets**) are stocks of foreign currencies held by central banks, typically used to influence the value of a currency.

◆ A **fixed exchange rate system** exists when the central bank (or monetary authority) buys and sells foreign currencies to ensure the value of its currency stays at a single, predetermined rate.

◆ **Foreign currency reserves** (or **reserve assets**) are stocks of foreign currencies held by a central bank, usually to influence the value of its currency.

There is a range of reasons why a government may wish to fix the value of its currency against one, or a basket of other currencies. Two common reasons for doing so relate to export competitiveness and the current account balance. The value of a currency helps to determine the price of an

economy's exports and the price of its imports. This, along with price elasticity of demand, can impact the revenues received from exports and the economy's expenditure on imports. Hence, the economic stability created by fixed exchange rates tends to have a positive impact on the current account balance, economic growth and employment in the domestic economy.

Devaluation and revaluation of a currency (AO2)

When a government deliberately attempts to lower the value of its currency against another, this is referred to as devaluation. **Devaluation** occurs when the price of a currency operating in a fixed exchange rate system is officially and deliberately lowered against another currency or basket of currencies. **Revaluation** occurs when the price of a currency operating in a fixed exchange rate system is officially and deliberately increased.

◆ **Devaluation** occurs when the price of a currency operating in a fixed exchange rate system is officially and deliberately lowered.

◆ **Revaluation** occurs when the price of a currency operating in a fixed exchange rate system is officially and deliberately increased.

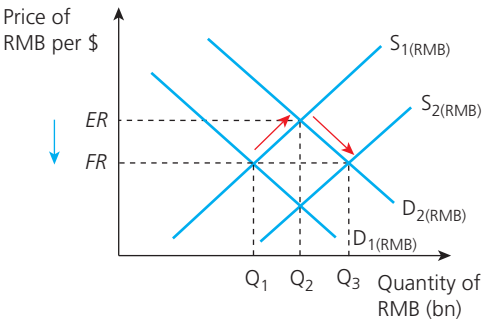
In theory, the international competitiveness of a country can be improved by a currency devaluation. This is because exports will become relatively cheaper and imports become more expensive. As domestic consumers and firms must exchange more of their own currency to buy imported goods and services, this is likely to reduce the expenditure on imports. By contrast, the devaluation will make exports relatively cheaper for foreign customers. Hence, a devaluation will improve the value of the country's net exports, ceteris paribus.

However, at other times, governments may wish to revalue their currency when they wish to import more essential goods and services. This can often occur after natural disasters, when the government needs to import resources for large reconstruction projects. These projects can cost hundreds of billions of dollars, so even a 5 per cent revaluation can lower the government's expenditure significantly.

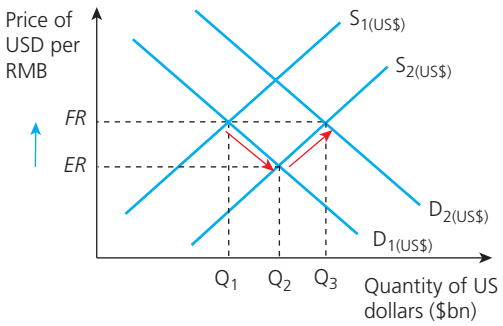
How fixed exchange rates are maintained (AO2, AO4)

As with a freely floating exchange rate system, the value of a currency in a fixed exchange rate regime is determined by the market forces of demand and supply. In a freely floating exchange rate system, a currency should always be moving from one equilibrium position to another. In the short run, excess demand or supply for a currency may exist but the price mechanism will use the signalling, rationing and incentive functions to return the currency to a point of equilibrium. However, in a fixed exchange rate system the government needs to match changes in the supply or demand for a currency in order to keep it at the fixed rate.

This can be explained through a hypothetical example, using the Chinese yuan renminbi (RMB) and the US dollar. These two countries are each other's largest trading partners. China exports vast amounts to the USA each year and may therefore wish to mitigate the impact of exchange rate fluctuations on the competitiveness of its exports. One method China has used to ensure this is to devalue its currency against the US dollar.



US customers buy the RMB to purchase Chinese exports. The Chinese government sells the RMB to maintain the fixed exchange rate.



US customers sell the USD to buy Chinese imports. The Chinese government buys the USD to maintain the fixed exchange rate.

Figure 31.18 Devaluation under a fixed exchange rate system

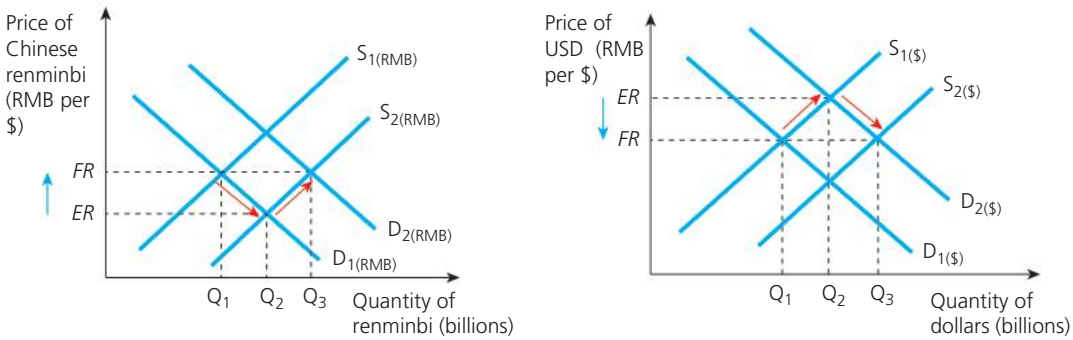
In Figure 31.18, the Chinese government via the People's Bank of China (the nation's central bank) fixes the value of the RMB against the USD at a fixed rate of *FR*. This means that one RMB will be worth *FR* units of USD. Suppose there is an increase in the demand from the USA for Chinese exports, which will lead to an increase in the demand for the RMB to purchase these exports. In a floating exchange rate system, this would lead to an appreciation of the RMB from *FR* to *ER*, but a potential loss of China's export competitiveness.

■ **Figure 31.19** Governments may choose to devalue their currencies to maintain export competitiveness

To maintain the value of the RMB at the fixed rate (*FR*), the Chinese government must devalue the currency. It does this by matching the increase in the demand for RMB (from D_{1RMB} to D_{2RMB}) with an increase in the supply of the RMB from $S_{1(RMB)}$ to $S_{2(RMB)}$. It does this through the purchase of the USD (that is, selling the RMB and buying the USD). When the Chinese government purchases more USD, they will increase the supply of RMB, thus lowering the value, but also increase the demand for USD, therefore increasing its value. The deliberate attempt by the Chinese government to shift the exchange rate from *ER* back down to *FR* is referred to as a devaluation.

Note that it is important that the Chinese government matches the increased demand for its currency with the supply exactly and simultaneously to maintain the fixed rate. If the government is unable to do this, then the exchange rate will not be fixed so it will fluctuate, causing uncertainty in the export market.

In Figure 31.20, suppose the Chinese government, or the People's Bank of China, wishes to fix the value of its currency at a rate above the market equilibrium (revaluation). A common reason for this is the need for Chinese firms to import a large amount of their factors of production. When customers in China wish to purchase imports from the USA, this increases the supply of Chinese RMB on the foreign exchange market so should, in a freely floating exchange rate system, lead to a depreciation of the RMB against the US dollar. To prevent this, the Chinese government must sell its foreign reserve assets of USD and buy RMB instead. This will increase the supply of USD on the foreign exchange market from $S_{1(\$)}$ to $S_{2(\$)}$, thus depreciate the US dollar, and increase the price of the Chinese renminbi from *ER* back to *FR*. The deliberate attempt by the Chinese government to shift the exchange rate from *ER* back up to *FR* is referred to as a revaluation.



Chinese customers sell the RMB to purchase US dollars for their purchase of US exports. The Chinese government buys back the RMB to maintain the fixed exchange rate.

Chinese customers buy the USD to buy US imports. The Chinese government sells the USD to maintain the fixed exchange rate.

■ **Figure 31.20** Revaluation under a fixed exchange rate system

■ **Figure 31.21** The Hong Kong dollar is pegged to the US dollar

Managed exchange rates (AO2, AO4)

A **managed exchange rate** is a system where the government or central monetary authority intervenes periodically in the foreign exchange market to influence the exchange rate. The exchange rate under a managed system is primarily determined in the foreign exchange market by the market forces of demand for and supply of a currency. The government intervenes only occasionally when deemed necessary to influence the exchange rate. The government achieves this by influencing the demand for and supply of the domestic currency. The main reason for such intervening when necessary is to prevent large and sudden or unexpected fluctuations in the exchange rate, which could happen if currencies were left entirely to the market forces of demand and supply. Hence, managing the exchange rate helps to create or maintain certainty and confidence in the economy.

A common method of managing a currency is through a '**crawling peg**' system. This involves the setting of two bands or limits, which the government strives to keep the currency value within by periodic intervention. For example, the State Bank of Vietnam (SBV) uses a crawling peg system where it allows the Vietnamese dong (VND) to fluctuate against the US dollar by +/– 3%.

◆ A **managed exchange rate** is a system where the government or central monetary authority intervenes periodically in the foreign exchange market to influence the exchange rate, when deemed necessary to maintain certainty and confidence in the economy.

◆ A **crawling peg** is a form of fixed exchange rate system in which a currency is permitted to fluctuate within predetermined bands of exchange rates.

In Figure 31.22, it can be seen that the State Bank of Vietnam is periodically intervening to keep the value of the Vietnamese dong within the two predetermined bands. Should there be a series of increases in the demand for the VND (cumulatively represented by the shift in the demand curve for the VND from D_1 to D_2), thus pushing the exchange rate up from ER_1 to ER_2. In response, the SBV will increase the supply of VND from $S_{1(VND)}$ to $S_{2(VND)}$ by purchasing other currencies, thereby reducing the value back to ER_1.

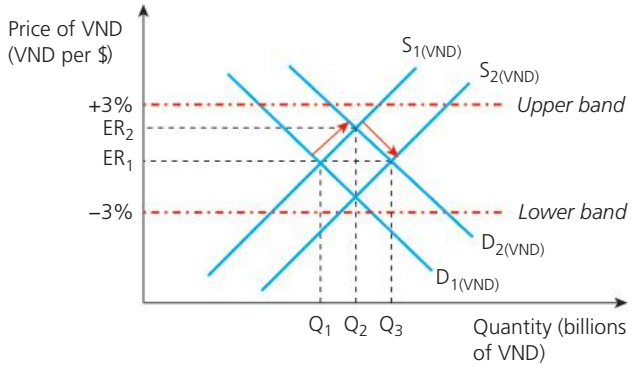


Figure 31.22 Managing the currency through a pegged exchange rate

On the contrary, if the exchange rate were to drop from ER_1 towards the lower band, the SBV would buy the Vietnamese dong using its currency reserves to increase the demand, thereby raising the exchange rate.

It is important that these interventions by the central bank are done on a periodic or ad-hoc basis in order to keep the currency within the upper and lower range, rather than at a fixed point. When a currency is overvalued, the government or central monetary authority will seek to reduce its value by selling the currency. By contrast, if the currency is undervalued, the government or central monetary authority will seek to increase the value by purchasing the currency. Through managing the exchange rate in this way, the government helps to prevent wild fluctuations in the exchange rate and build confidence in the economy.

Overvalued currencies

An **overvalued currency** occurs when the value of a currency is above its equilibrium value in the long run. In a freely floating currency exchange system, no currency should be overvalued in the long run. This is because the price mechanism will use the signalling, rationing and incentive functions to change the quantity demanded and supplied in order to bring the currency back to its equilibrium. In a fixed or managed system, the government or central monetary authority will intervene to correct this position.

◆ An **overvalued currency** occurs when the value of a currency is above its equilibrium value in the long run.

As a result of an overvalued currency, imported goods become cheaper while exports become relatively more expensive. This puts downward pressure on the rate of inflation in the economy. Domestic producers are forced to be more efficient in order to compete with the cheaper products from foreign producers and suppliers. Exports become less competitive, causing lower profits in export industries. As imports become cheaper and exports more expensive, there can also be a negative impact on the country's balance of payments.

Undervalued currencies

An **undervalued currency** occurs when the value of a currency is below its equilibrium value in the long run. As a result of an undervalued currency in an economy, exports become relatively cheaper, leading to economic growth and employment in export industries in the long run. Imports become more expensive for domestic consumers and firms, so they switch to buying domestic goods, ceteris paribus. This can lead to fewer leakages (withdrawals from the economy's circular flow of income). For some countries, this can lead to imported inflation when essential imports become more expensive, where imported raw materials and components are more costly, thus affecting the general price level.

◆ An **undervalued currency** occurs when the value of a currency is below its equilibrium value in the long run.

In a freely floating exchange rate system, an undervalued currency is corrected by the forces of demand and supply in the foreign exchange market. However, in a fixed or managed exchange rate system, the government or central monetary authority will intervene to correct this disequilibrium.

Common mistake

Students often use the terms 'managed' and 'fixed' exchange rates interchangeably – perhaps owing to the news media misusing these terms. It is important to remember that in a fixed exchange rate system, a currency is not permitted to move away from its fixed rate, therefore the government must match every change in demand with a corresponding change in supply, and vice versa. This must be done simultaneously otherwise the currency will move away from the fixed rate. A managed exchange rate system is more relaxed, with occasional interventions by the government or central monetary authorities aimed at maintaining the value of the currency within a predetermined range.

ATL ACTIVITY 31.1

With a partner, investigate why fixed exchange rate systems have become so rare in the modern global economy. Discuss the argument that there are more losers than winners in the economy if the government adopts a fixed exchange rate system. Be prepared to present your findings in class.

Fixed versus floating exchange rate systems (AO3) (HL only)

In general, fixed exchange rate systems bring greater benefits to firms than they do to governments. The disadvantages of fixed exchange rate systems can be seen as the advantages of a floating exchange rate system and vice versa. The arguments for and against fixed and floating exchange rate systems are examined below.

Certainty and stability

When a currency is fully fixed, investors and firms can be sure of the value of their investments. All investors seek returns on their investments, which can easily be eliminated by fluctuations in exchange rates. Therefore, a fixed exchange rate system reduces risks so makes investments in an economy more attractive. A fixed exchange rate regime also reduces some of the speculation that occurs in currency markets. For example, more than $6 trillion was exchanged each day on average in 2019 – only the USA and China had nominal GDP worth more than this amount.

Opportunity costs

To improve a country's long-term international competitiveness, governments need to engage with considerable reforms in their economy, rather than just relying on exchange rates. These reforms take a considerable amount of time, especially with supply-side policies (see Chapter 26). In a freely floating exchange rate system, any improvements made by the government can be negated by fluctuations in exchange rates. Therefore, a fixed exchange rate system reduces market volatility, thereby facilitating long-term improvements to the economy. However, this comes at the opportunity cost of the central monetary authority having to devote time and a significant amount of money to intervene in foreign exchange markets.

Currency liquidity

Currency liquidity refers to the availability of currency in a market. Governments that maintain a fixed exchange rate must accumulate vast foreign currency reserves and/or assets such as gold reserves in order to impact the demand and supply of their own currency. For example, when there is an increase in the demand for a fixed currency, this must be matched by an increase in its supply, which can be achieved by the government selling its currency to purchase the other. Over time, the purchasing of currencies by governments leads to a considerable reduction in the availability of currencies on global markets for private investors (both individuals and organizations). Hence, this crowds out or removes currency reserves for these private investors to invest in global economies.

A target for speculators

When governments have predetermined interventions in their currency markets, this represents predictable economic behaviour. Therefore, currency investors or speculators know that if they can generate a circumstance where a currency will fall below a predetermined level, the government will have to increase the value of the currency. Hence, consortiums of speculators, such as foreign exchange brokerage companies, can accumulate vast stockpiles of a particular currency over time in order to sell it all at once. In some cases, this may move the currency below the fixed amount (or the peg in a managed float exchange rate system) for a very short period of time, allowing the speculators to then buy the currency back and see an increase in the value of their holdings because the central monetary authority will need to increase their own demand for the currency to maintain the fixed rate or peg. The resources that governments use to counteract this represent a significant expense to the economy. The source of the Asian Financial Crisis is often stated as the failure of the Thai economy to maintain its fixed exchange rate to the US dollar, which was exacerbated by speculators trying to break the peg.

Figure 31.23 Currency speculators can cause fluctuations in exchange rates

CASE STUDY

Iranian Central Bank fights against currency speculation

In June 2020, the Iranian Central Bank announced it would defend the country's ailing currency against speculation, after the rial fell to historic lows. In 2018, the rial had fallen by about 75 per cent after the United States withdrew from a nuclear deal and re-imposed sanctions. The US dollar is used mainly by Iranians for imports of state subsidized food products and medicines.

Monetary policy

Under a fixed exchange rate system, the central bank often relies on the use of interest rate policy (see Chapter 23) to influence the exchange rate. In general, higher interest rates are needed to revalue or appreciate the value of the domestic currency, and vice versa. This means that economies operating under a fixed exchange rate regime have minimal if any use of monetary policy to influence the level of economic activity. By contrast, economies operating a floating exchange rate system are free to use monetary policy as a demand-side policy to influence the level of economic activity, growth, employment and the general price level (inflation).

Overall, fixed exchange rate systems, despite their popularity in the 1960s and 1970s are very rare indeed today. The globalized nature of trade and investment means that governments simply do not have the resources needed to match the changes in demand and supply of their currency. The closest example of a fixed exchange rate in economically developed economies is the Hong Kong dollar's peg to the US dollar. While this is technically a peg, the range of this is so small that essentially the Hong Kong Monetary Authority has to constantly intervene in the foreign exchange market, rather than periodically, to maintain the peg; essentially the margins are so insignificant that the system resembles a fixed exchange rate regime. The conditions of the Hong Kong economy, with its high concentration of international trade, makes this a unique example of an economy with the conditions needed to maintain such an exchange rate system. Today, very few economies around the world enjoy these conditions of trade and therefore would not find it viable to operate a fixed exchange rate system.

Theory of Knowledge

To what extent can ethical arguments be used to justify government intervention in foreign exchange markets?

WISE ChoICES

To what extent does government **intervention** in foreign currency markets improve **economic well-being**? Do the benefits of such intervention always result in improved **equity** in the economy?

WISE ChoICES

For a country of your choice, investigate how a **change** in the exchange rate through government **intervention** in the foreign exchange market affects different stakeholder groups. Discuss how these changes have impacted **efficiency** and **equity** in the economy.

PAPER 3 EXAM PRACTICE QUESTION 31.7 (HL ONLY)

■ Foreign currency reserves – June 2020

Country	USD millions ($)
China	3,101,700
USA	128,802

Source: Trading Economics – https://tradingeconomics.com/country-list/foreign-exchange-reserves?continent=g20

a Calculate the value of the US foreign currency reserve as a percentage of China's. *[1 mark]*

b Explain **one** possible reason why China's foreign currency reserves are much larger than those of the USA. *[2 marks]*

Chapter summary

- Exchange rates are the value of one currency in terms of another, or a basket of other currencies.
- A floating exchange rate is where the value of a currency is determined by the demand for and supply of the currency on the foreign exchange market. There should be no deliberate attempts by a government to manipulate the value of a currency in a freely floating exchange rate system.
- The value of a currency increasing in a floating exchange rate system is referred to as *appreciation*. The value of a currency decreasing in a floating exchange rate system is referred to as a *depreciation*.
- When the value of one currency appreciates against another, the value of the other currency must depreciate. For example, if the US dollar is worth more than euros, then the euro must be worth fewer dollars.
- Factors that can lead to an increase in the value of a currency (appreciation) include: foreign demand for exports, inward foreign direct investment, inward portfolio investment, remittances, speculation (buying), relatively lower inflation rates, higher interest rates, relative economic growth rates and central bank intervention.
- Factors that can lead to a depreciation of the currency include: domestic demand for imports, outward foreign direct investment, outward portfolio investment, the remittance of wages from the economy to another economy, speculation (selling), higher levels of inflation, lower interest rates, lower economic growth and central bank intervention.
- Fluctuations in the value of an economy's currency can lead to various forms of inflation, disinflation and, in extreme circumstances, deflation.
- An appreciation of a currency will lead to an increase in the price of exports and a decrease in the price of imports. The result of this may be a decrease in net exports (X – M).
- An appreciation of a currency may lead to a significant reduction in the demand for exports, as they are now more expensive in foreign economies. This reduction in export revenues may result in firms reducing their derived demand for labour.
- If the PED for exports and imports is greater than 1, the result of a currency appreciation will be a reduction in the value of the current account (that is, a worsening, a movement towards a deficit), and vice versa if PED is less than 1.
- A strong currency will bring benefits to consumers of imported goods as they will need to exchange less of their currency to purchase the imported items.
- A fixed exchange rate system exists when the central bank (or monetary authority) buys and sells foreign currencies to ensure the value of its currency stays at a predetermined rate.
- Devaluation occurs when the price of a currency operating in a fixed exchange rate system is officially and deliberately lowered. This causes exports to become relatively cheaper and imports to become more expensive as domestic consumers must exchange more of their currency to buy the imported goods. The opposite applies for revaluation of a currency.
- A managed exchange rate system is where the government or central monetary authority intervenes periodically in the foreign exchange market to influence the exchange rate.
- Managed exchange rates can lead to overvalued and undervalued currencies. As a result of an overvalued currency, imported goods become cheaper so there is downward pressure on the rate of inflation. An undervalued currency occurs when the value of a currency is below its equilibrium value in the long run. Exports become cheaper, leading to growth and employment in export industries.
- Arguments in favour of using a fixed exchange rate system are certainty and stability. Arguments against this are issues related to opportunity costs, currency liquidity, the inability to freely use monetary policy and speculation in foreign exchange markets.

32 Balance of payments

Real-world issue

Who are the winners and losers of the integration of the world's economies?

Conceptual understandings (WISE ChoICES)

- The increased **interdependence** of economies has benefits and costs.
- Increased economic integration may result in **efficiency**, welfare gains and improvements in **economic well-being** but the benefits may not result in **equity**.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- the balance of payments, and explain credit and debit items, and surplus or deficit on an account (AO1)
- elements of the balance of payments from a set of data (AO4)
- the components of the balance of payments: current account, capital account and financial account (AO2)
- the interdependence between the accounts: zero balance in the balance of payments, credits matched by debits and deficits matched by surpluses (AO2)
- the relationship between the current account and the exchange rate (AO2) (HL only)
- an exchange rate diagram to show the relationship between the current account balance and the exchange rate (AO4) (HL only)
- the relationship between the financial account and the exchange rate (AO2) (HL only)
- the implications of a persistent current account deficit (AO3) (HL only)
- methods to correct a persistent current account deficit: expenditure switching, expenditure reducing and supply-side policies (AO2) (HL only)
- the effectiveness of measures to correct a persistent current account deficit (AO3) (HL only)
- the Marshall-Lerner condition and the J-curve effect (AO2) (HL only)
- a J-curve diagram with reference to the Marshall-Lerner condition (AO4) (HL only)
- the implications of a persistent current account surplus (AO3) (HL only).

The balance of payments (AO1, AO4)

The **balance of payments** (BOP) is a financial record of a country's transactions with the rest of the world for a given time period, usually over one year. This includes the country's trade in goods and services with other countries. The government and economists have a keen interest in the economy's international transactions as these can have a major impact on domestic consumption and investment, exchange rates, inflation, export competitiveness, employment and economic growth.

◆ The **balance of payments** is a financial record of a country's transactions with the rest of the world, usually over one year. It includes the country's trade in goods and services with other countries.

Credit and debit items

In the broadest of terms, any money entering an economy can be referred to as an inflow or credit to the country's balance of payments, whereas an outflow or debit refers to any flow of money leaving the economy. For example, export sales are recorded as credit items on the balance of

payments, whereas import expenditures are recorded as debit items. **Credit items** are payments received from consumers, firms and institutions or governments located outside of the economy. **Debit items** are payments made to consumers, firms and institutions or governments located outside of the economy.

There are similarities between the balance of payments and a personal bank account. When money is deposited into a bank account, it is said to be 'credited' to that account. The balance of the account then increases. When money is withdrawn from the account, it is said to be debited, and the balance of the account decreases.

Credit items on the balance of payments (BOP) include revenues earned from the export of goods and services, foreign direct investment and capital transfers. For example, the expenditure of overseas tourists visiting Bangkok would be recorded as a credit item on Thailand's balance of payments.

◆ **Credit items** are payments received from consumers, firms and institutions or governments located outside of the economy.

◆ **Debit items** are payments made to consumers, firms and institutions or governments located outside of the economy.

Figure 32.1 The balance of payments records a country's transactions with the rest of the world

Figure 32.2 Bangkok, Thailand, is one of the world's most visited cities

Debit items on the BOP include the purchase of imports of foreign goods and services, income transfers overseas and the repatriation of profits from multinational companies in the economy. These transactions have a corresponding outflow of money from the domestic economy. Examples of credit and debit items for the UK economy are outlined in Table 32.1.

Table 32.1 Examples of credit and debit items for the UK economy

Credit (inflows or export earnings)	Debit (outflows or import expenditures)
• The purchase of UK produced chemical products by German companies • US tourists flying to the UK on British Airways • Global sales of *Harry Potter* books by British author JK Rowling • French tourists buy theatre tickets to see *Les Misérables* in London	• The UK government maintains foreign embassies located overseas • UK supermarkets purchase Australian wine and meat products • A large British firm buys a fleet of lorries (trucks) from South Korea • The UK government pays interest on its borrowing (debts) from overseas banks

Surplus or deficit on an account

Like with a personal bank account, when the total value of deposits (or credit items) over a period of time exceeds the total value of withdrawals (or debit items), the account is said to have a positive balance. In economics, this is referred to as a **surplus** on the account. By contrast, when the total value of withdrawals in a personal bank account exceeds the total amount of deposits, that account will have a negative balance. The account holder will therefore be in debt. In economics, this imbalance is referred to as a **deficit** on the account. The same logic applies to the balance of payments. So, for example, if the sum of all credit items exceeds the total value of debit items on the balance of trade in goods and services (that is, the value of exports exceeds the value of imports).

◆ A **surplus** on an account exists when the total value of credit items exceeds the total value of debit items, over a given period of time.

◆ A **deficit** on an account exists when the total value of debit items exceeds the total value of credit items, over a given period of time.

Components of the balance of payments (AO2)

The records of all international transactions for an economy, even in just one day, are enormous. It therefore makes sense for the government to separate these into separate sections or components called **accounts**. It helps to imagine the balance of payments as a book of transactions and the various accounts as chapters of that book, which allow the reader to find information easily. The balance of payments consists of three accounts: the current account, the capital account and the financial account. The structure of the balance of payments accounts is shown in Table 32.2.

Table 32.2 The structure of the balance of payments

1 Current account
• Balance of trade in goods • Balance of trade in services • Income • Current transfers
2 Capital account
• Capital transfers • Transactions in non-produced, non-financial assets
3 Financial account
• Foreign direct investment (FDI) • Portfolio investment • Reserve assets • Official borrowing

Source: IB Economics Guide, *page 49*

Top tip!

The structure of the balance of payments may vary from one country to another. However, the version in Table 32.2 is used for the IB Economics curriculum and assessment.

Current account

The **current account** is a record of trade flows, income flows and income transfers across borders. It records all exports and imports of goods and services, plus net investment income from overseas assets, and the net balance of transfers made between countries by individuals, firms and governments. The current account is typically reported quarterly and yearly. It is the largest account in the balance of payments, and as such is likely to be the account most often referenced in the new media.

There are four components of the current account: balance of trade in goods, balance of trade in services, income and current transfers.

◆ The **current account** is a record of all trade flows (exports and imports of goods and services), income flows and income transfers between countries by individuals, firms and governments.

Balance of trade in goods

This records all exports and imports of physical goods (such as rice, coffee, oil, computers and motor vehicles) between a country and the rest of the world. These are sometimes referred to as *visibles* as they can be physically seen and checked at border crossings.

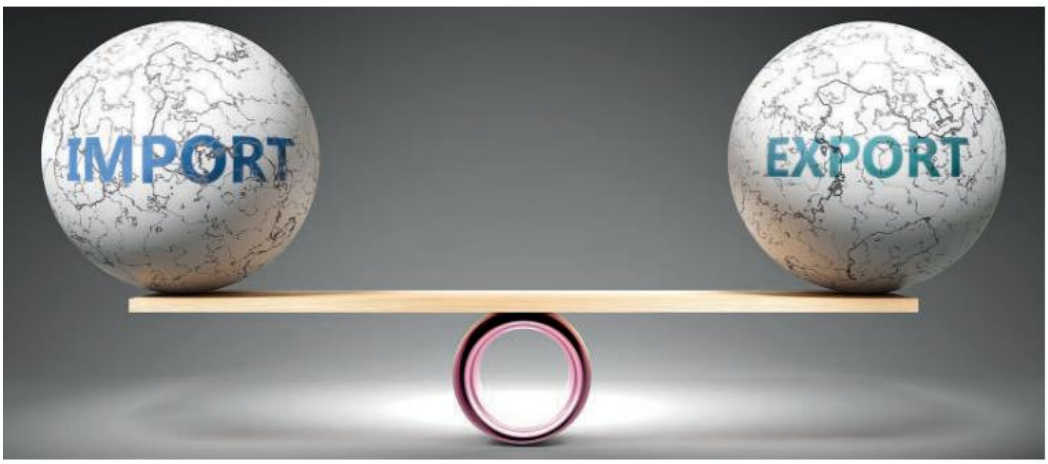


Figure 32.3 The balance of trade is the difference between the trade in exports and imports

Balance of trade in services

This records all exports and imports of services (such as insurance, banking, management consultancy and tourism) between a country and the rest of the world. These are sometimes referred to as *invisibles* as they cannot be physically seen or checked at border crossings. If a foreign tourist makes a payment for a taxi fare in Paris or purchases a ticket to visit Disneyland Paris, then these expenditures cannot be seen or checked by the authorities at border control.

The **balance of trade** (or the **trade balance**) is then the difference between the value of exported goods and services and imported goods and services. The balance of trade is typically the largest component of the current account. The balance of trade is commonly referred to as **net exports** or abbreviated to (X – M).

◆ The **balance of trade** (or **trade balance**) is the difference between a country's total export earnings and its total import expenditure on both goods and services, that is, the value of (X – M).

PAPER 2 AND 3 EXAM PRACTICE QUESTION 32.1

Study the data below and answer the question that follows.

Balance of payments for Country K ($ billion)

Exports	**85**
Goods 57	
Services 28	
Imports	**i....**
Goods 88	
Services 15	
Balance of trade in goods	ii....
Balance of trade in services	iii....
Trade balance	**iv....**

Calculate the missing figures in the data above. *[4 marks]*

◆ **Current transfers** are the inflows and outflows of money that are not made in exchange for trade or any corresponding output of goods or services.

Income

This section of the current account records the *income receipts* (inflows or credit items) earned from foreign investments minus the *income payments* (outflows or debit items) of factor incomes paid to foreign investors. Income consists of the inflows and outflows of payments to the factors of production, that is, the sum of wages, interest, rent and profit (WIRP). Examples can be residents of an economy who receive incomes from abroad, either for employment purposes or returns on various investments of overseas assets. Hence, the value of net income from abroad is income earned from abroad minus income paid to those abroad.

Current transfers

Current transfers are the inflows and outflows of money that are not made in exchange for trade or any corresponding output. Examples include Official Development Assistance (ODA), foreign aid, government grants, concessionary loans and donations between a country and the rest of the world. It also includes foreign workers' net remittances. Remittances are the movement of money when nationals working abroad (expatriate workers) send money back to their home country (see Chapter 31). Hence, the value of net current transfers is current transfers from abroad minus current transfers sent abroad.

The **current account balance** is therefore the sum total of all items listed in the current account, that is, the sum of the current account = Net exports of goods and services + Net income + Net current transfers. A **current account deficit** occurs when the sum of money flowing out of a country's current account exceeds the money flowing into the country's current account. A current account deficit can occur due to exports being relatively more expensive to foreign buyers. This can be caused by higher labour costs in the domestic economy and a higher exchange rate making exports more expensive (see Chapter 31). Domestic buyers also tend to buy more imports if they are relatively cheaper and/or of better quality. This can also lead to a current account deficit.

Current account formulae

- Net exports of goods and services = Exports of goods and services – Imports of goods and services

Or

- Net exports of goods and services = Balance of trade in goods and services = (X – M)
- Net income from abroad = Income from abroad – Income paid abroad
- Net current transfers = Current transfers from abroad – Current transfers sent abroad
- Current account balance = Sum of net exports of goods and services + Net income + Net current transfers

◆ The **current account balance** is the sum total of all items listed in the current account. The balance can be in deficit (when M > X), in surplus (X > M) or zero (X = M).

◆ A **current account deficit** occurs when the sum of money flowing out of a country's current account exceeds the money flowing into its current account, per time period.

By contrast, a **current account surplus** occurs when a country has a positive net balance on its current account. For example, a lower exchange makes it cheaper for foreign buyers to purchase the country's exports of goods and services.

The current account is said to be balanced when the sum of inflows (credit items) matches the sum of outflows (debit items), that is, the value of exports is equal to the sum of imports (X = M). This means the value of net exports is zero.

◆ A **current account surplus** occurs when the sum of money flowing into a country's current account exceeds the money flowing out of its current account, per time period.

Top tip!

Remember the GIST of the current account, that is, that it is made up of four components:

- **G**oods – the trade in physical goods
- **I**ncome – investment income from overseas assets
- **S**ervices – the trade in services
- **T**ransfers – current transfers (of private individuals and government) between countries.

Common mistake

Students frequently comment that high-income countries tend to have a current account surplus. Note that having a current account deficit or surplus is not necessarily a good indicator of how well an economy is performing. Economic powers like Germany, China, Japan, South Korea and the Netherlands usually have current account surpluses, whereas the United States, the United Kingdom and India usually have current account deficits.

Capital account

The **capital account** is a relatively small section of the balance of payments. It records the different forms of capital inflows and outflows of a country during a given time period. Examples include foreign currency flows, debt forgiveness or debt relief (see Chapter 36), and the buying and selling of non-produced, non-financial assets such as copyrights, trademarks and patents. This second component of the BOP is very small in comparison to the values of the current account (and the financial account). The capital account is made up of two sections: **capital transfers** and transactions in non-produced, non-financial assets.

◆ The **capital account** records the different forms of capital inflows and outflows of a country during a given time period, namely capital transfers and transactions in non-produced, non-financial assets.

Capital transfers

Capital transfers are the different forms of capital inflows and outflows of a country. Examples of transactions that could be seen in the capital transfers section of the capital account are payments associated with debt-forgiveness, project-aid for developmental purposes and sums of money/assets brought into an economy by immigrants (an inflow or credit item) or those leaving with emigrants (an outflow or debit item).

◆ **Capital transfers** are the different forms of capital inflows and outflows of a country, such as debt aid, debt forgiveness and the flow of money by emigrants and immigrants.

Transactions in non-produced, non-financial assets

Transactions in non-produced, non-financial assets are the legal property rights to natural resources (such as land rights, mineral rights and fishing rights) and intellectual property rights to intangible assets (such as trademarks, copyrights, brands and patents). These property rights are likely to produce an income stream for the country, so are included in its financial account. In the case of fishing rights, for example, fishing firms make payment to a foreign government to purchase the right to fish in their waters.

Hence, the capital account balance (or net capital transfers) = Capital transfers received from abroad – Capital transfers sent abroad.

Financial account

The third and final component of the balance of payments is the **financial account**, which is significantly larger than the capital account but generally smaller than the current account. The transactions that are recorded in the financial account relate to the change in ownership of assets, that is, cross-border investments. The financial account comprises four sections: foreign direct investment (FDI), portfolio investment, reserve assets and official borrowing.

Foreign direct investment (FDI)

Foreign direct investment (FDI) is spending by multinational corporations (MNCs) in economies that they are not headquartered in. Examples of such expenditure include MNCs setting up production and distribution facilities in overseas locations, or existing MNCs expanding their operations in foreign markets by purchasing more machinery, buildings, capital and other fixed assets. Net FDI is the difference between FDI inflows and FDI outflows.

Portfolio investment

Portfolio investment refers to the stock of investment assets, which can include equities (stocks and shares), government debt (bonds, treasury bills and securities) and debentures (corporate debts). It mainly consists of investors purchasing ownership of a company (stocks or shares) and the purchasing of government bonds (which is technically a purchase of debt). Note that portfolio investments are not direct investments as the investors do not generally have a lasting interest in the companies or governments but are speculating that such expenditure will generate positive returns in the future. Hence, these assets only represent lending and borrowing between households, firms and governments in international markets. Net portfolio investments are calculated by sales of domestic portfolio assets minus purchase of overseas portfolio assets.

Reserve assets

Reserve assets are stocks of foreign currencies and liquid assets such as gold reserves held by central banks used to balance international transactions and payments. A reserve asset must be readily available to the central or monetary authority, must be a physical asset and must be easily transferable or convertible. They are typically used to meet balance of payments financing needs or to influence the exchange rate.

Figure 32.4 Government bonds are a form of debt that are recorded in the financial account

◆ **Transactions in non-produced, non-financial assets** refer to the rights to natural resources and intellectual property rights. The income stream from these intangible assets for the country are recorded in its financial account.

◆ The **financial account** is a record of the transactions that relate to the change in ownership of assets, that is, cross-border investments. These include foreign direct investment, portfolio investment, reserve assets and official borrowing.

CASE STUDY

Foreign currency reserves

A central bank chooses to hold currency reserves in order to manage its own currency's value in the foreign exchange market. In 2020, seven countries in Asia featured in the top 10 countries with the largest amount of foreign currency reserves: China, Japan, Russia, Taiwan, Hong Kong, India and South Korea. The non-Asian countries in the top 10 list were Switzerland, Saudi Arabia and Brazil.

Figure 32.5 Gold is commonly held as part of a country's reserve assets

Official borrowing

Official borrowing refers to government borrowing rather than borrowing of private individuals, households and firms. Governments often borrow from creditors or institutions outside of their own economy. Examples of this can be borrowing from foreign banks and financial institutions or loans from the International Monetary Fund (IMF) for development purposes. In the short run, these loans would represent a credit to the financial account (as there is an inflow of money). However, as the repayments are made, along with interest charges, these would then represent a debit on the financial account. As the loan is likely to be on non-concessional terms, the final value of the loan will exceed what was originally borrowed. Governments of economically least developed economies often need to borrow from the IMF to finance their current account deficits (see Chapter 36).

WISE ChoICES

To what extent do the concepts of **change** and **interdependence** cause current account disequilibria?

Interdependence between the accounts (AO2)

The relationship between the three accounts in the balance of payments relates to an accountancy technique called 'double-entry'. The complexities of this are not studied in IB Economics, but this method of accounting means that for every transaction with an economy and its trading partners, a value will be credited, and a value will be debited. This means that an outflow of a value will be offset by an inflow of value. This is similar to individuals having their bank accounts debited when they purchase an item from an online retailer, and the business being credited the same amount of money in its corporate bank account.

Essentially, the overall balance of payments (BOP) must always balance because in the long-term a country can spend only what it earns, where the sum of all credit items equals the sum of all debit items. Therefore, it is possible to run a deficit in one component of the BOP as this will be

'balanced' by having a surplus on another account or component of the BOP. Hence, in theory, three interrelated conditions must hold:

- zero balance in the overall balance of payments
- credit items are matched by debit items
- deficits are matched by surpluses.

For example, owing to its geographical location, size and factor endowments, Hong Kong typically reports a deficit on its current account but funds this by having a surplus on its financial account owing to favourable conditions for FDI. In other words, Hong Kong must have a surplus in the financial account to provide the economy with the foreign exchange needed to pay for its excess of imports over exports.

By contrast, China, Norway, Germany and Saudi Arabia typically report a current account surplus. These countries can run a deficit on their capital account and/or financial account by directly investing the surplus from the current account in foreign countries and/or accumulate foreign currency reserves.

This means that the current account, capital account and the financial account are interdependent. A country with a current account deficit consumes more than it produces so has to pay for this extra output somehow, that is, through a surplus on its financial account (and vice versa). We can therefore demonstrate this relationship as:

Current account = Capital account + Financial account

However, in reality the balance of payments will not necessarily balance (have a zero balance) because there are just too many transactions to account for in such a precise way. To resolve this calculation problem, **errors and omissions** are used to represent statistical discrepancies when compiling the accounts. Therefore, the following condition must hold for the BOP to balance:

Current account = Capital account + Financial account + Errors and omissions

PAPER 2 AND 3 EXAM PRACTICE QUESTION 32.2

Study the data below for Country L and answer the questions that follow.

Trade in goods	– $18.3bn
Trade in services	+ $21.8bn
Net income	+ $6.7bn
Net current transfers	– $5.6bn

a Define the term *net income*. *[2 marks]*

b Calculate the value of the balance of trade for Country L. *[2 marks]*

c Calculate the value of the current account for Country L. *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 32.3

The data below show data from Country J's balance of payments.

Item	$ billions	Item	$ billions
Exports of goods	235	Net current transfers	–30
Exports of services	320	Net direct investment	65
Imports of goods	–440	Net portfolio investment	38
Imports of services	–235	Capital transfers	26
Net income	20	Trade in non-produced, non-financial assets	20

a Calculate the value of Country J's current account balance. *[2 marks]*

b Calculate the value of the financial account for Country J. *[2 marks]*

c Calculate the value of the capital account for Country J. *[2 marks]*

d Calculate the value of errors and omissions on the balance of payments for Country J. *[2 marks]*

ATL ACTIVITY 32.1

Research the balance of payments position for two countries of your choice (preferably one with a BOP surplus and the other with a deficit). Investigate the causes and consequences for these economies, and evaluate the extent to which you feel the position these countries are in should be a matter of concern.

Relationship between the current account and the exchange rate (AO2, AO4) (HL only)

Another way of explaining why the balance of payments must balance (equal 0) is with a freely floating exchange rate diagram. In Figure 32.6, Australia and the UK trade with each other, so whenever there is a supply of Australian dollars, this would be for the sole purpose of buying UK pounds (£). The supply of Australian dollars would therefore represent a debit to the Australian economy and a credit to the UK economy. When there is a demand for Australian dollars, this would have to be the result of an increase in the supply of UK pounds. This would mean that the demand for Australian dollars would represent credits to the Australian economy and a debit to the UK economy.

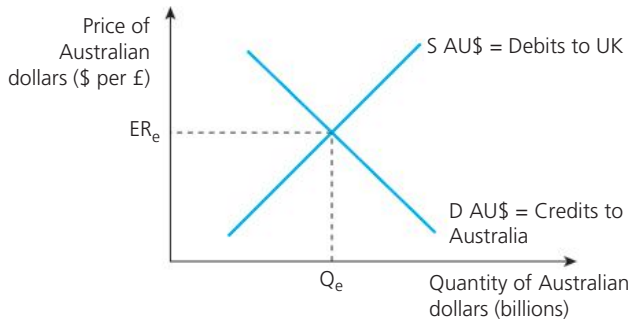


Figure 32.6 Equilibrium position in a freely floating exchange rate system

As a freely floating exchange rate market should always find equilibrium in the long run (at ER_e and Q_e), this means that debits and credits in both the British and Australian balance of payments will be equal. This means that the BOP should be equal to zero. Note that this is conditional on a freely floating exchange rate (see Chapter 31). Should a currency be overvalued, the quantity demanded will exceed the quantity supplied resulting in more debits than credits and a resulting balance of payments deficit. The opposite is true should the currency be undervalued.

When looking at the current account specifically, it can be seen that the value of an exchange rate will have a considerable impact on the demand for imports and exports. A current account deficit exists when the sum of the outflows from the current account exceeds the inflows into the account. This means the total import expenditure on goods and services is greater than total export earnings for the country. Hence, a current account deficit shows the country is spending more than it is earning, so it must require more foreign exchange while it faces relatively lower demand for its own currency as exports are in lower demand.

In theory, the deficit is automatically resolved under a freely floating exchange rate system because it makes exports relatively cheaper and imports relatively more expensive. Export earnings should therefore rise, and imports expenditure should fall, thereby restoring equilibrium.

However, this is more problematic in a fixed exchange rate system unless a devaluation of the currency is approved. If so, this should improve the current account in the same way as with a depreciation of the currency in a freely floating exchange rate system.

In a managed exchange rate system, a current account deficit is resolved by combining market forces and central bank intervention by buying and selling of foreign currencies to reduce the value of the domestic currency in overseas markets.

Given that changes in the value of a currency are likely to change the value of export earnings and the value of import expenditures, it can be said that the balance of trade in the current account has an inverse relationship with the value of the domestic currency. This means that a lower exchange rate causes a rise in the price of imports and a fall in the price of exports. Assuming that this causes a greater than proportionate change in the quantity demanded for exports and imports, the value of exports will increase, and the value of imports will decrease.

By contrast, a higher exchange rate causes a fall in the price of imports and an increase in the price of exports. Assuming that this causes a greater than proportionate change in the quantity demanded for exports and imports, the value of exports will decrease, and the value of imports will increase (see the Marshall-Lerner condition and the J-curve effect covered later in this chapter).

Relationship between the financial account and the exchange rate (AO2) (HL only)

Recall that the financial account handles mainly investments (FDI and portfolio investment) but also includes reserve assets and official borrowing. This means that a surplus in the financial account is likely to be due to credit items from investments being made in the domestic economy by those outside of the economy. To make these investments, investors will need to demand the domestic currency, thus causing an increase in the demand for that currency and therefore an increase in the value of a currency, ceteris paribus.

In the case of a financial account surplus, this would likely be represented by domestic investors investing in portfolio investments abroad. This would mean that these investors would need to supply their own currency to purchase the currency of the foreign country that they are investing in. All things being equal, this would result in a fall in the value of the domestic currency.

When portfolio investments, such as stocks and bonds, yield dividends for their investors, this would represent a debit in the current account as the money leaves the economy. This needs to be paid to investors in their foreign currencies, thus increasing the supply of the local currency, which would cause a currency depreciation, ceteris paribus. The opposite is true for domestic investors who receive dividend payments from holding foreign assets.

Implications of a persistent current account deficit (AO3) (HL only)

The implications of a persistent current account deficit can be examined in terms of exchange rates, interest rates, foreign ownership of domestic assets, debt, credit ratings, demand management and economic growth.

Exchange rates

It is important to remember that whenever a good is imported, there has been a demand for a foreign currency to pay for the product by supplying the domestic currency. Therefore, persistent current account deficits will lead to a weaker domestic currency. For countries reliant on key imports such as oil, a declining currency can fuel imported inflation (see Chapter 20), which increases the costs of production and erodes consumer surplus. There may subsequently be negative impacts on national output and employment.

Figure 32.7 A persistent current account deficit will tend to cause the exchange rate to fall

Interest rates

As a persistent current account deficit puts downward pressure on the exchange rate, governments may be tempted to raise interest rates to attract foreign currency and capital inflows. Investors looking to save in foreign bank accounts will be attracted by the higher interest rates. However, higher interest rates can have a contractionary impact on aggregate demand in the local economy. If this is prolonged, it could even lead to a recession with detrimental impacts on the economy, including higher unemployment.

Foreign ownership of domestic assets

A current account deficit has to be financed by a surplus on the financial account, often in the form of foreign direct investment (FDI). However, this may mean that more domestic assets become owned by foreign multinational corporations. There are opposing implications of this as FDI can help to boost aggregate demand in the economy, thus increase economic growth. However, FDI and foreign ownership of domestic assets can also cause leakages in the circular flow of income as MNCs remit income and profits back to their home countries. There is also no guarantee that these foreign MNCs will remain located in the domestic economy as they may choose to relocate for cost advantages in other countries. Hence, foreign ownership of domestic assets could leave the local economy in a vulnerable position.

Debt

If a country does not have sufficient financial reserves to fund its persistent current account deficit, it will have to borrow money to pay for this. One way of financing this current account deficit is by taking loans from the International Monetary Fund (IMF). This would represent a credit to the financial account, but the country therefore goes into debt with its related problems, such as mounting interest repayments, lower economic growth and job losses.

Credit ratings

Credit ratings refer to the creditworthiness of a borrower based on the likely ability to repay the debt. A persistent current account deficit tends to reduce a country's credit rating as it can signal underlying structural economic problems for the country. These economies may need to restructure their debt by agreeing new terms, such as longer repayment periods, or negotiate lower interest rates. However, these measures will make such countries much less attractive to foreign lenders in the future. This could mean that loans they receive in the future are at a much higher interest rate, if they manage to secure any further loans at all.

Demand management

An economy often experiences a current account deficit when the economy is growing rapidly, as those with higher incomes purchase more imports. Contractionary demand management policies (see Chapters 24 and 25) might be used to reduce the demand for imports to correct this persistent current account deficit. These policies tend to focus on incentivizing households to purchase domestically produced goods or generic attempts at lowering expenditure on all goods and services (and therefore also targeting imported goods and services). However, this can negatively affect growth and employment in the economy.

Economic growth

A common theme of the previous implications of a persistent current account deficit has been the impact on economic growth. For example, in order to close the current account deficit, a government may choose to use contractionary demand-side policies to lower aggregate demand, although this is likely to reduce economic growth. Similarly, a poor credit rating due to high levels of indebtedness (to fund the deficit) will also impact on consumer and producer confidence levels. This will hinder both consumption and investment – key components of aggregate demand – so could lead to unemployment and lower economic growth.

Theory of Knowledge

Should high-income countries have an ethical and moral obligation to support low-income countries suffering from a persistent current account deficit?

◆ **Credit ratings** refer to the creditworthiness of a borrower based on the likely ability to repay the debt. It is what lenders look at before they approve any new credit or loans to the customer.

Methods to correct a persistent current account deficit (AO2) (HL only)

Governments use various policies to correct a persistent deficit on the current account. These policies can be grouped into three categories: expenditure switching policies, expenditure reducing policies and supply-side policies.

Expenditure switching policies

Expenditure switching policies are measures designed to cut a current account deficit by encouraging households and firms to buy domestically produced goods and services rather than imported alternatives by raising the relative price of imports or reducing the relative price of exports. To achieve this, a government might try to engage in:

- **Export promotion** – These are policies that stimulate the demand for exports. The government decides which industries could become export competitive. They then issue subsidies and

◆ **Expenditure switching policies** are measures intended to cut a current account deficit by encouraging households and firms to buy domestically produced goods and services rather than imported alternatives.

direct spending towards these industries to help them develop. However, export promotion involves a large amount of public sector spending, for which there are opportunity costs.

- **Trade protection** – These measures reduce the competiveness of imports thereby making domestically produced products more attractive. Examples of trade protectionist measures include tariffs (import taxes that raise the price of imports) and quotas (quantitative limits on the amount of available imports). However, trade protection could cause trading partners to retaliate by imposing their own restrictions and cause economic inefficiency as governments protect uncompetitive domestic industries. Trade protection is covered in more detail in Chapter 28.
- **Currency devaluation** – The central bank or monetary authority of a country can devalue its exchange rate in order to reduce the price of exports and make imports relatively more expensive. This will lead to an increase in net exports (exports exceeding imports), ceteris paribus. Whether this policy will help to correct the balance of payments deficit depends on the value of the price elasticity of demand for both imports and exports (see the Marshall-Lerner condition covered later in this chapter).

Expenditure reducing policies

Expenditure reducing policies are measures designed to cut a current account deficit by lowering disposable income in order to limit aggregate demand and import expenditure in particular. This is usually achieved through the use of contractionary monetary and fiscal policies:

◆ **Expenditure reducing policies** are measures designed to cut a current account deficit by lowering disposable income to limit aggregate demand and import expenditure in particular.

- **Contractionary monetary policy** – Higher interest rates make new and existing loans more expensive for borrowers, while greater incentives to save will withdraw more money from the circular flow of income. Hence, both households and firms will tend to reduce their demand for imports. Monetary policy is covered in more detail in Chapter 24.
- **Contractionary fiscal policy** – These measures use a combination of higher taxes and reduced government spending in order to reduce consumption, investment and the amount of money available to spend on imports. The effect of this policy will be more significant for economies that have a high marginal propensity to import (see Chapter 25). Fiscal policy is covered in more detail in Chapter 25.

Supply-side policies

Supply-side policies are long-term measures that strive to raise the productive capacity of the economy and to increase the country's international competitiveness. Examples of such policies (which are covered in more detail in Chapter 26) include the following:

- **Investments in education and healthcare** – These measures help to improve the economy's human capital (investments to improve the skills and productivity of the workforce), thereby raising the quality of exports and improving the country's long-term international competitiveness. This should then increase the demand for exports, ceteris paribus. There could also be a resulting reduction in the demand for imported goods because domestic firms become more productive owing to their healthier and more skilled workers.
- **Investment in infrastructure** – This policy measure would help to support businesses and industries, especially those engaged in export-orientated markets, owing to improved transportation and communications networks. This will improve the productivity of domestic firms and make them more competitive against foreign firms. This should then lead to a reduction in the demand for imports in favour of greater demand for exports (for overseas customers) and domestically produced products (for local customers).

- **Measures to encourage export-driven business start-ups and industries** – These policies include the use of government subsidies and tax incentives for targeted businesses. The measures focus on the long-term but will potentially or eventually develop home-grown industries that will rival those found in foreign economies.

The biggest limitation of using supply-side policies to correct a persistent deficit on the BOP is the time involved. While these policies can be extremely expensive to fund, they take a significant amount of time to have any meaningful impact on the economy and, hence, the BOP. There is no guarantee that the above examples will correct the disequilibrium in the current account.

WISE ChoICES

To what extent does **intervention** (to correct a balance of payments disequilibrium) improve **economic well-being** at the opportunity cost of **efficiency**?

Effectiveness of methods to correct a persistent current account deficit (AO3) (HL only)

Expenditure switching policies do not always work because price is not always the deciding factor in determining the level of demand for goods and services. Customers will consider non-price factors that determine the demand for products, such as quality, functionality and brand loyalty. This can apply to any good or service that is being exported or imported, but is particularly applicable in the case of luxury products such as jewellery, designer clothes and accessories, fine wines or high-performance sports cars. The effectiveness of any policies to correct a persistent current account deficit will therefore also depend on the price elasticity of demand (PED) for both exports and imports. This point is discussed further under the theory of the J-curve effect, later in this chapter.

Trade protection can also cause problems, as other countries are likely to retaliate by creating their own trade barriers (see Chapter 29). Such political and economic disharmony in the global economy will reduce the benefits of free international trade and lower economic efficiency. Domestic consumer surplus will be reduced, due to higher prices and reduced choice, which can lead to a reduction in the standard of living, especially in ELDCs.

Attempts by governments to devalue their currency will work only under a system of fixed exchange rates, yet the falling value of the currency can cause imported inflation (higher priced imports of essential products such as crude oil, raw materials and foodstuff). This can result in inflation and a reduction in the standard of living in that economy as wages often will not rise with inflation, therefore reducing the purchasing power of households and firms.

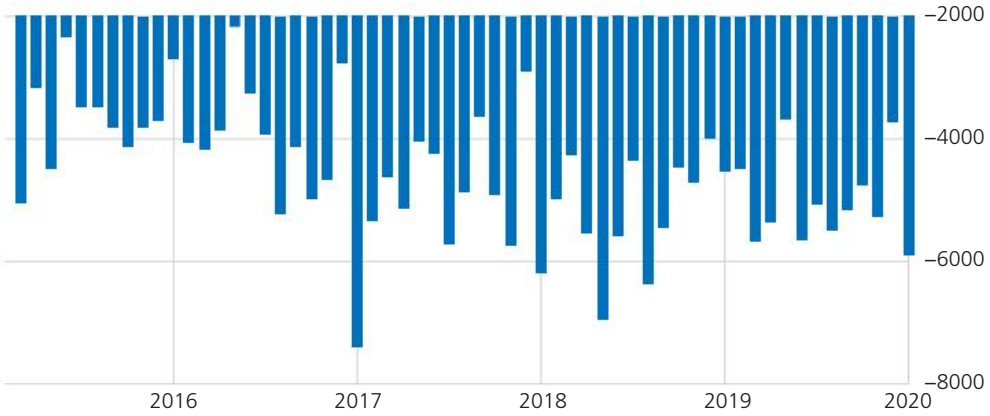
Expenditure reducing policies, such as higher direct taxes, can cause huge disincentive effects in the economy, thereby negatively affecting economic growth and development in the long-term. Contractionary demand-side policies can also cause mass-scale unemployment in the economy. The impact of these policies will depend on the macroeconomic position of the economy at the time that they are implemented. Should the economy be experiencing demand-pull inflation, then the impact of contractionary policies will be less damaging (see Chapters 24 and 25).

While supply-side policies, such as investments in education and healthcare, aim to make domestically produced goods and services more competitive, they tend to take a very long time to materialize and require significant government expenditure. Hence, the implementation of supply-side policies to rectify a persistent current account deficit tends to be very costly to the economy.

Figure 32.8 Importers of American 'muscle cars' prioritize performance over price

PAPER 2 EXAM PRACTICE QUESTION 32.4

France is a major exporter of medicines and a large importer of European, US and Japanese cars. The chart below (Figure 32.9) shows the balance of trade for France from 2016 to 2020.



Source: Trading Economics https://tradingeconomics.com/france/balance-of-trade

Figure 32.9 Balance of trade for France ($bn) 2016-20

a Define the term *balance of trade*. *[2 marks]*

b Explain **two** possible causes of France's persistent balance of trade deficit. *[4 marks]*

The Marshall-Lerner condition and the J-curve effect (AO2, AO4) (HL only)

Under a fixed exchange rate system, a government can devaluate its exchange rate to resolve a current account deficit. In theory, this will make exports relatively more attractive to overseas buyers and make imports more expensive for domestic citizens, thus improving or correcting the current account deficit.

Similarly, a country with a current account deficit and operating under a floating exchange rate system should experience an improvement if there is downward pressure (currency depreciation) on its exchange rate. Assuming customers are responsive to the relative change in prices of exports and imports, the fall in the exchange rate should help to correct the current account deficit.

However, a devaluation or depreciation of the currency will work in rectifying a balance of trade deficit only if the sum of the price elasticity of demand for exports (X) and imports (M) is greater than 1, that is, price elastic. This rule is called the **Marshall-Lerner condition**:

◆ The **Marshall-Lerner condition** states that a devaluation or depreciation of the currency will work in rectifying a balance of trade deficit only if the sum of the price elasticity of demand for exports and imports is greater than 1 (that is, price elastic).

$$PED_x + PED_m > 1$$

Hence, if $PED_x + PED_m = 1$, then a currency devaluation or depreciation has no impact on the current account deficit. Likewise, if the $PED_x + PED_m < 1$, then the fall in the exchange rate will actually worsen the deficit further. This is because the currency is worth less, yet exporters and importers are not relatively responsive to the fall in the exchange rate. The Marshall-Lerner condition was named after classical economist Alfred Marshall (1842–1924) and Abba P Lerner (1903–82). Essentially, the Marshall-Lerner condition is an extension of Alfred Marshall's theory of the price elasticity of demand applied to foreign trade.

The impact or significance of the Marshall-Lerner condition can be illustrated by using a **J-curve diagram** (or **J-curve effect**). The J-curve effect is an economic model that shows the impact on a country's balance of trade following a currency depreciation. Initially, the balance of trade worsens before it gets better (see Figure 32.10). The J-curve effect is caused by time lags, such as:

◆ The **J-curve effect** is an economic model that shows the impact on a country's balance of trade following a currency depreciation, namely that the trade balance worsens before it gets better.

- Limited access to information – individuals, households and firms are not fully aware of the currency devaluation or its relevance to them.
- Ongoing contracts – being bound by existing contracts means that some firms cannot take advantage of the currency devaluation immediately.
- Existing habits and tastes – in the short run, buyers might not be responsive to the currency devaluation so continue to buy imports at a higher price (and export at a lower price). This is particularly the case for customers who have a high degree of customer loyalty for foreign brands.

Owing to the time lags, individuals, households and firms in both domestic and foreign markets need time to adjust to the change in the relative prices of exports and imports. Again, the J-curve will materialize (that is, a fall in the value of the exchange rate rectifying a balance of trade deficit) only if the Marshall-Lerner condition holds.

A lower exchange rate causes a relative rise in the price of imports and a relative fall in the price of exports. This worsens the deficit, as shown by the movement from point A to point B in Figure 32.10, as the PED_x and PED_m remain relatively price inelastic in the short run. In the long run, the PED_x and PED_m will increase as households and firms adjust to the relative price changes of exports and imports, resulting in the lower exchange rate boosting exports and reducing imports over time. This enables the trade balance to move from point B to point C, thereby eliminating the deficit on the balance of trade.

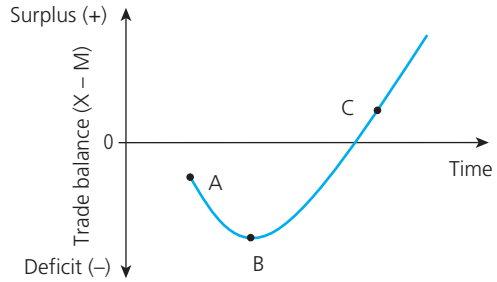


Figure 32.10 The J-curve

Top tip!

The previous analysis can apply to countries with a persistent balance of trade surplus by using an inverse or 'upside down' J-curve. This happens by raising the value of the exchange rate (either through a currency revaluation or currency appreciation). Eventually, if the Marshall-Lerner condition holds, the economy will spend more on imports and earn less from exports, thereby eliminating the balance of trade surplus.

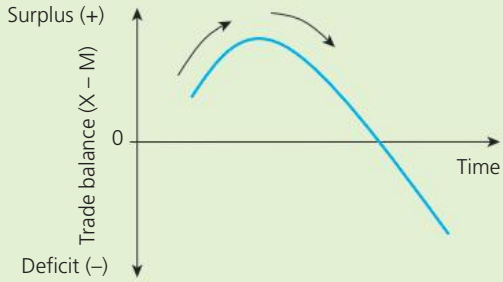


Figure 32.11 The inverse J curve

Implications of a persistent current account surplus (AO3) (HL only)

It may seem at first that a surplus is preferable to a deficit on the current account, given that this can help to offset a deficit on the country's capital or financial accounts. However, a persistent current account surplus is not necessarily or always desirable. There are implications of this on the level of domestic consumption and investment, exchange rates, inflation, employment and export competitiveness.

Domestic consumption and investment

In the short run, a current account surplus means the (X – M) component of aggregate demand is positive, so this helps to boost AD from AD_1 to AD_2, as shown in Figure 32.12. This causes the economy to grow, with real GDP rising from Y_1 to Y_2. As a result, employment will also increase. However, there is likely to be some inflationary pressure on the economy owing to the higher level of AD, causing the general price level to rise from PL_1 to PL_2. Hence, a persistent current account surplus can create inflationary pressures on the economy so can deter future investments in the country.

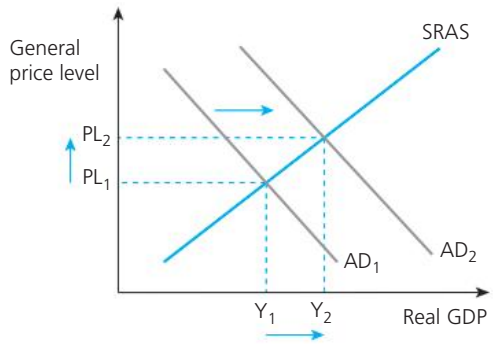


Figure 32.12 The impact of a persistent current account surplus on consumption and investment

Alternatively, a persistent current account surplus can be caused by lower domestic expenditure on imports, but this is not necessarily good for households and firms if it is due to a lower level of income. As the current account is in surplus, it could mean that the additional money coming into the country is not being spent on imported goods and services, which could represent less choice for domestic consumers and therefore a lower standard of living.

Exchange rates

Demand for exports represents demand for the currency of the exporting nation. A current account surplus implies there is greater demand for the domestic currency as export sales exceed import expenditure. This can then cause the domestic currency to appreciate in value. Subsequently, exporters will find it increasingly difficult to sell to foreign buyers who have to pay higher prices for these products. This also means that domestic consumers will find domestically produced goods less competitive to foreign goods, which hinders domestic output and employment opportunities. Furthermore, a higher exchange rate, which results from a persistent current account surplus, can make the economy less attractive as a destination for foreign direct investments.

Inflation

A persistent current account surplus implies there is an increase in net exports for the country, (which is a key component of aggregate demand). A large surplus will result in a significant increase in aggregate demand, leading to demand-pull inflation, as shown in Figure 32.12. This will tend to reduce the international competitiveness of the economy in the long run.

On the other hand, the appreciation of a currency that follows a persistent current account surplus could lead to a reduction in the costs of production for firms that import their raw materials. If the consumer price index (see Chapter 20) includes a significant number of imported goods for the average household in the country, the price of these goods will also fall due to the currency appreciation. This could then lead to lower inflation figures and a lower general cost of living for citizens in the country.

Employment

The export-led growth enjoyed by an economy with a persistent current account surplus may also lead to reductions in unemployment. A common cause of structural unemployment (see Chapter 19) is a change in the geographical location of jobs. An economy that has a large and persistent current account surplus is likely to be the location where the demand for labour has relocated as multinational companies relocate to different economies in search of the best returns on their investments. The demand for labour is derived from the demand for goods and services, therefore when firms experience a sustained increase in their export revenues, this will result in an increase in their demand for labour. This will also reduce cyclical unemployment in the economy.

From the perspective of workers, a sustained current account surplus can be desirable as higher export sales help to create jobs. Major oil exporters such as Saudi Arabia, Kuwait, Qatar and the United Arab Emirates have consistently enjoyed current account surpluses, thus boosting their GDP and standards of living.

However, a consequence of a sustained current account surplus is that job losses are created in other countries. For example, the USA has blamed China for sustaining a large current account surplus thereby causing large-scale unemployment in the USA. After all, a current account surplus in one country is another's deficit.

Export competitiveness

Higher demand for exports and a higher exchange rate can lead to higher export prices. Therefore, a persistent current account surplus can diminish the international competitiveness of the country over time. Imports also become relatively cheaper, so the long run implication of a current account surplus is likely to be a further reduction in the competitiveness of domestic firms.

Domestic firms that rely on imports of raw materials and components can become complacent and over-reliant on favourable exchange rates so do not invest in measures to improve their efficiency and competitiveness. When these conditions disappear owing to a correction of the current account surplus, these firms are left uncompetitive in international markets.

ATL ACTIVITY 32.2

Investigate the causes of a persistent current account deficit for a country of your choice. Evaluate the macroeconomic policies used by the government to correct this deficit.

Top tip!

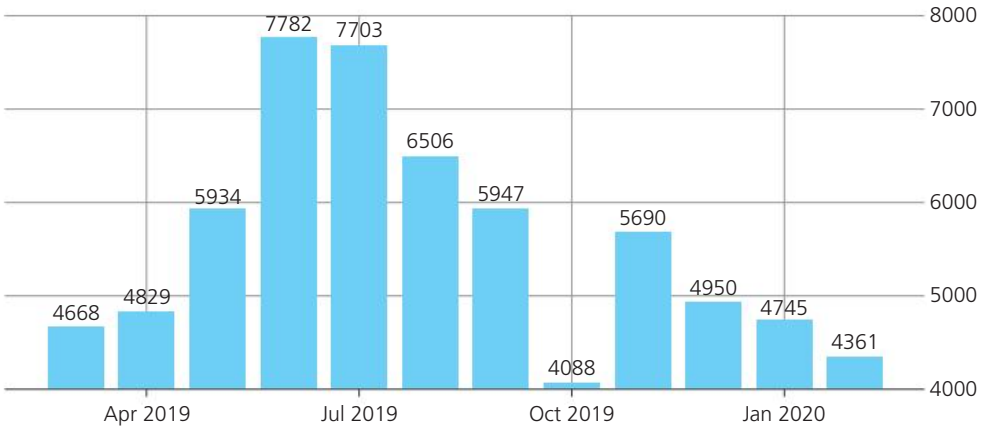
Whether a persistent current account surplus is desirable largely depends on the causes of the surplus. For example, if the surplus is due to increased use of trade protection measures, then there are welfare and efficiency losses to society. If the surplus is due to export-orientated policies, then the country should experience economic growth and more employment opportunities.

Common mistake

Don't assume that a balance of payments disequilibrium refers only to a deficit. While this is not entirely incorrect, disequilibrium also refers to a surplus.

PAPER 3 EXAM PRACTICE QUESTION 32.5 (HL ONLY)

Australia is one of the world's largest net exporters of coal. The chart below (Figure 32.13) shows Australia's balance of trade for the period of March 2019 to February 2020 (figures in millions of Australian dollars).



Source: Trading Economics https://tradingeconomics.com/australia/balance-of-trade

Figure 32.13 Australia's balance of trade (millions of AUD), March 2019–February 2020

a Define the term *current account surplus*. *[2 marks]*

b Explain **two** consequences of Australia's persistent current account surplus throughout the period shown. *[4 marks]*

Chapter summary

- The balance of payments (BOP) is a financial record of a country's transactions with the rest of the world for a given time period.
- Any money entering an economy can be referred to as an inflow or an outflow, or credit items and debit items.
- A surplus on an account occurs when the total value of credit items into an account exceeds the total value of debit items during a given time period.
- A deficit occurs when the total value of debit items into an account exceeds the total value of credit items during a given time period.
- There are four components to the balance of payments: current account, capital account, financial account and errors and omissions. The accounts of the balance of payments are interdependent. They must amount to zero, facilitated by a freely floating exchange rate system.
- The *current account* is a record of a country's exports and imports of goods and services, plus its net investment income from overseas assets, and the net balance of transfers made between countries by individuals, firms and governments.
- The *capital account* records the different forms of capital inflows and outflows of a country during a given time period, such as capital transfers and transaction in non-produced, non-financial assets.
- The *financial account* records transactions related to the change in ownership of assets, consisting of foreign direct investment (FDI), portfolio investment, reserve assets and official borrowing.

- The overall BOP must always balance because in the long run a country can spend only what it earns, that is, the sum of all credit items equals the sum of debit items. Therefore, it is possible to run a deficit in one component of the BOP as this will be 'balanced' by having a surplus on another account.
- It can therefore be said that the Current account = Capital account + Financial account + Errors and omissions.
- The value of an exchange rate will have a considerable impact on the demand for imports and exports. There is an inverse relationship between the value of a domestic currency and the country's BOP.
- Persistent current account deficits will lead to weaker currencies, higher interest rates, further leakages, debt, lower credit ratings, harmful demand-management policies and lower economic growth.
- Methods to correct persistent current account deficits are expenditure switching policies, expenditure reducing policies and supply-side policies. (HL only)
- A currency devaluation is only effective in resolving a balance of trade deficit if the sum of the price elasticity of demand for imports and exports is greater than one (that is, price elastic). This is referred to as the Marshall-Lerner condition. (HL only)
- The J-curve effect is an economic model that illustrates the impact of time lags on the balance of trade following a currency depreciation. It shows the balance of trade worsens before it gets better due to the time lags involved in adjusting to the lower exchange rate. (HL only)
- A persistent current account surplus will tend to lead to appreciating exchange rates, demand-pull inflation, increased growth, lower unemployment and higher export prices in the long run. (HL only)

33 Sustainable development

Real-world issue

Why is economic development uneven?

Conceptual understandings (WISE ChoICES)

- Perceptions of the meanings of development and **equity** change over time and vary across cultures.
- Governments and other economic agents may intervene in an attempt to promote **economic well-being** and **equity** in societies.
- The pursuit of **sustainability** is subject to various constraints.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:
- the meaning of sustainable development (AO2)
- the Sustainable Development Goals (AO2)
- the relationship between sustainability and poverty (AO2) (HL only).

The meaning of sustainable development (AO2)

The United Nations Educational, Scientific and Cultural Organization (UNESCO) states that **sustainable development** is the overarching paradigm of the United Nations. Since the Second World War, the world has generally experienced an unprecedented scale of economic prosperity and well-being. With the help of science, technology, international trade and innovation, among other factors, the world's per capita income was quintupled in the twentieth century.

◆ **Sustainable development** is economic development that meets the needs of the present generations without compromising the ability of future generations to meet their own needs.

Although two world wars and a few financial crises disrupted the continuous progress that was made, the world has seen a sharp rise in standards of living, longevity, access to clean water, access to capital, literacy rates, access to healthcare and education. The numbers of child deaths under one year of age, deaths of women while giving birth and other death-related indicators have decreased considerably with the improvement in medical and healthcare facilities. Cities, roads, ports, airports and communications networks have been built using mostly exhaustible resources. Both economic growth and economic development permeate through all the countries in the world.

However, the benefits have not been shared or distributed evenly across and within countries, and the progress comes with irrevocable costs. As they strive for economic growth and prosperity, countries have used both non-renewable natural resources and renewable resources to achieve economic growth and development.

Figure 33.1 Pollution in Shanghai, China – the opportunity cost of rapid economic growth

- Non-renewable natural resources are not naturally replenished once they have been used. Examples of non-renewable natural resources include carbon-based fossil fuel, minerals and groundwater.
- Renewable resources are naturally replenished if they are used or harvested sustainably. Examples include timber (lumber), animal husbandry, renewable energy sources and food crops.

However, this constant exhaustion of the planet's scarce resources leads to greater scarcity and unavailability of certain resources. With the insatiable desire to grow, countries have been using increasingly more resources to achieve further growth and development. Some countries use economic growth for poverty alleviation and to allow more people to have access to basic necessities. For example, China was able to move approximately 600 million people out of extreme poverty through its radical economic policies since the 1980s. However, by-products of such rapid and sustained economic growth are environmental exploitation, pollution and climate change. There have also been growing income inequalities in China as a result of its high growth rates over the past four decades. The rate of economic growth and development may meet the needs of today's demand, but it could hinder the ability to fulfil the needs of future generations. Hence, economists and policymakers are interested in the pursuit of sustainable development.

ATL ACTIVITY 33.1

In a group of two or three, identify some issues of sustainability in your own country or a country of your choice. Brainstorm how sustainability can be improved and the economic costs of doing so. Present your findings with the rest of your class.

Sustainable development is not a new concept in economics. US marine biologist and conservationist Rachel Louise Carson's seminal work *Silent Spring* (1962) alluded to the idea that economic growth and development may have adverse effects on the environment. In 1968, several heads of state, diplomats, scientists, economists and UN bureaucrats founded the international

group Club of Rome and in 1972 published a report on the negative impacts of growth. Following the Club of Rome report, several academics from the Massachusetts Institute of Technology (MIT) prepared a hearing named 'Growth and its implication for the Future' (1975) for the United States Congress. By the late 1980s, the United Nations World Commission on Environment and Development (UNWCED) published the groundbreaking report *Our Common Future* (1987), commonly referred to as the Brundtland Report.

Perhaps the most widely acceptable definition of sustainable development comes from the Brundtland Report, which states:

> *Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs. It contains within it two key concepts:*
>
> *The concept of 'needs', in particular, the essential needs of the world's poor, to which overriding priority should be given; and*
>
> *The idea of limitations imposed by the state of technology and social organization on the environment's ability to meet present and future needs.*
>
> *Source: World Commission on Environment and Development,* Our Common Future *(1987)*

Table 33.1 Selected timeline leading to declaration of the UN SDGs

Time	Event
17th century	Establishment of seventeenth-century European forest management practices related to timber use and depletion.
1962	Rachel Carson's seminal work *Silent Spring* suggests that economic growth and development may have adverse effects on the environment.
1970s	The Club of Rome publishes its report on the negative impacts of growth. The Organization for Economic Co-operation and Development (OECD) proposes the polluter-pays principle to compensate for the negative costs of pollution.
1987	United Nations World Commission on Environment and Development (UNWCED) publishes the groundbreaking report *Our Common Future*, commonly referred to as the Brundtland Report. The report coins the term 'sustainable development'.
1987	The Montreal Protocol, an international treaty, is adopted to protect the ozone layer by phasing out the production of numerous substances responsible for ozone depletion, such as Chlorofluorocarbons (CFCs).
1992	The UN Conference on Environment and Development (UNCED), or the Earth Summit, is held in Rio de Janeiro. Agreements are reached on the Rio Declaration, a non-legally binding agreement on Forest Principles (recommendations for conservation and sustainable development forestry).
1995	The World Trade Organization (WTO) is formed and the impacts of trade on the environment and sustainable development come into consideration.
1997	International agreement on an emissions reduction target are set in the Kyoto Protocol, an international treaty that extends the 1992 United Nations Framework Convention on Climate Change (UNFCCC). The agreement commits governments to reduce their greenhouse gas emissions.
2000	The United Nations' Millennium Development Goals (MDGs) are finalized.
2012	The United Nations Conference on Sustainable Development (UNCSD) holds its third international conference on sustainable development called Rio 2012 or Earth Summit 2012, aimed at reconciling the economic and environmental goals of the global community. The UN's Sustainable Development Goals are formed.
2015	The Paris Agreement builds upon the UNFCCC, with the specific goal being to fight against climate change by participant countries. The **Sustainable Development Goals** (SDGs) are adopted by all United Nations Member States as the most comprehensive international pact to take necessary actions to end world poverty, to protect the planet and to ensure that everyone enjoys peace and prosperity by the year 2030.

◆ The **Sustainable Development Goals** (SDGs) of the United Nations Development Programme consist of 17 international development targets aimed at achieving peace and prosperity by all UN member countries by 2030.

Since the Brundtland Report in 1987, the whole world converged in acceptance of sustainable development beyond a restorative and environmental framework. The UN Conference on Environment and Development (UNCED), also known by its name the Earth Summit, published the Earth Charter in 1992, which raised the need to construct a 'just, sustainable and peaceful' global society. In 2000, the world leaders under the UN charter identified economic development, social development and environmental protection as three pillars of sustainable development. These events culminated in the United Nations announcing its Sustainable Development Goals (SDGs) in 2015.

There are three core and interrelated pillars or categories of sustainable development. These are social sustainability, economic sustainability and environmental sustainability (see Figure 33.2).

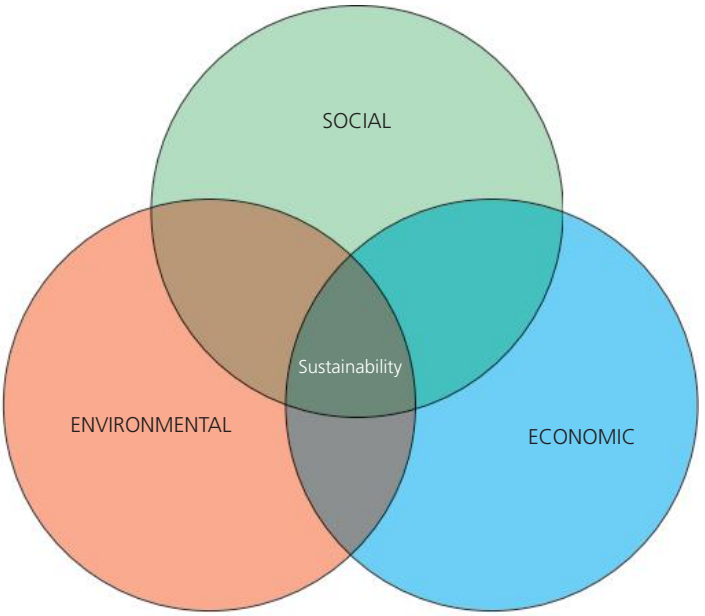


Figure 33.2 The three pillars of sustainability

Social sustainability

Social sustainability is defined as the ability of an economy to develop social processes and structures that enable its current members to live optimally and to support the ability of future communities to do the same. For example, population growth around the world results in greater levels of consumption and depletion of the Earth's natural resources. Social sustainability helps people to achieve a better quality of life through better access to shelter, healthcare and education. Essentially, policies that achieve social sustainability enable people of the current generation to live a respectable life without hindering the quality of life of future generations.

◆ **Social sustainability** refers to the ability of an economy to develop social processes and structures that enable its current population to live optimally and to support the ability of future populations to do the same.

Internationally renowned welfare economist Amartya Sen (b.1933) suggests the following for achieving social sustainability:

- Society should offer equitable opportunities for all members of the community, especially the most vulnerable members, like the poor (including low-income households), women and children.
- Communities should provide ways that support processes, systems and structures that deliver interconnectedness within and beyond the community. Economic theory suggests that an open economy and one that is conducive to trade is more likely to achieve sustainable growth and development for its citizens.
- Society should ensure that the basic needs of its members are met to aid a better quality of living through the provision of a good healthcare system, stable employment opportunities, social housing and universal education.
- All the members of the community should be able to have a voice and raise their concerns in order to facilitate a better functioning and impartial government.
- The society should promote and support diversity, to bind the community for improved prosperity and sustained well-being.

Economic sustainability

Economic sustainability refers to the optimal use of scarce resources in such a way to ensure future generations are not disadvantaged in favour of today's generation. It involves safeguarding

◆ **Economic sustainability** refers to the optimal use of scarce resources in such a way to ensure future generations are not disadvantaged in favour of today's generation.

and maintaining economic resources in a sustainable way, such as by recycling, reducing and reusing resources in the production process.

Economic sustainability is a key pillar of overall sustainability. It enables sustained and equitable growth that leads to economic development through poverty alleviation, universal access to education and skills development, a more equitable distribution of income and wealth, and widespread access to credit and finance. Economic sustainability also requires a system of good governance that promotes economic well-being for all. This can be achieved by enabling the economy to grow in such a way so that the social, environmental and cultural aspects of the community are not compromised. Essentially, this means that economic activities should promote long-term benefits for all without harming the environment and the sociocultural aspects of the community. This highlights the interrelatedness of the three pillars of sustainability.

There are a few important benefits of achieving economic sustainability:

- Economic sustainability attracts domestic and foreign investment as investors and entrepreneurs see this as an opportunity to increase long-term economic prosperity.
- Local communities can function more efficiently and productively as well as more innovatively.
- Firms can enjoy lower costs of production owing to reductions in environmental costs. Without being economically sustainable, environmental costs such as pollution and damage to ecosystems are likely to grow exponentially thereby hindering economic activity in the future.
- A stable economy is a prerequisite to sustained social welfare. Economic suitability helps households to maintain a stable level of income, rather than facing the challenges of volatility and uncertainty.

Environmental sustainability

Environmental sustainability refers to the responsible use of the planet's natural resources so that future generations are not compromised on their use of these natural resources. This may

◆ **Environmental sustainability** refers to the responsible use of the planet's natural resources so that future generations are not compromised on their use of these natural resources.

Figure 33.3 Waste management systems are an essential aspect of environmental sustainability

involve better use of renewable resources, waste management and environmental protection. Environmental sustainability ensures that nature is not exploited as a non-renewable resource in the pursuit of economic growth and development.

CASE STUDY

The economic costs of deforestation

According to Natural Resources Canada (**www.nrcan.gc.ca**), Brazil's continual reduction of forest areas is harmful for the country's long-term environmental sustainability. By contrast, Bhutan has continually increased its forest area. NRC states that the ratio of forest areas to the total land area of the country is a key indicator of environmental sustainability. NRC, a Canadian government agency, states that forestry helps to sustain ecological functions such as carbon storage, nutrient cycling, water and air purification, and maintenance of wildlife.

Figure 33.4 Deforestation of eucalyptus forests in Minas Gerais, Brazil

Examples of environmental sustainability issues include:

- Encouraging the use of renewable energy resources, such as wind, solar or biomass thereby reducing the reliance on production of energy from non-renewable fossil fuels.
- Global warming and climate change have caused a gradual increase in the overall temperature of the planet. This is one of the biggest threats to environmental sustainability and is a major challenge faced by the entire world.
- Environmental sustainability helps our ecosystems to be healthy and productive in the future to meet the needs of future generations. Protection of ecosystems also helps to ensure sufficient agricultural food supplies, fishing stocks and animal husbandry for generations to come.
- Environmental sustainability laws that protect common access resources (see Chapter 11) from being overly exploited. In general, the inability to dissuade people from over-consumption of natural resources, including forests, fisheries, public beaches and natural parks will lead to the tragedy of the commons and destruction and depletion of non-renewable and common pool resources.

ATL ACTIVITY 33.2

Achieving sustainable development is a time-consuming, multilayered and complex process that needs initiatives from all the governments, policymakers, scientists, economists and experts from academia. It also requires the common knowledge and understanding of the dangers of not making progress towards a more sustainable future.

Investigate the various ways to achieve sustainable development. Ensure you make reference to social, economic and environmental sustainability.

Sustainable development goals (AO2)

The **Sustainable Development Goals** (SDGs) of the United Nations consist of 17 international development targets, and 169 associated objectives, aimed at achieving peace and prosperity for every human being on the planet. The SDGs were adopted in 2015 and the 17 goals are all targeted to be achieved by each and every UN member country by 2030.

Prior to the agreement on the 17 SDGs in 2015, the UN focused on the eight **Millennium Development Goals** (MDGs). These were originally planned to be achieved by all 193 UN member countries by 2000 (hence, the name). The target for achieving the MDGs were extended to 2015, prior to UN members adapting the new SDGs.

The eight Millennium Development Goals of the United Nations Development Programme (UNDP) were to:

1. Eradicate extreme poverty and hunger; that is, reducing the number of people whose income is below US$1.25 a day and those who suffer from hunger (such as underweight children). Note: The World Bank revised this figure – representing the international poverty line – to $1.90 per day in 2015.
2. Achieve universal education to primary level (at least); that is, to raise the school enrolment rate and the literacy rate.
3. Promote gender equality and empower women; that is, social, economic and ethical reasons for promoting gender equality, including equal access to education and employment.
4. Reduce child mortality rates; that is, reducing the infant mortality rate such as by raising the number of young children who are immunized against measles.
5. Improve maternal health; that is, raise the proportion of births supported by skilled healthcare professionals such as qualified doctors and nurses.
6. Combat diseases such as HIV/AIDS and malaria; that is, to stop the spread of HIV/AIDS and major diseases such as malaria and tuberculosis.
7. Ensure environmental sustainability; for example, by reducing energy use and carbon emissions while improving access to clean water sources and better sanitation.
8. Develop a global partnership for development; that is, to address the needs of low-income countries, including debt servicing, debt relief, sustainable debt, infrastructure and employment opportunities for women and the youth.

◆ The **Sustainable Development Goals** (SDGs) of the United Nations Development Programme consist of 17 international development targets aimed at achieving peace and prosperity by all UN member countries by 2030.

◆ The **Millennium Development Goals** (MDGs) of the United Nations committed member countries to combat poverty, hunger, disease, illiteracy, environmental degradation and discrimination against women. The MDGs were scheduled to be achieved by 2015.

Figure 33.5 Millennium Development Goals

Although the MDGs were not all fully achieved, there have been some major achievements that enabled the UNDP to create the SDGs. For example, although extreme poverty and hunger have not been completely eradicated, more than 1 billion individuals were lifted out of extreme poverty between 1990 and 2015. The infant mortality rate has also decreased by more than half since inception of the MDGs. The number of children who do not attend any form of education has also dropped by more than half from its pre-1990 level.

The UNDP's Sustainable Development Goals (SDGs) were conceived in Rio De Janeiro, Brazil during the 2012 UN Conference on Sustainable Development (UNCSD). The SDGs were developed and scheduled to supersede the Millennium Development Goals (MDGs) in 2015. The 17 SDGs, to be achieved by 2030, are:

1. End poverty in all its forms everywhere.
2. End hunger, achieve food security and improved nutrition and promote sustainable agriculture.
3. Ensure healthy lives and promote well-being for all ages.
4. Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.
5. Achieve gender equality and empower women and girls.
6. Ensure availability and sustainable management of water and sanitation for all.
7. Ensure access to affordable, reliable, sustainable and modern energy for all.
8. Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all.
9. Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation.
10. Reduce inequality within and among countries.
11. Make cities and human settlements inclusive, safe, resilient and sustainable.
12. Ensure sustainable consumption and production patterns.
13. Take urgent action to combat climate change and its impacts.
14. Conserve and sustainably use the oceans, seas and marine resources for sustainable development.
15. Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation, and halt biodiversity loss.
16. Promote peaceful and inclusive societies for sustainable development, provide access to justice for all, and build effective, accountable and inclusive institutions at all levels.
17. Strengthen the means of implementation and revitalize the global partnership for sustainable development.

Source: **https://sustainabledevelopment.un.org/post2015/transformingourworld**

Figure 33.6 Sustainable Development Goals

ATL ACTIVITY 33.3

With another student, reflect on whether the Sustainable Development Goals (SDGs) are truly or completely different, or just an extension of the Millennium Development Goals (MDGs). Share your opinions with the rest of the class to see which line of argument is in the majority. In your discussions, consider the merits and limitations of changing from 8 to 17 goals.

ATL ACTIVITY 33.4

According to the most recent estimates from the World Bank, in 2015, there was still around 10 per cent of the world's population who lived on less than US$1.90 a day. Back in 2013, this figure was 11 per cent.

Investigate the main reasons why so many governments are not on target to achieve the first SDG by the end of the decade.

WISE ChoICES

Discuss how the **choice** of postponing consumption for the future makes the present more sustainable, leading to economic **well-being** and **sustainability**.

Relationship between sustainability and poverty (AO2) (HL only)

According to data from the United Nations, the level of world poverty has decreased more than half since the start of the millennium. However, hundreds of millions of individuals are still living on less than US$1.90 per day, the universally accepted poverty line. The World Bank states that around 800 million people in the world are still living in extreme poverty, while struggling to fulfil the most basic necessities to live a decent life, such as access to healthcare, education, clean water and sanitation. Perhaps even more staggering is that around 46 per cent of people

Theory of Knowledge

With reference to the UN's SDGs, discuss the extent to which ethical thinking plays a role in deciding if poverty leads to unsustainable development.

ATL ACTIVITY 33.5

Consider which one of the 17 Sustainable Development Goals you feel *most* attached to. Which of the 10 IB Learner Profile attributes influenced your choice?

Common mistake

It is incorrect to assume that each of the SDGs is a standalone goal. In reality, any progress made to achieve one goal, such as ending poverty in all of its forms everywhere (SDG1), leads to improvements in other SDGs or key indicators, such as ending hunger (SDG2) and reducing inequality within and among countries (SDG10). Hence, the interconnected nature of the goals makes it relatively easier to achieve other SDGs.

worldwide, or 3.4 billion people, are still living below $5.50 per day according to the World Bank. The situation is worse of Africa, where 85 per cent of Africans live on less than $5.50 per day.

According to the World Bank:

> *As the world grows wealthier and extreme poverty becomes more concentrated, there are legitimate questions over whether $1.90 is too low to define whether someone is poor in all countries of the world. In half of the countries in the world, extreme poverty is at or below 3 percent, but that doesn't mean the fight to eradicate poverty is over in these countries. The World Bank now reports on two higher-value poverty lines: $3.20 and $5.50 per day. These lines, which are typical of standards among lower- and upper-middle-income countries, respectively, are designed to complement, not replace, the $1.90 international poverty line. Data suggest that the rapid gains against extreme poverty have not been matched by reductions in the number of people living below these higher levels of income. In 2015, over a quarter of the world's population survived on $3.20 per day and nearly half of the world still lived on less than $5.50 per day.*

Table 33.2 Percentage of population living under $1.90 – bottom 10 countries

Country	Percentage of population
Zambia	57.50
Lesotho	59.70
Uzbekistan	62.10
Mozambique	62.40
Central African Republic	66.30
Guinea-Bissau	67.10
Malawi	70.30
Burundi	71.80
Democratic Republic of the Congo	76.60
Madagascar	77.60

Source: https://data.worldbank.org/indicator/SI.POV.DDAY

Figure 33.7 More than 77 per cent of people in Madagascar live on less than $1.90 a day

Poverty and the UNDP Sustainable Development Goals

Poverty and sustainability are closely interlinked as reflected in the first of the UNDP's 17 SDGs. It is only through sustainable development that poverty can be completely eradicated in the long run. The progress made in meeting the MDGs by 2015 fell short of the targets. Regional progress was uneven, substantial inequalities remained within and across countries, so there has been an inconsistent degree of poverty reduction across the globe. Ending poverty in all forms is the most important task recognized by the UN that leads to adopting the first of its 17 SDGs. The new agenda has unprecedented scope and capacity to address the issue of poverty through sustainable development, in terms of social, economic and environmental sustainability.

The SDGs recognize that eradicating poverty in all its forms, overcoming inequalities, creating inclusive economic growth and achieving environmental sustainability are intricately linked and the relationships between all these components are reciprocal, dynamic as well as mutual. It is not possible to achieve economic sustainability without protecting the planet. Similarly, it is not possible to eradicate hunger and poverty without being socially sustainable. Hence, the UNDP's

seventeenth SDG calls for the global partnerships for sustainable development by formulation of targets and policies in collaboration with the international, national and local governments and non-governmental organizations (NGOs).

The UNDP's plan of action, titled 'Transforming our World: the 2030 Agenda for Sustainable Development' set the following five targets related to poverty eradication. This is the first of the 17 SDGs, that is, to end poverty in all its forms everywhere:

1.1 By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than US$1.90 per day.

Figure 33.8 Climate change is a hindrance to sustainability and poverty eradication

1.2 By 2030, reduce by at least half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions.

1.3 Implement nationally appropriate social protection systems and measures for all, and by 2030 achieve substantial coverage of the poor and the vulnerable.

1.4 By 2030, ensure that all men and women, in particular the poor and the vulnerable, have equal rights to economic resources, as well as access to basic services, ownership and control over land and other forms of property, inheritance, natural resources, appropriate new technology and financial services, including microfinance.

1.5 By 2030, build the resilience of the poor and those in vulnerable situations and reduce their exposure and vulnerability to climate-related extreme events and other economic, social and environmental shocks and disasters.

Economists argue that inadequate access to basic needs, education and suitable employment opportunities, along with conflicts, inequalities and climate change (caused mainly by the economic activities of high-income countries) are major reasons behind the sustained level of acute and mild poverty. Furthermore, low-income countries are unable to pursue sustainable growth and development without adequate infrastructure such as transportation and communications networks.

Poverty and the over-reliance on primary sector activities

The **primary sector** of the economy involves industries in the primary stage of production, such as the extraction and collection of natural resources. Examples include agricultural farming, forestry, hunting, fishing, oil exploration and mining. Data from the World Bank show that the majority of the people living below the international poverty line live mainly in southern Asia and sub-Saharan countries. Most of the countries in these regions are agriculturally based where the secondary and tertiary sectors are relatively undeveloped. These countries tend to be heavily reliant on agriculture, fisheries and forestry. However, the growth of these sectors is largely dependent on soil conditions, land availability and climate conditions. Sustainable agriculture is therefore vital for poverty eradication for farmers and farm workers in these countries.

◆ The **primary sector** of the economy involves industries in the primary stage of production, such as the extraction and collection of natural resources. Examples include agricultural farming, forestry, hunting, fishing, oil exploration and mining.

Poverty and access to education

The lack of equal access to education adds more to the problem of sustainable development and poverty. Education leads to greater employability and skills development, that in turn leads to an increase in household income and a reduction in poverty. Technological change and globalization forces workers to adapt and develop a diversified portfolio of skills and higher education. It also facilitates research and development, as well as innovation of new techniques or ideas. However, in many countries, girls do not have equal access to schooling because of limited financial means and an orthodox mentality. Nevertheless, boys from the poorest families tend to drop out of school in favour of working on farms or unorganized industrial labour markets in order to support their families.

Interestingly, statistics reveal that girls from marginalized societies tend to perform better in middle and high schools than boys with the same socioeconomic condition and background. The gender gap and gender discrimination hinder the social and economic sustainable developmental process. Even in middle- and high-income countries, it is the poorest members of society who suffer the most. For example, during the closure of schools across the world in 2020 owing to the coronavirus pandemic, poorer families struggled the most with online/distance learning, as not all families own laptops or desktop computers, and many children had to share devices. Hence, the lack of equal access to education does little to promote the eradication of poverty or social and economic sustainability.

Theory of Knowledge

What knowledge questions might be encountered in deciding the impact of global pandemics, like Covid-19, in slowing down the progress of attaining the Sustainable Development Goals?

Figure 33.9 Equal access to education, for girls and boys, is vital for poverty eradication

Poverty and the gender gap

The **gender gap** is the difference between women and men as reflected in social, political, intellectual, cultural and economic activities, attitudes and opportunities. Correcting the gender gap is important for sustainability. Hundreds of millions of people could come out of extreme poverty if women were given equal access rights to productive resources. Some studies suggest that around $12 trillion could be added to the world economy if women were given access to opportunities on par with those for men (only the USA and China have annual GDP in excess of $12 trillion).

◆ The **gender gap** refers to the difference between women and men in terms of social, political, intellectual, cultural and economic activities, attitudes and opportunities.

Women in low-income countries typically work in agricultural industries, but these are vulnerable to external shocks such as floods and droughts. Hence, it can be argued that women are particularly vulnerable to climate change as their livelihood depends on favourable weather conditions. In economically less developed countries, women tend to be engaged with work such as water collection, growing and harvesting of crops, transportation of agricultural output to market stalls as well as the highly demanding task of raising children and looking after the family. Much of this vital work is unaccounted for as there are no direct financial returns for doing so. This commitment to the family also gives women fewer opportunities to engage with mainstream economic activities that could help them achieve more social and economic opportunities, including the empowerment of women.

Closing the gender gap could also facilitate sustainable population growth, as more women focus on their professional careers rather than other priorities expected by society. Most economists are in favour of population control for sustainable development from social, economic and environmental perspectives, especially in low-income countries. Uncontrollable population growth is a major hindrance to equal access to scarce resources, education and healthcare, thereby making it more difficult to achieve sustainable development.

Figure 33.10 High population growth puts a strain on the planet's scarce economic resources

ATL ACTIVITY 33.6

Raising a child is considered a responsibility, or even an obligation, of a mother in many parts of the world. With a partner, discuss the extent to which a government should pay a stipend (a fixed regular sum of money) to mothers for raising a child. What are some major challenges the nation would face if the government introduced the stipend?

PAPER 2 EXAM PRACTICE QUESTION 33.1

Development in Rwanda

Small and landlocked, Rwanda is hilly and fertile with a densely packed population of about 12.5 million people (2018). It borders the far larger and richer Democratic Republic of the Congo, as well as its closest East African neighbours, Tanzania, Uganda and Burundi. With the support of the International Monetary Fund (IMF) and the World Bank, Rwanda has been able to make important economic and structural reforms and sustain its economic growth rates over the last decade.

Public investments have been the main driver of growth in recent years. External financing through grants, and concessional and non-concessional borrowing has played an important role in financing the public investments. Going forward, the private sector will play a bigger role in helping to ensure sustainable economic growth. However, low domestic savings, inadequate skills and the high cost of energy are major constraints to private investment in Rwanda. Promoting domestic savings is therefore viewed as critical.

The percentage of the population living below the international poverty line declined from 59 per cent to 39 per cent between 2001 and 2014, but was almost stagnant between 2014 and 2017. The official measure of income and wealth inequality, the Gini index, declined from 0.52 in 2006 to 0.43 in 2017. Gender inequality has reduced considerably since 2006.

Source adapted from: **https://www.worldbank.org/en/country/rwanda**

a Explain how a reduction in gender inequality might help Rwanda to achieve sustainable economic development. *[4 marks]*

b Using information from the text/data and your knowledge of economics, discuss methods that the Rwandan government might employ to achieve sustainable development. *[15 marks]*

CASE STUDY

Ping An helping China by offering innovative insurance schemes

Ping An is one of the largest insurance companies in the world. Based in China, the for-profit organization developed a model for poverty alleviation for diverse business sectors through different insurance schemes, such as the 'Poverty Alleviation Insurance Scheme' or the 'Hydropower Support Scheme'. The company presented this innovative model at the World Economic Forum (WEF) in Davos, Switzerland in 2020. Ping An offered a 'credit guarantee and interest subsidy' model that could spread the risk to a wider group of stakeholders, such as governments and commercial banks.

Top tip!

Remember that economic growth does not always lead to economic development. Indeed, economic growth can help to eradicate extreme poverty, but it does not necessarily promote gender equality or other international development goals.

WISE ChoICES

Discuss whether the pursuit of economic **sustainability** always leads to **economic well-being** in an **equitable** way. Provide real-world examples to justify your answers.

Chapter summary

- Sustainable development is economic development that meets the needs of the present generation without compromising the ability of future generations to meet their own needs.
- There are three interlinked pillars or dimensions to sustainability: social, economic and environmental sustainability.
- Most of the countries in the world strive to achieve the 17 Sustainable Development Goals of the United Nations Development Programme, enacted in 2015, with the aim of these goals being achieved by 2030. The SDG replaced the previous eight Millennium Development Goals of the UNDP.
- Poverty and sustainable development are intricately interwoven, and some theories believe that one is the cause and another is the effect. Through sustainable development, by social, economic and environmentally sustainable means, poverty could be eradicated entirely. (HL only)
- A reduction in poverty is one of the most important factors that ensure sustainable development. (HL only)
- Most of the poorest people in low-income countries live in acute poverty (less than US$1.90 per day) and are dependent on primary sector activities. (HL only)
- Education leads to employability and skills development, which can help to increase household income and reduce poverty. (HL only)
- A gender gap worsens the income distribution, which leads to poverty and threatens sustainable development. (HL only)

34 Measuring development

Real-world issue

Why is economic development uneven?

Conceptual understandings (WISE ChoICES)

- Perceptions of the meanings of development and **equity** change over time and vary across cultures.
- Governments and other economic agents may intervene in an attempt to promote **economic well-being** and **equity** in societies.
- The pursuit of **sustainability** is subject to various constraints.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:
- the multidimensional nature of economic development (AO2)
- single indicators as measures of development (AO2)
- composite indicators as measures of development, with specific reference to the Human Development Index (HDI), the Gender Inequality Index (GII), the Inequality-adjusted Human Development Index (IHDI), and the Happy Planet Index (AO2)
- the strengths and limitations of approaches to measuring economic development (AO3)
- the possible relationship between economic growth and economic development (AO3).

The multidimensional nature of economic development (AO2)

The concept of economic development has evolved in the last few centuries, especially from the middle of the twentieth century, that is, the end of the Second World War. The term has been referred to as modernization, Westernization and industrialization in many academic and political contexts.

One of the oldest lines of thought linked to the theory of economic development was the theory of Mercantilism, prevalent during the seventeenth century in modern Europe. Mercantilism was a political and economic policy that focused on wealth creation through the use of international trade. Countries tended to maximize exports and minimize imports, ensuring an incoming flow of wealth into the nation.

During the same time in the UK, a line of distinguished economists influenced the academic landscape of economics. The likes of Adam Smith, Thomas Robert Malthus, Jean-Baptiste Say, David Ricardo and John Stewart Mill, collectively known as the

Figure 34.1 The monument of Adam Smith, the father of Classical economics, in Edinburgh, Scotland

classical economists, championed a self-regulating system of production, trade and exchange without the intervention of government authorities. Their ideology believes that equilibrium is achieved in an economy when people are driven by self-interest. Classical economists also advocated international trade and were against all types of protectionist measures taken by governments; it is this that differentiates the theories of classical economists from the mercantilists' line of thought.

◆ **Classical economics** was the main economic school of thought during the eighteenth and nineteenth centuries, which focused on self-regulating markets to allocate resources efficiently.

In the nineteenth century, with the advent of the Industrial Revolution, countries including the United Kingdom, Germany and the USA experienced an unprecedented level of economic prosperity. Having market presence in international markets led to the dominance of Japan in East Asia during much of the twentieth century. The economies known as the 'Asian Tigers' (comprising South Korea, Taiwan, Hong Kong and Singapore) experienced phenomenal growth and development in the mid and late twentieth century, owing to their export-oriented approach and focus on human and physical capital, enabling these nations to become classed as more economically developed countries (MEDCs).

Owing to the **multidimensional nature of development**, during the 1970s, economists started to look beyond economic growth as a solution to eradicate poverty, reduce malnutrition, increase job opportunities and provide access to basic needs. They started to look into the multidimensional nature of development, which includes more than economic growth, such as consideration of qualitative determinants of quality of life. These factors include the importance of a reduction in poverty, income inequality, gender inequality, political oppressions and unemployment. American economist Professor Michael Todaro (b.1942) conceived development as a multidimensional process in his PhD thesis titled 'The Urban Employment Problem in Less Developed Countries – An Analysis of Demand and Supply' (1968). Todaro argued that development, being a multidimensional concept, requires major changes in social structures, popular attitudes and national institutions, as well as the acceleration of economic growth, the reduction of inequality and the eradication of absolute poverty.

American economist Professor Denis Goulet (1931–2006), credited for coining the term development ethics, indicated three basic components as core values that should be considered to achieve economic development. The components of his work, featured in *The Cruel Choice: A new concept in the theory of development* (1971) are sustenance, self-esteem and freedom. Goulet is noted for broadening the notion of development to include social objectives and values that societies should strive for, in addition to economic objectives. Goulet's work on development ethics is often viewed as the precursor to welfare economist Amartya Sen's (b.1933) work that culminated in the Human Development Report of the United Nations Development Programme (UNDP) in 1990 and the adaptation of the UNDP's Human Development Index (HDI). Sen advocated the importance of human development in the form of freedom from hunger, illness, unemployment and social stigma. He also advocated the right of every human being to develop a fulfilling and productive life.

Single indicators (AO2)

As economic development is a complex and multilayered issue, measuring development is a challenging task. Different economists and theorists have approached measuring development in different ways. **Single indicators** are used to describe individual facets of development. For example, there is a direct correlation between a country's gross domestic product (GDP) per capita and its ranking of economic development. Although, the total scope of development can hardly be measured by a single indicator (owing to the multidimensional nature of development), they serve as relatively straightforward measures to gauge the level of development. After all,

◆ A **single indicator** refers to a statistical measure of economic development that uses one particular gauge, such as literacy rates, income per capita or life expectancy.

economic models make assumptions and simplify matters to examine cause of effects on the economy (see Chapter 2). Given that it is not pragmatically possible to account for all aspects of development, single indicators serve the purpose of allowing economists and policymakers to see the impact of these factors on the economy's development.

The Sustainable Development Goals (SDGs) of the United Nations, adopted in 2015, consist of 17 international development goals, 169 associated targets and 232 indicators aiming for the peace and prosperity of the people and the planet to be achieved by all UN member countries by 2030. As far as composite indicators are concerned, the SDGs are as comprehensive as it gets. However, out of the 232 indicators, only some have been used frequently by policymakers and economists while some have been used only sparsely. Hence, the use of single indicators is often considered as more practical and purposeful.

Theory of Knowledge

Of the 232 indicators used to measure the progress of the SDGs, how do we determine which indicators are more important than the others?

Figure 34.2 How feasible is it to have 232 single indicators to measure the progress of the SDGs?

ATL ACTIVITY 34.1

If you are given a job of President of one of the economically least developed countries on the planet, which two IB Learner Profile attributes would you choose to adopt to ensure better implementation of existing macroeconomic policies to achieve economic development?

The IB Economics syllabus specifies five categories of single indicators that can be used to measure development GDP per person (or GNI per capita) as purchasing power parity, health and education indicators, economic/social inequality indicators, energy indicators and environmental indicators.

GDP/GNI person (per capita) at PPP

Recall from Chapter 3 that gross domestic product (GDP) is the value of all final output of goods and services produced within a country in a year. Gross national income (GNI) is the value of a country's final output of all goods and services plus net factor income earned from abroad.

GDP per capita calculates the value of national output of a country (its gross domestic product) divided by its population. Essentially, this expresses national income as **GDP per person**. It is the most used single indicator of standards of living within a country (see Table 34.1). GDP per capita includes productive assets owned by foreign companies operating within the country. However, it does not include the output of national companies operating outside of the country; this is included in the calculation of **GNI per capita** (gross national income per person). Either of these methods can be used as a single indicator to measure economic development.

Comparing nominal GDP and GNI figures of different countries can be meaningless owing to variations in exchange rates and costs of living. For international comparisons, GNI and GDP figures are expressed in terms of **purchasing power parity** (PPP). This refers to the exchange rate that equates the price of a basket of the same traded goods and services in different countries. Using GDP per capita at PPP exchange rates enables a more meaningful comparison of differing costs of living (and, hence, standards of living) across different countries.

Hence, expressing GNI and GDP per capita at PPP enables economists and policymakers to measure the degree of economic development in different countries.

◆ **GDP/GNI per capita** (or **GDP per head**) is the average income in an economy, i.e. the total GDP/GNP divided by the population size.

◆ **Purchasing power parity** (PPP) refers to the exchange rate that enables residents to purchase a common basket of goods and services in different countries.

Table 34.1 GDP/GNI per capita (selected countries), 2020

Country	GDP (PPP) per capita ($)	GDP (PPP) per capita ranking	Nominal GDP per capita ($)	Nominal GDP per capita ranking
Qatar	138,910	1	70,737	6
Macao SAR	114,363	2	80,065	4
Luxembourg	108,951	3	116,727	1
Singapore	103,181	4	64,829	9
Ireland	83,399	5	80,265	3
Brunei Darussalam	80,384	6	28,397	35
Norway	76,684	7	78,333	5
United Arab Emirates	69,435	8	37,375	26
Kuwait	66,387	9	28,883	33
Switzerland	66,196	10	86,674	2

Source: http://statisticstimes.com/economy/projected-world-gdp-capita-ranking.php

Using GDP (or GNI) per capita at PPP is regarded as the most common single measure of the standard of living (or well-being) of a nation. It is calculated by adjusting the country's GDP per capita for differences in the cost of living between countries. Notice in Table 34.1 that Brunei Darussalam, the UAE and Kuwait are not within the top 25 nations based on nominal GDP per capita but that they are all within the top 10 once the figures are adjusted for purchasing power parity.

Top tip!

Using GDP per capita to measure the level of economic development in low-income countries can be misleading as the value of parallel markets (unofficial transactions) is unrecorded.

ATL ACTIVITY 34.2

India tops the total amount of remittances sent back home in 2019. However, India's GDP per capita is higher than its GNI per capita. Think about the problems this creates, then pair with your partner and share your explanations with the rest of the class.

PAPER 2 AND 3 EXAM PRACTICE QUESTION 34.1

Suppose the exchange rate between Demland and Danland is 1 dems to 3.5 dans. If a Big Mac burger is priced at 3.5 dems and there is purchasing power parity, calculate the price of a Big Mac burger in Danland. *[2 marks]*

PAPER 2 AND 3 EXAM PRACTICE QUESTION 34.2

Assume in an economically less developed country (ELDC) that the real gross national product (GNI) increases from $50 billion to $54 billion and that its population increases from 68 million to 71 million. Calculate the change in the ELDC's real GNI per capita. *[2 marks]*

Top tip!

Whichever measure of income is used to measure economic development (real GDP/GNI per capita or GDP/GNI per capita at PPP exchange rates), it is worth noting that other essential components of development are ignored, such as income and wealth inequalities, gender inequalities, political freedom and environmental issues.

Health indicators

Health indicators are a category of single indicators of measuring economic development through the important attributes of health-related determinants of the quality of life, such as: longevity (life expectancy at birth), expenditure on healthcare as a percentage of GDP, and under-five mortality rates (see Table 34.2).

◆ **Health indicators** are a category of single indicators of measuring economic development by using health-related determinants of the quality of life, such as life expectancy and the under-five mortality rate.

Table 34.2 Life expectancy at birth and mortality rate*

Country	Life expectancy at birth (years)	Mortality rate (per 1,000 live births)
Japan	85	3
Iceland	83	2
United Kingdom	81	4
Argentina	77	10
Jordan	74	16
India	69	37
Pakistan	67	69
Somalia	57	122
Nigeria	54	120
Sierra Leone	54	105

* measured by those aged under 5 per 1,000 live births, 2018

Source: World Bank https://data.worldbank.org/indicator/SP.DYN.LE00.IN?year_high_desc=false

Life expectancy at birth indicates the number of years a person in a country is expected to live, expressed as the average number of years for the entire population of the country. Economically developed countries like Japan (85), Hong Kong SAR (85) and Western European countries are

among the highest for life expectancy at birth, whereas West African countries like Sierra Leone (54) and Central African Republic (53) have among the lowest life expectancy rates.

Perhaps unsurprisingly, the countries with the highest life expectancy at birth also spend more on healthcare as a percentage of their GDP. Table 34.3 shows data from the World Bank for selected countries, expressing their healthcare expenditure as a percentage of GDP.

Table 34.3 Healthcare expenditure as percentage of GDP, selected countries

Country	% of GDP	Country	% of GDP
USA	17.06	Venezuela	1.18
Sierra Leone	13.42	Bangladesh	2.27
Switzerland	12.35	Brunei Darussalam	2.37
Afghanistan	11.78	Qatar	2.61
Cuba	11.71	Gabon	2.78
France	11.31	Angola	2.79
Germany	11.25	Pakistan	2.90
Sweden	11.02	Republic of Congo	2.93
Japan	10.94	Indonesia	2.99
Canada	10.57	Ghana	3.26

Source: Word Bank (https://bit.ly/2yRJKfy)

Countries with a high GDP/GNI tend to have a lower infant mortality rate and a higher life expectancy rate. Countries with high per capita income are also the ones that tend to spend more money on healthcare. But this does not necessarily apply to all countries. For example, Sierra Leone spends more than 10 per cent of its GDP on healthcare even though the West African country has one of the highest infant mortality rates in the world. Countries that fare better in health indicators provide universal healthcare for all citizens of the country, including the prevention of infectious diseases and provision of general healthcare information and education for the population. They also make available safe drinking water, sanitation and better government provision of basic services by addressing poverty and gender inequalities.

Education indicators

Education indicators are a category of single indicators of measuring economic development. They measure education-related indicators of the quality of life, such as literacy rates and the mean average years of schooling. They help to monitor the level of participation in education, such as enrolment, at a particular age group to identify the scope for improvement in major aspects of education. Goal 4 of the UNDP's Sustainable Development Goals is to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all. This signifies the importance of measuring education indicators to inform policymakers in their pursuit of development.

◆ **Education indicators** are a category of single indicators of measuring economic development by using education-related determinants of the quality of life, such as literacy rates and the mean years of schooling.

The UNESCO Global Education Monitoring Report (**http://gem-report-2019.unesco.org**) suggests that 264 million children of school age are not attending school. This jeopardizes the chances of the country achieving the Sustainable Development Goals. For example, the report found that the adult literacy rate was only 35 per cent in Mali, 31 per cent in Niger and just 22 per cent in Chad.

Figure 34.3 Investment in education is vital for economic development in the long run

Table 34.4 Mean years of schooling and adult literacy rate* for selected countries

Country	Mean years of schooling (years)	Adult literacy rate (%)	HDI rank
Norway	12.6	100.0	1
Switzerland	13.4	100.0	2
Australia	12.9	100.0	3
Montenegro	11.3	99.0	50
Bulgaria	11.8	98.0	51
Romania	11.0	99.0	52
Suriname	8.5	94.0	100
Botswana	9.3	87.0	101
Rwanda	4.1	73.0	156
Sudan	3.7	61.0	167

* Literacy rate, adult total (% of people ages 15 and above)

Sources: http://hdr.undp.org/en/composite/HDI *and* https://data.worldbank.org/indicator/SE.ADT.LITR.ZS

ATL ACTIVITY 34.3

Investment in education is considered as a long run supply-side policy. With a partner, debate why a government of a low-income country might choose to prioritize investment in education, as opposed to providing capital for other investments that would provide quicker results. Consider the winners and losers from such decisions.

Improvements in all aspects of education indicators will facilitate countries in meeting the UN Sustainable Development Goals through higher labour market participation and productivity needed for achieving economic growth and development. In the long run, this also help to reduce poverty and hunger (Goals 1 and 2), and other relevant goals such as good health and well-being (Goal 3), achieving gender equality and empowering all women and girls (Goal 5), full and productive employment and decent work for all (Goal 8), reduced inequalities (Goal 10) and partnerships for sustainable development (Goal 17).

Economic/social inequality indicators

Economic indicators are a category of single indicators of measuring economic development using different economic parameters, such as real GNI per capita, real GDP per capita (both of which were discussed in the first part of this chapter), per capita consumption and gross fixed capital formation.

Per capita consumption (PCC) is measured as the value of consumption of goods and services per person for a given period of time. PCC is calculated by the sum of total consumption in the economy divided by the population of the country. The higher the PCC, the higher the level of development tends to be.

Gross fixed capital formation (GFCF) refers to the capital investment of the country as a percentage of the national income. The greater the value of GFCF, the more likely the country is to enjoy greater levels of development in the future. This could be depicted by using a production possibility frontier diagram (see Chapter 2), with the PPF curve shifting outwards over time due to the increased productive capacity of the economy. Without capital investments, it is not possible to sustain development in the long run.

Figure 34.4 Some countries like Denmark place a higher degree of trust in the community

Social indicators use measures that determine the extent to which societal factors contribute to development. These include some of the health and education indicators discussed above, but also include indicators such as shelter and housing, crime rates, homicide rates, safety and the degree of trust in a neighbourhood or community. Some of the most prominent social inequality indicators are income inequality and wealth inequality, such as the Gini coefficient, which is used as a universal indicator of income inequality (see Chapter 23).

However, a single economic or social indicator such as real GNI/GDP per capita, the PCC, GFCF or GNI can hardly indicate the overall quality of development of the country as such measures ignore qualitative information that affects the human development of people living in the country. Nevertheless, using these single indicators can certainly shed some light on a specific and important aspect of the overall level of development in the country.

Energy indicators

Energy indicators are a category of single indicators of measuring economic development by using indicators that identify the factors needed to create affordable, reliable, sustainable and modern energy for all citizens. Energy played a pivotal role in economic growth achieved in the last few centuries. Energy is still considered as a one of the most important factors to achieve sustainable development. With only finite reserves of fossil fuels, countries are looking for new sustainable sources of energy. There are various energy indicators that can be used to measure development, such as the indicators below:

◆ **Energy indicators** are a category of single indicators of measuring economic development by using factors that create affordable, reliable, sustainable and modern energy for all citizens.

- Proportion of population with access to functioning electricity.
- Proportion of population with primary reliance on clean fuels and technology.
- Renewable energy usage as a proportion of total energy consumption.
- Energy use as a proportion of GDP.

There has been a move towards production and consumption of green energies, that is, energy made from natural sources such as sunlight, wind, rain, tides, plants, algae and geothermal heat. These alternative energy sources are complex and have had varying degrees of success. China, the USA, the Scandinavian countries and several other economically developed countries are leading the initiative to move from traditional sources of energy to renewable sources of energy. The transition to new energy sources requires huge set-up costs and ongoing investments in research

and development. That may increase the price of energy, which many ELDCs and emerging countries may not be able to afford.

Nevertheless, energy indicators are vital for development, and are part of the UN's seventh SDG: to ensure access to affordable, reliable, sustainable and modern energy for all. There is a clear correlation between access to electricity and the level of economic development (see Table 34.5).

Table 34.5 Access to electricity (% of population) for selected countries, 2017

Country	Access to electricity (% of population)
Saudi Arabia	100.0
Thailand	100.0
Fiji	96.0
Bolivia	91.8
South Africa	84.4
Eswatini	73.5
Senegal	61.7
Uganda	22.0
Malawi	12.7
Burundi	9.3

Source: https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS

ATL ACTIVITY 34.4

Based on the data in Table 34.5, reflect if countries having a greater access to electricity tend to be more developed.

Environmental indicators

Environmental indicators are a category of single indicators of measuring economic development by using environmental issues that influence or determine the quality of life. These interrelated indicators include:

- global warming and climate change
- desertification
- deforestation
- waste disposal management systems
- loss of biodiversity and ecosystems.

◆ **Environmental indicators** are a category of single indicators of measuring economic development by using environmental issues that influence or determine the quality of life, such as global warming and climate change.

Figure 34.5 Environmental indicators of development became commonplace after the first Earth Summit in Rio

Environmental indicators are a relatively new form of measuring development. It was after the United Nations Conference on Environment and Development held in Rio de Janeiro, Brazil in June 1992, that environmentalists, developmental economists and governments became more conscious of the long-term impacts of environmental situations all across the world. It was at the Rio Conference (often referred to as the Earth Summit) that nations took a joint decision to initiate the raising of public awareness of the need to integrate environmental issues and sustainable developmental initiatives. Some of these indicators are outlined below.

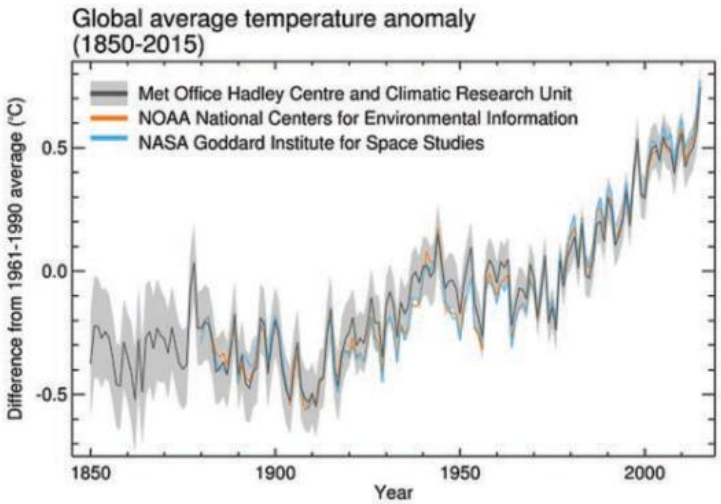


Figure 34.6 Global average temperature anomaly (1850–2015)

Global warming and climate change

Scientists measure global warming from the air above the land and the ocean surface, recording and monitoring variances in annual temperatures. A temperature anomaly is the difference between the current and average temperature, calculated from data collected over many years (see Figure 34.6). A positive anomaly signals that the observed temperature was warmer than the average, and a negative anomaly points towards a cooler temperature than the average. Using data for annual temperature anomaly offers more meaningful comparisons between different locations over time and provides more accurate calculations of temperature trends.

Desertification

Desertification is a category of land degradation caused by natural processes or human-induced activities that result in fertile land becoming desert; that is, the over-exploitation of soil and land through human activity. Typically, desertification is caused by droughts, forest fires, global warming, overgrazing, deforestation or inappropriate agriculture (unsustainable farming). It is a major threat to sustainability and development because fertile and arable land becomes arid and unproductive. Desertification has been a major cause of famine in many least economically developed countries.

ATL ACTIVITY 34.5

Take a look at The World Counts (**https://bit.ly/2Vp9XtI**) to see how the Earth's average temperature is continually increasing. According to the website (**www.theworldcounts.com**), if no action is taken to tackle the global problem of climate change, temperatures could increase by 5 degrees by 2100. Use the website to help answer the question below.

What are the economic implications of global temperatures increasing by around 5 degrees within the next eight decades?

Figure 34.7 Fertile land becomes desert due to droughts or inappropriate agricultural practices

Desertification not only causes drylands and a loss of fertile soil, but the process cannot be reversed for many decades. Land erosion and soil degradation are intertwined with threats to sustainability and development, including water erosion, wind erosion, salinization, chemical and biological degradation, and physical poverty.

◆ **Desertification** is a category of land degradation caused by natural processes or human-induced activities that result in fertile land becoming desert; that is, the over-exploitation of soil and land through human activity.

Deforestation

Deforestation is the permanent destruction of trees and forests to use the space for alternative purposes, such as for preparing land for commercial agriculture or grazing, construction or manufacturing facilities. Forests cover more than 30 per cent of the Earth's land surface and provide sustenance to natural ecosystems and billions of people. The deforestation rate is an environmental indicator used to measure the pace and scale of permanent destruction of trees and forests. A positive deforestation rate indicates changes in the landscape and generation of trees, whereas a negative deforestation rate indicates loss of trees and forests as a result of human and natural deforestation. The countries with the greatest number of forests in the world are also the most affected by deforestation, which include Russia, Brazil, Canada, the USA, China, Australia, Democratic Republic of the Congo, Argentina, India and Indonesia.

Waste disposal management systems

Waste disposal is another environmental indicator of development that forces people to change their behaviour and lifestyle. The picture of a seahorse swimming with a human cotton bud (swab) by the coast of Borneo, clicked by Justin Hofman (2017), became the centre of discussion in every part of the world. It is a just a tip of the iceberg of the huge problem we are facing as a result of unsustainable waste disposal management and irresponsible use of plastics.

■ **Figure 34.8** Deforestation in the pursuit of economic growth destroys natural habitats

Figure 34.9 Justin Hoffman showed that plastic waste in our oceans is a major threat to sustainable development

World Bank figures indicate that 2 billion tonnes of municipal solid waste is generated worldwide each year, of which one third is not managed in an environmentally responsible manner. The World Bank predicts that 2.2 billion tonnes of municipal solid waste will be produced by 2025, or 1.42 kilograms per person per day. This is largely due to the increase in consumerism, causing the continual and perhaps unsustainable increase in solid waste per capita. According to the United Nations Department of Economic and Social Affairs, the world's population could reach 9.8 billion people by 2050, meaning that the issue of sustainable waste management systems will become one of the most challenging but critically important tasks faced by governments around the world.

Some of the basic environmental indicators used to study and measure trends in waste management include:

- The amount of municipal solid waste (MSW) generated per capita (see Table 34.6).
- The volume of physical waste sent to landfills or incineration plants/facilities.
- The recycling rate, expressed as the percentage of household waste that is recycled.

Table 34.6 Top five countries for municipal solid waste (MSW) per capita, 2019

Country	MSW per capita per day (kg)
Kuwait	5.72
Antigua and Barbuda	5.50
St Kitts and Nevis	5.45
Guyana	5.33
Sri Lanka	5.10

Source: World Bank

Loss of biodiversity and ecosystems

Maintaining or improving biodiversity is another challenging task for governments in order to repair the damage done from uncontrolled growth and the pursuit of profits over the planet. Several environmental indicators are used to measure and monitor the pressures or threats on biodiversity and ecosystems, such as trends in land and water use, loss of natural habitat and the state (condition or quality) of ecosystems.

Some of the strategies used by governments to protect biodiversity and ecosystems below water and on land (the UNDP's Sustainable Development Goals 14 and 15 respectively) include:

- Protection of important biodiversity areas and ecosystems, such as the Great Barrier Reef off the coast of Queensland, Australia.
- Proportion of important sites for terrestrial and freshwater biodiversity and ecosystems, including mountain biodiversity.

ATL ACTIVITY 34.6

Visit the World Bank's website (**https://bit.ly/3eg9faR**) to read about how different countries around the world are tackling the global issue of solid waste management, such as the investment projects in solid waste management in Vietnam and the Philippines in order to prevent clogging of drains, which could result in flooding.

- Sustainable agriculture and farming, including fish stocks.
- Law enforcement and sanctions against illegally traded wildlife, including poaching or illicit trafficking of animals or animal products.
- Official Development Assistance and public expenditure on conservation and sustainable use of biodiversity and ecosystems.
- International co-operation to further reinforce laws, rules and systems, such as membership of the Convention on International Trade in Endangered Species (CITES), which regulates the international trade in animals and plants. There are heavy penalties including imprisonment for crimes that contravene these international agreements between governments.

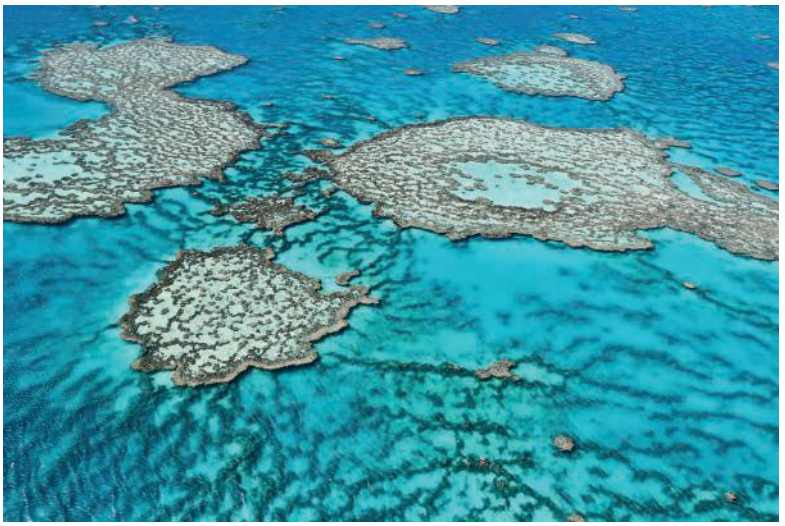


Figure 34.10 The Great Barrier Reef is the world's largest coral reef system

Composite indicators of measuring development (AO2)

Composite indicators of measuring development use a range of measures to calculate the degree of development in a country. A **composite indicator** is a statistical method that combines single indicators of economic development into a combined index such as the Human Development Index. Although they are complicated to compile and measure, composite indicators of measuring development encompass greater breadth and depth of quantifying development within a country. Hence, they are considered to be better measures of economic development than single indicators as they are more comprehensive.

◆ A **composite indicator** is a statistical method that combines single indicators of economic development into a combined index such as the Human Development Index.

The IB Economics syllabus states four composite indicators used to measure development. These are the Human Development Index (HDI), Gender Inequality Index (GII), Inequality-adjusted Human Development Index (IHDI), and the Happy Planet Index (HPI).

The Human Development Index (HDI)

The most widely used composite indicator of economic development is the Human Development Index (HDI), created by Pakistani economist Mahbub ul Haq and Indian economist Amartya Sen in 1990 as part of the United Nations Development Programme (UNDP). The **HDI** is a composite indicator of life expectancy, educational attainment and income, used as an alternative to real GDP or GNI per capita as a measure of economic development. The HDI measures three dimensions of human development:

◆ The **Human Development Index** (HDI) is a composite indicator as a measure of economic development, comprising real income, life expectancy and educational attainment.

- **Healthcare** – This measures life expectancy at birth. The better the healthcare in a country, the greater the social and economic well-being tends to be for its citizens.

- **Education** – This indicator measures the mean (average) years of schooling and the expected years of schooling in the country. The higher the mean years of schooling for a country, the greater its level of economic development tends to be.
- **Income levels** – The measures the national income of a country (using purchasing power parity). As with the GNI per capita single indicator of development, the higher the income level the greater human development tends to be.

The three dimensions of the HDI are assigned equal statistical weights. The value of the HDI is between zero (extreme underdevelopment) and one (very high human development). For example, Norway has a HDI of 0.953 whereas Sierra Leone's HDI is 0.419. Economically least developed countries such as Niger, South Sudan, Chad, Burkina Faso and Mali have a very low Human Development Index. By contrast, more economically developed countries (MEDCs) such as Switzerland, Australia, Ireland, Singapore and Canada have a very high HDI.

Economic development can occur if there is a reduction in inequality and absolute poverty. This is reflected by an improvement in the HDI, but it does not mean that the country's GDP or GNI per capita necessarily increases. Similarly, although a country's GDP or GNI can increase without any improvement in the provision of healthcare and education, its HDI ranking will fall relative to its GDP/GNI ranking. These reasons make the HDI as a composite indicator a better measure of development than GDP/GNI per capita as a single indicator.

However, there are several limitations in using the HDI as a composite measure of development and standards of living. These drawbacks include:

- Qualitative factors – The HDI ignores qualitative measures affecting standards of living, such as gender inequalities and human rights.
- Income distribution – The HDI does not take account of inequitable income distribution, thus is less accurate in measuring living standards and human development for the 'average' person.
- Environmental issues – The HDI ignores the impact of environmental and resource depletion on development. For example, it ignores the negative externalities associated with increased output and consumption associated with economic growth, such as pollution and environmental degradation.
- Cultural differences – Although the HDI is a composite indicator of economic development, it ignores cultural differences in the interpretation of the meaning of 'standards of living' and 'quality of life'; both of which are vital aspects of human development.
- Sustainable development – The HDI ignores the concept of sustainable development (see Chapter 33), that is, consuming more today can jeopardize the standard of living for future generations owing to resource depletion and destruction of natural habits and ecosystems caused by higher levels of economic activity.

Theory of Knowledge

The HDI measures three dimensions of human development in terms of healthcare, education and income. How do we know that these are the most important aspects of development to include, and is there real meaning in them being equally weighted in the index to calculate human development?

PAPER 2 AND PAPER 3 EXAM PRACTICE QUESTION 34.3

Study the information below, which shows the Human Development Index for four countries: Australia, Ethiopia, Russia and Vietnam.

a Define the term *Human Development Index*. *[2 marks]*

b Identify the four countries based on the given HDI figures in the table and explain your reasoning behind the answers. *[4 marks]*

Country	HDI	Country	HDI
A	0.939	C	0.649
B	0.816	D	0.463

Common mistake

It is incorrect to assume that countries with a high GNI per capita have a high HDI. This is not necessarily the case. For example, Kuwait, Qatar and Liechtenstein have a higher GNI per capita than the USA and Hong Kong, but a lower HDI.

Gender Inequality Index (GII)

The UNDP's Human Development Report in 2010 introduced the **Gender Inequality Index** (GII), which calculates gender disparities as a measure of development. The GII is calculated by using three measures or dimensions of gender inequality – reproductive health, empowerment and labour market participation.

◆ The **Gender Inequality Index** (GII) is a composite measure of development by calculating gender disparities through three dimensions, namely reproductive health, empowerment and labour market participation.

Reproductive health

The first dimension of the GII composite indicator of development is reproductive health. This dimension uses two measures: the Maternal Mortality Ratio (MMR), which can be defined as the number of deaths of mothers per 100,000 live births, and the Adolescent Fertility Rate (AFR), which measures the number of births per 1,000 women aged between 15 and 19. Having children at an early age exposes young girls to medical problems, exploitation, social burdens and possible death caused by medical complexities. This could also discourage women from going to schools and colleges for education, subsequently impeding them from taking part in formal employment.

Empowerment

The second dimension of the GII is the empowerment, which is determined by two indicators: the share of parliamentary seats held by each gender and higher education attainment levels. The empowerment and reproductive health elements of the GII are interlinked as a higher participation and attainment in post-secondary education link to a lower adolescent fertility rate (AFR) and child mortality rate. Higher representation of women in parliament or government ensures that issues related to equality, equity, education, health and other developmental aspects are debated from alternative perspectives. This increases the awareness on the issues related to gender inequality and development, and also helps to formulate policies in favour of greater equality. Table 34.7 shows that female representation in parliament is more than half in Rwanda, Cuba, Bolivia and the UAE.

Table 34.7 Highest female representation in parliament, 2020

Rank	Country	Female participation rate (%)
1	Rwanda	61
2	Cuba	53
3	Bolivia	53
4	United Arab Emirates	50
5	Mexico	48
6	Sweden	47
7	Grenada	47
8	Andorra	46
9	South Africa	46
10	Nicaragua	45

Source: https://data.worldbank.org/indicator/SG.GEN.PARL.ZS

Labour market participation

Labour market participation of women is the third and final element of the Gender Inequality Index. The female labour market participation rate indicates the percentage of females in the country aged over 15 and below 64 who are employed or willing to work. Table 34.8 represents the top 10 countries with the highest female labour force participation ratios, based on data from the World Bank in 2019. In theory, the higher the labour market participation rate, the greater the level of household income and the lower poverty tends to be.

Note that the top 10 countries with the highest female labour participation rates are all low-income nations. They are also heavily dependent on the primary sector. Therefore, using the labour market participation rate itself cannot be a good measure of economic development.

Table 34.8 Highest female labour force participation ratios

Rank	Country	Value (%)
1	Rwanda	84.16
2	Madagascar	83.70
3	Nepal	81.69
4	Burundi	80.36
5	Tanzania	79.20
6	Zimbabwe	78.73
7	Mozambique	77.28
8	Lao PDR	76.82
9	Togo	75.83
10	Angola	75.35

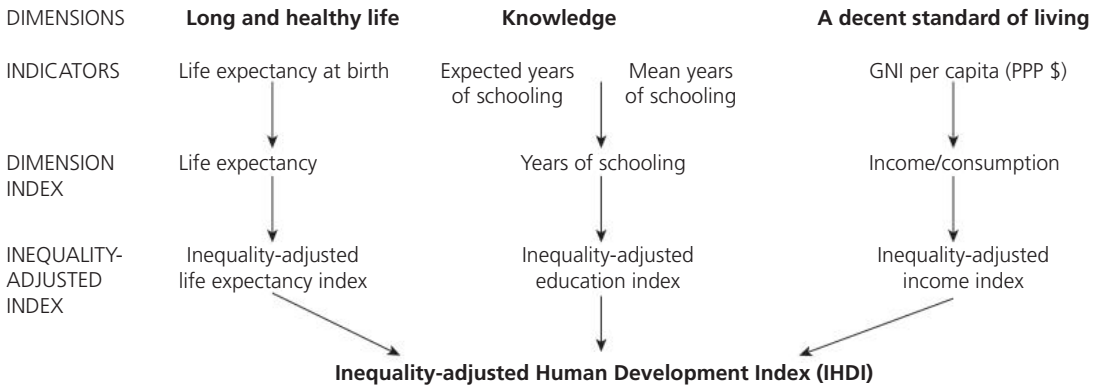
Source: https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS

Inequality-adjusted Human Development Index (IHDI)

The UNDP use the **Inequality-adjusted Human Development Index** (IHDI) as a composite measure of economic development for the average person in a society by accounting for inequalities. The IHDI combines a country's indicators of health, education and income with consideration of how these measures are distributed among the country's population. Hence, the IHDI measures the average level of human development of the average person by discounting (or accounting for) inequalities in society. The coefficient of inequality is calculated as an unweighted average of inequality across the three dimensions of the HDI (Figure 34.11).

The difference between the IHDI and HDI is the social and economic cost of inequality, that is, the overall loss to human development due to inequality. Under a situation of perfect equality in the economy, the IHDI is equal to the HDI. However, in reality, the IHDI is likely to be below the HDI when inequality rises. Some countries may have a high HDI, but a low IHDI when inequalities are accounted for. Examples include Liechtenstein, Saudi Arabia, the United Arab Emirates, Andorra, Qatar, Brunei, Bahrain, Oman, The Bahamas, Kuwait and Malaysia. This suggests that the average person may not enjoy such a high quality of life, as originally suggested by the HDI.

◆ The **Inequality-adjusted Human Development Index** (IHDI) is a composite measure of the average level of human development by accounting for inequalities in society.



Source: http://hdr.undp.org/en/content/inequality-adjusted-human-development-index-ihdi

Figure 34.11 The Inequality-adjusted Human Development Index (IHDI)

CASE STUDY

Gender equalities in Rwanda

Rwanda has enjoyed strong economic growth rates, creating new business prospects and lifting people out of poverty. Rwanda also ranks number one in terms of highest female representation in parliament and the highest female labour force participation ratio (see Table 34.8). However, Rwanda's HDI ranking is abysmally low. Rwanda's HDI value for 2017 was 0.524, which puts the country in the low human development category in a position ranked 158 out of 189 countries and territories.

■ **Table 34.9** Countries with highest IHDI, 2019 report

Rank		Country	Value	
IHDI	HDI		IHDI	HDI
1	1	Norway	0.889	0.954
2	6	Iceland	0.885	0.938
3	15	Japan	0.882	0.915
3	2	Switzerland	0.882	0.946
5	12	Finland	0.876	0.925
6	8	Sweden	0.874	0.937
7	11	Denmark	0.873	0.93
8	10	Netherlands	0.87	0.933
9	3	Ireland	0.865	0.942
10	6	Australia	0.862	0.938

Source: http://hdr.undp.org/sites/default/files/hdr_2019_overview_-_english.pdf

Notice that Japan, while ranked fifteenth in the HDI, is ranked third highest in the IHDI owing to favourable improvements in its Gender Development Index and Gender Inequality Index. Finland is in a similar situation, being ranked twelfth in the HDI but within the top five for the IHDI.

◆ The **Happy Planet Index** (HPI) is a measure of sustainable human well-being, that is, how individuals and countries are able to achieve long, happy and sustainable lives.

Happy Planet Index (HPI)

The **Happy Planet Index** is a measure of sustainable human well-being, that is, how individuals and countries are able to achieve long, happy and sustainable lives (also see Chapter 16). The traditional measures of economic development focus on how the welfare of citizens within a country can be maximized, failing to account for the way development today may impact on the development of tomorrow.

The Happy Planet Index (HPI) not only adopts a utilitarian approach (that is, that actions to maximize happiness and well-being on a personal level lead to maximum well-being of a country), but also considers the element of sustainable well-being to ensure a minimal trade-off between present development and development in the future.

■ **Figure 34.12** A happier world goes hand in hand with higher economic development

The HPI comprises four elements:

1. **Well-being** – how citizens of a country feel about their own life overall.
2. **Life expectancy** – the average number of years a person is expected to live.
3. **Inequality of outcomes** – a measure of the inequalities based on the distribution of the life expectancy and well-being data for the country.
4. **Ecological footprint** – the impact of the average person in the country on the environment.

The formula for calculating the HPI is:

$$\text{HPI} = \frac{(\text{Well-being} \times \text{Life expectancy} \times \text{Inequality})}{\text{Ecological footprint}}$$

Although the HPI is a composite indicator that accounts for the interest of future generations, it does have some limitations as an indicator of economic development. In particular, terms such as 'happiness' and 'well-being' are highly subjective and are difficult to measure as the responses

can change from day to day. Another criticism is that HPI is mistakenly considered as an indicator to measure personal happiness. In reality HPI is an indicator that measures the way citizens of a country could lead life in an ecologically sustainable way.

Strengths and limitations of approaches to measuring economic development (AO3)

Gross Domestic Product (GDP) is one of the most used indicators of the economic progress of a country over a prolonged period of time. However, there are many limitations in accepting this as a single indicator or measure of economic development as it does not cover every individual of the country. While indicators of income can be tweaked, such as the use of real GDP per capita expressed in terms of purchasing power parity to give a more accurate representation of development, they fail to consider qualitative factors that influence quality of life and economic development. At the beginning of this chapter, the multidimensional nature of economic development was discussed, including how the idea of economic development evolves over time. Given the diverse nature of development, any single measure of development will have its shortcomings.

It is therefore quite evident that economists and policymakers tend to use more than one indicator to measure economic development. After all, the best way to measure development could be very different for different members of society, so the accuracy of the various methods used to measure development is highly contested. For example, different countries place different priorities on environmental issues or gender equality. Furthermore, the UNDP's SDGs, introduced in 2015, have influenced economists and policymakers in numerous discussions and speculations about the approaches used to measure economic development.

The following points are relevant to addressing and evaluating the different approaches to measuring development, and whether these are a success or a failure:

- Measuring economic well-being is a multilayered, complex process that changes over time. For example, the number of indicators used to measure development as part of the UNDP's Millennium Development Goals increased significantly when they were replaced with the Sustainable Development Goals in 2015 (see Chapter 33).
- Some of the indicators may better portray a particular strength of a country, but this does not necessarily reveal the full picture. For example, Madagascar ranks second highest in female labour participation rates, but is ranked 162nd in the HDI. By contrast, the United Arab Emirates has a relatively high HDI, being ranked 34th, but has a female participation rate of only 52 per cent.
- Some of the indicators used to measure development (such as healthcare and education) appear in different measures of development (such as the HDI and GII). This makes some economists and policymakers speculate or question whether these indicators have any deeper and more significant relevance to measuring economic development.
- Economic development is measured using qualitative factors, rather than only measuring economic growth using a quantitative approach. Nevertheless, governments, statisticians and economists all attempt to quantify these qualitative factors, including measures such as the IHDI, HPI, life expectancy and mean years of schooling. The statistical weights used in some of these measures are applied arbitrarily.

ATL ACTIVITY 34.7

On a piece of paper, write down four ingredients of individual happiness. Collate the responses from the whole class and share the four most frequently recurring responses. Now, compare the collated answers with the elements of the Happy Planet Index (HPI). To what extent do the answers from your class support the four elements of the HPI (well-being, life expectancy, inequality of outcomes and ecological footprint)?

- Economic development cannot be delinked from the political, cultural and social aspects of the country in question. This makes measuring development difficult to quantify and subjective. For example, the Second Amendment to the United States Constitution, passed in 1791, protects the individual right to keep and bear arms. This is deemed as important to the American people, although in most other parts of the world there are alternative views on the rights to owning armed weapons.
- Political and social conflicts, corruption and external shocks make it challenging to access data in many countries in different times to measure development. Examples include internal and international conflicts, major natural disasters, terrorism, war and the spread of endemic diseases, such as HIV/AIDS and coronavirus. Some countries are also more prone to such external shocks than others.

Given that all countries are different and diverse in their own ways, there are various attributes to development that do not always apply in the same way across nations and regions of the world. Therefore, economists and policymakers can find it challenging to find one distinct measure to form the basis of measuring and comparing development. Nevertheless, the fact that there are numerous ways to measure development that interrelate and complement one another make such measures a strength, be they single or composite indicators of economic development.

Possible relationship between economic growth and economic development (AO3)

Economic growth is clearly defined by economists by using quantitative measures (see Chapter 18). Essentially, economic growth is the increase in the value of a country's real gross domestic product (GDP) per capita over time, which is a quantitative indicator of economic well-being. Economic development includes economic growth in addition to qualitative determinants of quality of life. Hence, development is a much broader and deeper measure of economic well-being, as it includes qualitative indicators such as a reduction in poverty, income inequality, gender inequality, political oppressions and long-term unemployment. Economic development is multidimensional in nature, encompassing tangible and intangible factors that raise the general standard of living in a country. It is generally accepted that economic growth is a precondition to achieving economic development, although this is not always necessary.

Figure 34.13 Approaches to measuring and addressing development are more than just about monetary measures

Figure 34.14 Not all households have benefited from Hong Kong's economic growth

There is a positive relationship between a country's economic growth and its economic development. For example, the wealthier a country is, the higher its standard of living tends to be (see Table 34.9). However, it is possible for some people to live happier lives in the absence of higher incomes as economic growth can bring about negative consequences such as pollution, climate change and environmental damage. In addition, economic growth does not always lead to a higher standard of living for the majority of people, so there is no economic development in such cases. For example, economic growth can be caused by productivity gains due to the increased use of technology and automation (such as self-service checkouts in supermarkets and other retail outlets), but this results in technological unemployment. Although growth is a necessary component of development, it does not necessarily always lead to economic development. Another example is if a country's GDP has increased due to spending significantly more on national defence then growth may occur but not necessarily development. Economic growth can also lead to greater income inequality in the economy (as in the case of Hong Kong, which has one of the world's largest wealth gaps).

PAPER 2 AND 3 EXAM PRACTICE QUESTION 34.4

Assume an economically less developed country has a real GDP per capita of $2,000 and real GDP per capita is $35,000 in a higher income country. If both countries have economic growth of 2.5%, calculate the change in the per capita income gap between the two countries. *[2 marks]*

Common mistake

It is incorrect to use the terms 'economic growth' and 'economic development' interchangeably. Economic growth is often regarded as a prerequisite to economic development, but there are certainly far more factors that contribute to development than simply an increase in real incomes for the average person in the country.

Economic development is also possible without achieving economic growth in the short run if policies are directed towards a more equitable distribution of income, wealth and resources to all the members of the society. However, sustained economic growth is generally regarded as a prerequisite to achieving a higher standard of living. This can be explained with the help of a production possibility curve (PPC) diagram as illustrated in Figure 34.15.

- Moving the economy from P to Q shows no economic growth but some development.
- Moving the economy from P to R shows no economic development but some growth.
- Moving from P to S shows both economic growth and economic development.

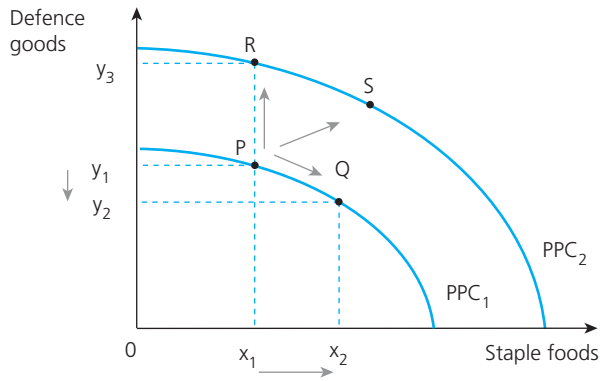


Figure 34.15 Economic growth and development using the PPC diagram

In Figure 34.15, the country in question can direct its resources to the production of staple foods and/or defence goods. Point P on the production possibility curve PPC_1 indicates the country can produce $0y_1$ units of defence goods and $0x_1$ units of staple foods. Assuming there are no changes to the resources available, the state of technology or level of productivity, economic development can be achieved if the economy moves from point P to point Q, as people will be able to enjoy more units of staple foods (which increase from $0x_1$ to $0x_2$), sacrificing manufactured defence goods (which fall from $0y_1$ to $0y_2$). Essentially, economic development can occur in the short run when countries reallocate more resources to provide goods and services that reach a greater number of people in the economy, including the poorest members of society, and cut expenditure on production of goods and services that benefit only a minority of citizens.

An outwards shift of the production possibility curve from PPC_1 to PPC_2 indicates that the country is now able to produce more staple foods and/or defence goods, as a result of innovation, technological improvements or a rise in productivity. The outwards shift of the PPC indicates higher production of both staple foods and defence goods due to economic growth. However, economic growth does not always result in an increase in living standards, shown by the movement from point P to point R. At point Q, prior to economic growth, the economy was producing $0x_2$ units of staple foods and only y_2 units of defence goods. If the economy achieves growth and reallocates production in favour of defence goods by moving from point Q to point R on PPC_2, this means the country produces fewer units of staple foods ($0x_1$) in favour of a greater number of units of defence goods ($0y_3$). Hence, growth is achieved without any contribution to economic development.

However, if the economy chooses to move from point P to point S, this indicates the ideal situation when economic growth also creates economic development. In this case, the economy can enjoy more output of both defence goods and staple foods by operating at point S on PPC_2. In any case, economic growth is usually needed over the long-term for there to be economic development. After all, to meet their infinite wants (desires), people need to earn more income.

WISE ChoICES

Using real-world examples, explain the types of **intervention** a government should initiate for a country that experiences high economic growth but low economic development.

Theory of Knowledge

What knowledge questions should be considered if governments are to decide on a given choice between achieving economic development without economic growth, or the other way around?

Chapter summary

- Economic development is multidimensional in nature, as there are many facets to what contributes to the economic and social well-being of individuals and societies.
- The concept of economic development has evolved during the past few centuries, from the perception that equates to economic growth to one that considers the ability to live without any social, economic and political constraints.
- Economic development can be measured by several single indicators including GDP or GNI per capita in purchasing power parity terms, health and education indicators, economic and social inequality indicators, energy indicators and environmental indicators.
- GDP per capita is the value of national output produced in a country divided by the number of people living in the country. It is the most commonly used single indicator of economic development.
- A common variation used to measure economic development is GDP per capita at purchasing power parity. Purchasing power parity (PPP) equates the cost of living between countries to enable more meaningful comparisons of standards of living across different countries.
- Health indicators are a single indicator of measuring development through health-related statistics, such as life expectancy at birth, infant mortality rates and expenditure on healthcare as a percentage of GDP.
- Education indicators of development involve measuring education-related determinants that affect standards of living, such as literacy rates and the mean (average) years of schooling.
- Some important environment indicators of development are measures to tackle critical environmental issues like global warming, ozone depletion, desertification, deforestation, disposal of wastes, and loss of biodiversity and ecosystems.
- Economic development, which increases the quality of life for all members of society, can also be measured by using composite indicators such as the Human Development Index (HDI), Gender Inequality Index (GII), Inequality-adjusted Human Development Index (IHDI) and Happy Planet Index (HPI).
- Measuring economic growth and economic development are complex and different approaches are taken to measure these. Growth generally brings about economic development but cannot ensure it.
- Many countries with high economic growth fare badly in many developmental indicators. Nevertheless, economic growth generally helps a country to achieve economic development.

35 Barriers to growth and/or economic development

Real-world issue

Why is economic development uneven?

Conceptual understandings (WISE ChoICES)

- Governments and other economic agents may **intervene** in an attempt to promote **economic well-being** and **equity** in societies.
- The pursuit of **sustainability** is subject to various constraints.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:
- poverty traps/poverty cycles (AO2)
- a poverty cycle diagram to show any linked combination of factors that perpetuate poverty (AO4)
- economic barriers to growth and/or development (AO2)
- political and social barriers to growth and/or development (AO2)
- the significance of different barriers to economic growth and/or economic development (AO3).

Poverty traps/poverty cycles (AO2, AO4)

◆ The **poverty trap** (or **poverty cycle**) is a vicious cycle of poverty and deprivation causing even greater poverty, from one generation to the next.

Economists attribute different economic, social and political barriers as the reasons for poverty. One of the main reasons for poverty is the existence of poverty itself, that is, the poor get poorer. This is known as the poverty trap or poverty cycle. The **poverty trap** (or **poverty cycle**) is a vicious cycle of poverty and deprivation causing even greater poverty, from one generation to the next. Low-income earners spend most, if not all, of their income in order to meet their essential needs but this also means they have insufficient funds to invest in their future so are trapped in poverty (see Figure 35.1).

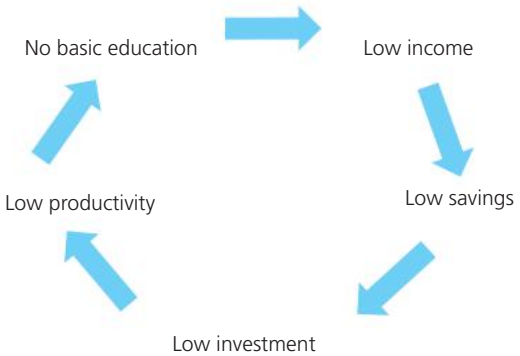


Figure 35.1 The poverty trap

Figure 35.1 shows how poverty caused by low income and/or no basic education leads to a cycle of poverty. The lack of basic education for the poorest members of society prevents them from getting into the mainstream or formal workforce because they lack the skills and requisite knowledge and understanding to get a job. Hence, the lack of basic education is directly linked to low incomes of individuals and households. Therefore, they may struggle to earn sufficient income to meet their day-to-day expenditure needs.

Having a low income (real GDP per capita) also means a low **savings ratio**. This is the amount of savings expressed as a proportion of total disposable income. A low savings ratio means that these low-income countries have limited funds from savings to fund investment expenditure on basic infrastructure to support education and healthcare. The low savings ratio therefore hinders their

◆ The **savings ratio** is the amount of savings expressed as a proportion of total disposable income in an economy.

future productive capacity and economic development. This leads to a lower level of productivity in the economy due to low skills and poor healthcare provision.

Individuals, households and countries suffering from low income are often unable to break the poverty trap owing to the vicious cycle that keeps them in poverty. This is because there is little, if any, savings in the economy that can be used to fund the necessary investment needed for sustainable economic development.

American economist Professor Jeffrey Sachs (b.1954) suggests there are six major types of capital necessary to fund the investment needed for economic development. ELDCs lack the following interrelated categories of capital assets, thereby forcing these countries to remain in a poverty trap:

- Physical capital, such as tools, equipment, machinery and buildings.
- Human capital – ELDCs typically suffer from low levels of education and skills, and poor health, which will have a harmful impact on rates of infant mortality, maternal mortality and life expectancy.
- Natural capital – the stock of indispensable natural resources from the Earth's ecosystems are depleted and are unable to be replenished without the necessary capital investments to sustain their use (see Chapter 33).
- Business capital – the financial assets needed for a firm to produce goods and/or services to sell for earning profit.
- Public institutional capital – government owned assets and infrastructure to aid productivity, such as transportation networks, telecommunications networks and public utilities, including sewage systems.
- Knowledge capital – all intangible attributes that contribute to development, such as ideas, creativity and innovations.

In his seminal book *The End of Poverty* (2005), Jeffrey Sachs describes the process of the poverty trap spanning generations as follows:

> *The poor start with a very low level of capital per person, and then find themselves trapped in poverty because the ratio of capital per person actually falls from generation to generation. The amount of capital per person declines when the population is growing faster than capital is being accumulated... The question for growth in per capita income is whether the net capital accumulation is large enough to keep up with population growth.*
>
> *'The End of Poverty' (2005) Jeffrey Sachs*

Figure 35.2 Economist, Jeffrey Sachs, suggested there are six major types of capital investments necessary for development

Low-income individuals and households also struggle to break the poverty trap because banks are unlikely to lend money to very poor people as there is a high risk of them defaulting on the loans, that is, failing to repay the money borrowed owing to their financial position. It is also common that poverty is transmitted from generation to generation, so a poor family that is unable to access education has a relatively low chance of getting out of poverty. Very poor individuals are often unable to feed their family, which also leads to malnutrition and physically weak labourers.

Government intervention is therefore required to bring people out of extreme poverty. This is because they are unable to do so on their own due to the poverty trap. For example, the government can invest in education and healthcare to increase the quality of human capital or improve employment opportunities for females. In the case of highly indebted poor countries (HIPCs), where the nation is so poor that government intervention is not possible, then it will need to rely on foreign aid (see Chapter 36). Many economists and policymakers argue that foreign aid and grants from

Figure 35.3 Government intervention is essential to get people out of the poverty trap

more economically developed countries (MEDCs) could make up for the lack of capital in HIPCs and ELDCs, which could eventually help them to escape the poverty cycle.

Theory of Knowledge

Here is a classic chicken or egg quandary. With reference to the poverty cycle, discuss whether low income leads to low levels of savings or whether it is a result of this. To what extent do you need money to earn money?

CASE STUDY

Live Aid

After the release of the film *Bohemian Rhapsody* in 2018, a new generation took interest in the Live Aid concert, which took place in 1985. The event was planned to raise money to help ongoing famine in Ethiopia, in the Horn of Africa. The show was a big hit, simultaneously taking place at Wembley Stadium in London, UK, and John F Kennedy Stadium in Philadelphia, USA. It was one of the most widely televised broadcasts of all time, with an estimated audience of 1.9 billion people, across 150 nations, making it the largest show on the planet. The show has seen an unparalleled line-up of performances, compared only to the famous Woodstock music festival in 1969, and included some of the biggest names in the industry – such as David Bowie, Elton John, Sting, Led Zeppelin, The Who, U2, Paul Young and Queen. Roughly £150 million ($195m) was raised as a result of the massive fund-raising operations. Although the Live Aid concert was not devoid of controversies or criticisms, this effort of helping the persistent famine is considered to be a form of intervention to break the poverty cycle.

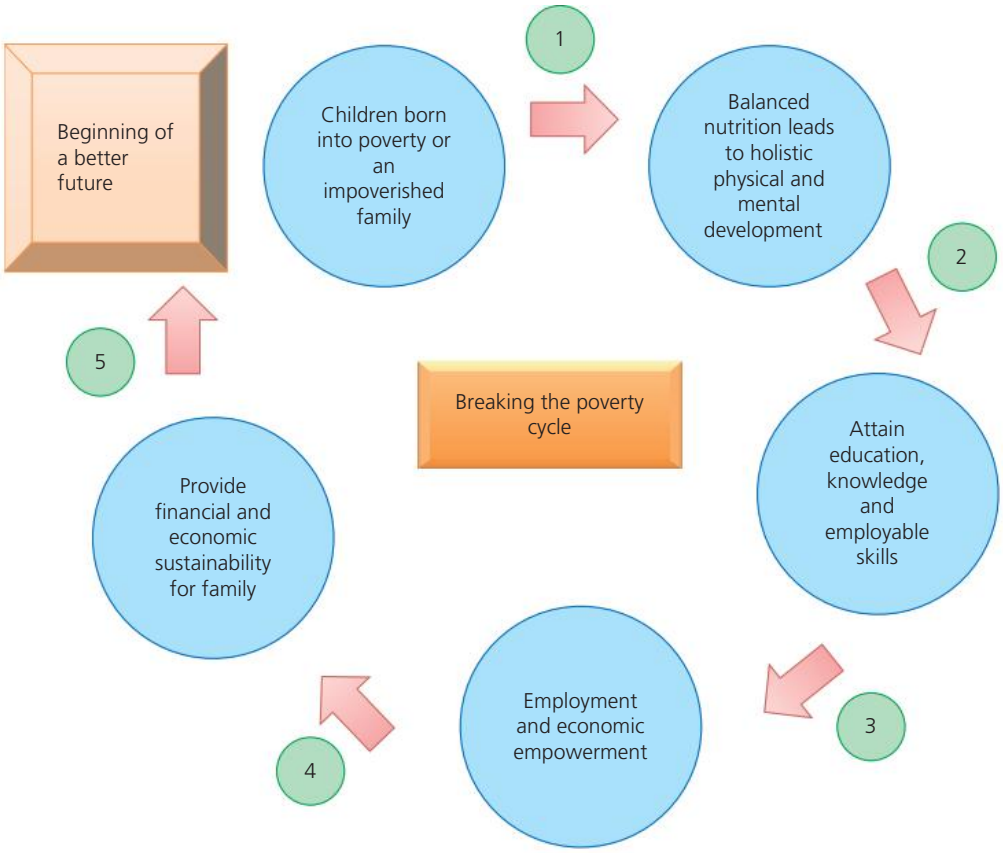


Figure 35.4 Breaking the poverty cycle

Figure 35.4 illustrates how external intervention in the form of government policies or private initiatives can help to break the poverty cycle. Government intervention can offer a child born into an impoverished family the opportunities of nutritious food and education. This helps to ensure the child has a chance to access a good level of education and skill development, thereby increasing the child's employability in the future. Having a paid job empowers individuals not only with income to meet their own daily needs but also the livelihoods of their family members and savings that could be used to provide nutritious food and education to the next generation, all without the need for further government assistance.

ATL ACTIVITY 35.1

In a group of two or three, select one poor country, such as an HIPC or an ELDC. Research the various causes or barriers to the low level of economic growth and development in the country. Identify the most important barrier to economic growth, and compare and contrast your findings with other groups in the class.

Economic barriers (AO2)

Barriers to economic growth and development can be categorized as two broad factors – economic barriers and political and social barriers. The IB Economics syllabus specifies ten economic barriers to economic growth and/or economic development. These are: rising economic inequality, lack of access to infrastructure and appropriate technology, low levels of human capital (lack of access to healthcare and education), dependence on primary sector production, lack of access to international markets, the informal economy, capital flight, indebtedness, geography (including landlocked countries), and tropical climates and endemic diseases.

Rising economic inequality

The most widespread measure of economic inequality is income distribution in the economy, whereby the national income of a country is shared out among its population. An alternative measure is wealth distribution, that is, how the national wealth of a country is distributed between all members of the society. Other widely used measures of rising economic inequality include the Gini coefficient (see Chapter 23) and the Inequality-adjusted Human Development Index (see Chapter 34).

In 1820, the ratio between the income of the top and bottom 20 per cent of the world's population was 3:1. By 1991, it was 86:1. Unequal distribution of national income exists when the relatively rich minority accounts for the majority of the country's national income, that is, the economy's GDP is not distributed proportionately. Hence, economic growth does not ensure economic development if the poorer majority of society does not have sufficient access to education, healthcare and capital (credit), which leads to further inequalities in the distribution of income and wealth.

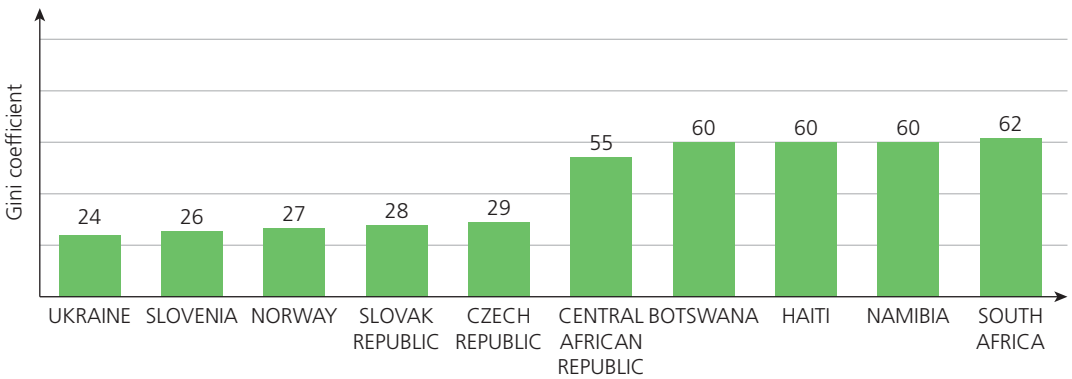
Figure 35.5 shows that the five most equal countries are from Europe and the majority of the five most unequal countries are from Africa, based on World Bank data on the Gini coefficient. A higher Gini index indicates greater inequality, with high-income individuals receiving a much larger percentage of the total income of the population.

A high and rising degree of economic inequality in many ELDCs is regarded as a major barrier to economic development. After all, the eradication of poverty does not only depend on growth in national income but how that income is adequately distributed. A more equal distribution of national income will mean that those in extreme poverty are more likely to be able to access education and healthcare. This helps to improve human development and raise the level of productivity in the economy. Greater income equality also reduces the likelihood of corruption, which is a major obstacle to human development in many ELDCs. Corruption and civil unrest distort market forces and create disincentives for individuals and firms to take entrepreneurial risks and investments.

The main way to redistribute income from the rich to the poor is to implement an effective progressive tax system (see Chapter 23) by increasing marginal tax bands as income levels rise. This will mean that those earning higher incomes are taxed at a higher rate of tax, thereby redistributing income to lower income earners and supporting those from the poorest households.

WISE ChoICES

Given that many ELDCs have experienced a high and rising degree of economic inequality, discuss whether change helps to improve equity and economic well-being in societies.



Source: adapted from https://databank.worldbank.org/reports.aspx?source=2&series=SI.POV.GINI&country=

Figure 35.5 Top and bottom five countries as measured by the Gini coefficient

However, there are three key reasons why it is difficult for ELDCs to raise tax revenues:

- Only a small proportion of the population pay income taxes, as so many people are poor and because ELDCs tend to have a large unofficial (parallel) market.
- Similarly, as the level of official economic activity is low, the revenue collected from corporation tax tends to be low. This is especially the case if ELDCs offer financial assistance to domestic firms and/or tax incentives to encourage foreign direct investment (FDI).
- Low incomes and the lack of international trade in many ELDCs mean there is minimal tax revenue earned from tariffs (taxes on imports).

Ineffective tax systems, coupled with a high degree of corruption, in many ELDCs have hindered income redistribution and, hence, economic and human development.

PAPER 1 EXAM PRACTICE QUESTION 35.1

Explain how economic, political and social factors work as barriers to economic growth and development. *[10 marks]*

Lack of access to infrastructure and appropriate technology

Infrastructure refers to the physical structures and facilities required for the efficient running of a country. This includes:

- transportation networks, such as ports, airports, bridges, tunnels, roads, railroads and highways/motorways
- telecommunications networks
- schools and hospitals
- social infrastructure such as parliament, courts, regulatory bodies and prisons
- public utilities, including energy, power supplies, water supplies, dams and sewage systems.

Most of a country's infrastructure is provided by the government. A good infrastructure system enables greater productive capacity, leading to economic growth and development in the long run. However, many ELDCs have very limited infrastructure, which reduces efficiency and productivity. It also makes it far more challenging for the country to engage in international trade and to attract foreign direct investment. For example, a country with poor road and transportation facilities may suffer from problems with food distribution as it will take more time to transport food products to consumers. A lack of proper road networks will slow down all aspects of economic activity. Moreover, children will take more time to get to and from school while patients

Figure 35.6 Infrastructure is one of the key drivers of growth and development

will find it difficult to get to a hospital for their medical needs. Hence, poor infrastructure in a country will obstruct its economic development.

The lack of access to appropriate technology is a related key economic barrier to development. At the most basic level, the inability to use essential tools and equipment or the lack of access to power limits the country's industrial production. As a result, agricultural production is labour intensive, inefficient and time consuming. This could also discourage international companies from investing in the country, resulting in missed opportunities to attract foreign direct investments (FDI). Limited access to power in many ELDCs also limits or discourages primary education and the country's ability to offer adequate healthcare services. In the case of Nepal, the country's tourism industry has been adversely affected as a result of frequent power cuts.

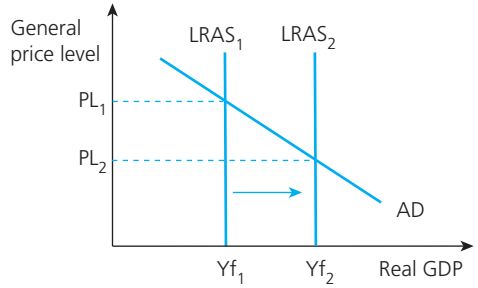


Figure 35.7 Impact of greater access to infrastructure and appropriate technology on productive capacity of the economy

Economic activity in ELDCs relies more on labour-intensive methods to create jobs and income for the population of the country. However, the use of undeveloped labour-intensive technologies (such as basic tools and equipment such as bicycles, wheelbarrows and hand-powered water pumps) is not sufficient for the country to escape extreme poverty. By contrast, capital-intensive technologies rely on the use of physical capital in the production process. For example, workers in ELDCs might be able to use sewing machines enabling them to be far more productive than doing the same tasks by hand.

In Figure 35.7, advances in technology and the use of appropriate technologies lead to a favourable change in the productive capacity of the economy. This is shown by an outwards shift of the economy's long run aggregate supply from $LRAS_1$ to $LRAS_2$, thus increasing national income from Yf_1 to Yf_2. However, many ELDCs suffer from a lack of infrastructure, such as transportation and telecommunications networks, and access to appropriate technologies. This therefore hinders their chances of economic growth and development.

Top tip!

Remember that not all economic growth leads to economic development. For example, technological progress that boosts capital productivity and causes mass unemployment (technological unemployment) can create economic growth, but this does not equate to human development.

Low levels of human capital – lack of access to healthcare and education

A lack of access to healthcare and education are major barriers to economic growth and development. Recall that a low level of education and poor access to healthcare provision are reasons that trap people in poverty. Zimbabwe has one of the highest rates of HIV and one of the highest maternal mortality rates.

The quality of labour is largely dependent on the standard of healthcare, educational attainment, skills development and training. Hence, investments in education and healthcare will shift an economy's long run aggregate supply curve to the right, ceteris paribus. This can be seen in Figure 35.7. Alternatively, the increase in the economy's productive capacity can be shown by using a production possibility curve diagram, as shown in Figure 35.8.

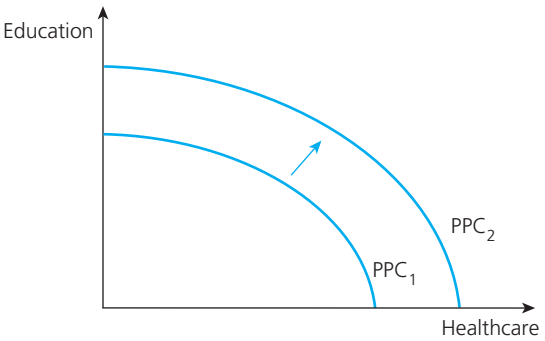


Figure 35.8 Impact of greater access to education and healthcare on development

Investment in education and healthcare are prerequisites for economic development. Investment expenditure on education and healthcare will improve the quality of labour and increase labour productivity, thus shift the economy's productive capacity outwards from PPC_1 to PPC_2, ceteris paribus.

According to children's charity UNICEF, educated women are at least twice as likely to encourage their children to attend school at primary level. In some cultures, parents prefer boys to attend school while girls are expected to remain at home to help with household responsibilities.

In 2000, the United Nations Human Development Report: Education Index indicated that 9 out of 10 countries with the highest percentages of children denied access to school were in Africa. In 2017, almost two decades later, the 10 lowest-ranked countries in the HDI rankings were all African nations, all with a low Education Index (Table 35.1). The Education Index is calculated using the mean years of schooling and expected years of schooling, with the range from 1.0 (very high rate of access to education) to 0.0 (extremely low rate of access to education). There is clearly a correlation between a country's level of development (as indicated by its HDI ranking) and the Education Index.

Table 35.1 The UNHDR: Education Index – lowest ranked countries

HDI Rank	Country	UN Human Development Education Index
178	Mozambique	0.372
179	Guinea	0.294
180	Burundi	0.370
181	Burkina Faso	0.250
182	Eritrea	0.228
183	Sierra Leone	0.305
184	Chad	0.256
185	Central African Republic	0.318
186	Congo (Democratic Republic of the)	0.372
187	Niger	0.198

Source: **http://hdr.undp.org/en/content/education-index**

Top tip!

Remember that education and health are important aspects of the United Nations Sustainable Development Goals (see Chapter 33), such as eradicating extreme poverty and hunger, reducing child mortality, improving maternal health, combating diseases such as HIV/AIDS, universal access to quality education for all, and gender equality.

Dependence on primary sector production

The primary sector of an economy is involved with the extraction, cultivation and collection of natural resources. Examples include agricultural farming, animal husbandry, mineral extraction, commercial fishing and forestry.

All nations were dependent on primary sector production before the Industrial Revolution. In the late eighteenth century, countries like Britain, Germany and the USA prospered owing to a shift from primary sector production to the secondary sector (manufacturing) output brought about by the production processes that involved the introduction of machinery, steam power and assembly lines. In the twentieth century, European countries and the USA achieved growth and development through diversifying into the tertiary sector.

The data in Table 35.2 show the composition of gross domestic product (GDP) for the different sectors of the economy, ranked by their level of involvement in the primary sector expressed as a percentage of GDP. Unsurprisingly, countries that are highly dependent on primary sector production tend to be the ones with the lowest human development indicators. Hence, it is evident that the dependence, or over-reliance, on primary sector production is a barrier to a country's economic development.

Table 35.2 Composition of GDP by production sector, 2017

Country	Agricultural (% of GDP)	Secondary (% of GDP)	Services (% of GDP)
Sierra Leone	60.70	6.50	32.90
Somalia	60.20	7.40	32.50
Chad	52.30	14.70	33.10
Guinea-Bissau	50.00	13.10	36.90
Comoros	47.70	11.80	40.50
Central African Republic	43.20	16.00	40.80
Mali	41.80	18.10	40.50
Niger	41.60	19.50	38.70
Sudan	39.60	2.60	57.80
Burundi	39.50	16.40	44.20
Ethiopia	34.80	21.60	43.60

Source: The World Factbook, Central Intelligence Agency, https://www.cia.gov/library/publications/resources/the-world-factbook/fields/208.html

Figure 35.9 The majority of Sierra Leone's GDP comes from the primary sector

Most of the resources used in the primary sector are exhaustible, so cannot be sustained in the long run. As such, the resources used to produce primary sector output will eventually dry up if the economy is overly dependent on primary sector production. A country rich in minerals may be completely dependent on the production and export of minerals or crude oil and get to enjoy a high national income in the short-term but will eventually run out of these resources in the long-term. This is why countries like Qatar and the United Arab Emirates have been investing very large sums of money in infrastructure, education and healthcare so that the country can continue to experience economic growth and development in long-term. In addition, as primary sector output is usually dependent on the climate and weather conditions, production is susceptible to natural disasters, climate change and global warming.

Lack of access to international markets

Market access is the ability of a nation to sell its goods and services across borders and to enter foreign markets. Globalization and trade liberalization policies (see Chapter 33) during the 1980s helped many economically underdeveloped countries to foster trading relationships with other countries. Subcontracting and outsourcing brought in a new era of global trade where multinational companies from more economically developed countries (MEDCs) looked overseas for cheaper and more efficient production methods, especially in countries where labour costs were substantially lower than in their home nations. Countries including Bangladesh, Brazil, China, India, Indonesia and Vietnam reaped the benefits from the increase in global trade in goods and services. Countries with highest value export earnings tend to have a greater level of economic development owing to their ability to access international markets (see Table 35.3).

◆ **Market access** refers to the ability of a country to enter foreign markets to sell its goods and services.

However, the intensity of global competition from MEDCs and their various protectionist measures have made it difficult for many ELDCs to compete, that is, the lack of access to international markets means it is challenging for low-income countries to trade with more economically developed countries. Governments from these countries and regions, such as the USA or European Union, impose various barriers to entry in the form of tariffs, embargoes, quotas and administrative barriers. Such practices restrict the ability of ELDCs access to many international markets, so limit their ability to develop.

Table 35.3 Countries ranked by highest value of exports, 2019

Rank	Country	Value of exports ($'000)	Share of world exports (%)
1	China	2,498,569,866	13.3
2	USA	1,645,174,335	8.8
3	Germany	1,486,462,772	7.9
4	Netherlands	721,301,085	3.8
5	Japan	705,842,013	3.8
6	France	555,100,606	3.0
7	South Korea	542,333,337	2.9
8	Italy	541,695,309	2.9
9	Hong Kong SAR, China	535,711,019	2.9
10	Mexico	472,272,871	2.5

Source: International Trade Centre (https://bit.ly/2K7yqyh)

A loss of trade and potential income due to the lack of access to international markets leads to lower potential development. Income from export earnings is vital to fund a country's purchase of imports of capital goods (required for production and expanding the country's infrastructure) and consumer goods and services. Low-income countries also lose out because the inability to access international markets takes away the opportunity for these countries to specialize in the export of a particular good or service. Specialization results in better quality output and economies of scale, leading to higher profits as a result of the lower average costs of production.

Informal economy

The **informal economy** (or **parallel markets**) can be defined as the various economic activities that are not officially recorded as part of a country's gross domestic product (GDP). The activities of businesses and workers in the informal sector of the economy are not monitored or protected by the government as these economic activities are not officially recorded so are not formally known to the government and so such activities are not liable for tax. Note that not all activities in the informal economy are necessarily illegal, such as subsistence farming, self-employment in small, unregistered business activities, and workers in casual jobs who are paid in cash and for which employment records are not kept.

◆ The **informal economy** (or **parallel markets**) can be defined as the various economic activities that are not officially recorded as part of a country's gross domestic product (GDP).

A study by the International Labour Organization (ILO) in 2018 revealed that 61 per cent of all workers, that is, 2 billion people worldwide, are engaged in some kind of economic activity in the informal economy. However, this issue is more apparent in ELDCs (see Table 35.4).

Table 35.4 Informal sector employment as a percentage of rural and urban employment

Countries by income level	Total	Rural	Urban
World	61	80	44
Developing	90	90	79
Emerging	67	83	51
Developed	18	22	17

Source: International Labour Organization (ILO) in 2019 (https://bit.ly/2xqzdYC)

Although the informal economy exits in every nation, it is clear from Table 35.4 that informal sector employment is highest in developing countries, accounting for up to 90 per cent of total

employment. A significant proportion of this is contributed by the large number of female labourers who work in the informal sector, such as agriculture.

However, the informal economy is a disruption to the country's economic growth and a major obstacle to achieving development because unofficial economic activities are very difficult to monitor, regulate and tax. The individuals and enterprises involved in the informal economy enjoy many of the services offered by the government but do not pay any form of income tax to help contribute towards the expenditure incurred by the government. Hence, the government has less tax revenue than otherwise to support and facilitate economic development in the economy.

Nevertheless, the informal sector can provide important employment opportunities for the poorest members of society, such as women and other members of the underprivileged population. These vulnerable groups are not represented in the formal sector of the economy because they do not have the required level of education, experience or skills to find meaningful employment. However, most economists argue for a continual reduction in the informal sector because it creates some substantial and ethical challenges such as the issue of child labour, or the exploitation of workers in ELDCs. This is of particular concern owing to the nature of the informal sector, so employees are not represented or protected by any form of trade union. According to the International Labour Organization, 'Informal economies are typically characterized by a high incidence of poverty and severe decent work deficits. Without formalization, decent work for all and equity in society will remain an illusion.'

Capital flight

Capital flight occurs when businesses take their money and resources out of the country owing to economic and political uncertainties. As it involves the large-scale withdrawal of assets and capital from a country, this can be extremely harmful to the host country. Capital flight is a cause and consequence of a lack of investor confidence in the economy, so can be a major hindrance to economic development. Capital flight is often associated with countries where there are culturally embedded political and economic problems such as Iran, Syria, Venezuela and Zimbabwe, and countries troubled by civil war and international conflict.

The main interconnected reasons for capital flight and the lack of investor confidence are:

- political turmoil, upheaval and instability
- social unrest and civil conflict, and a subsequent increase in crime rates
- economic instability and concerns over the country's banking system due to a weak credit rating
- exchange rate volatility and uncertainty
- expectations of the nationalization of privately owned assets
- anticipation of higher taxes being imposed on income, wealth and assets
- environmental degradation and destruction, and the resulting air, land and water pollution.

Capital flight can be classified as legal or illegal. *Legal capital flight* takes place in the form of repatriation of the capital or assets by a foreign entity. *Illegal capital flight* (or *illicit capital flight*) occurs when money is taken out of the country using dishonest or unlawful means, such as money laundering. Illegal capital flight can affect any country, but ELDCs are more prone to illegal capital flight owing to weaker regulations and high levels of corruption.

Capital flight is considered undesirable as the withdrawal of assets and investments in the country impedes the nation's growth and development. Capital flight can have negative ripple effects, giving the country a reputation thereby discouraging other multinational firms from

Common mistake

It is incorrect to assume that all trading activities in unofficial (parallel) markets are illegal. While these are unrecorded, many of these trades in ELDCs may be in the form of barter (swapping) or payments made in kind (using goods and services to trade rather than money). Hence, unofficial trade is not necessarily illegal trade, but has a detrimental impact on economic development.

◆ **Capital flight** is the withdrawal of money, assets and resources from a country owing to economic and political uncertainties, thereby hindering investor confidence and economic development.

investing in the country. Hence, capital flight also leads to unemployment and a reduced national income. It is particularly harder to stop in ELDCs.

Capital flight typically threatens the stability of the country's currency on foreign exchange markets, which results in a lower exchange rate. As a result, the country would need to pay more for essential imported goods and services. An ELDC will depend on the import of tools, equipment and machinery but capital flight would hurt the purchasing power of the country. Furthermore, capital flight could leave the country with fewer foreign currency earnings, which might impede their ability to finance investment expenditure in the economy.

Figure 35.10 Capital flight is the withdrawal of money, assets and resources from a country due to a lack of confidence

CASE STUDY

Asia's capital flight

Following the Asian Financial Crisis in 1997, international investors lost confidence in the economies of Thailand, Indonesia, Taiwan, South Korea and, in particular, Hong Kong so initiated an exodus of foreign capital and assets. This resulted in these economies and their respective stock markets collapsing. For example, approximately 600,000 foreign workers left Thailand, along with billions of dollars.

ATL ACTIVITY 35.2

Investigate one of the following events and determine the significance of capital flight on the economy.

- The collapse of Icelandic banks during the global financial crisis
- Political turmoil in Iran
- The drugs war in Mexico
- The civil war in Syria
- The Russian economic crisis
- The US conflict with Afghanistan
- Venezuela's oil crisis
- Zimbabwe's hyperinflation

Indebtedness

Indebtedness is the condition of owing money. It is the overall debts owed to creditors by all individuals, businesses and the government of the country. All forms of debt, including public sector debt, corporate debt and the debts of private individuals and households, can slow down the economic growth and development of a country. A lower level of indebtedness can be brought about by reduced interest rates (which reduces the amount of debt interest owed by households, firms and the government) or by higher national income. Hence, a lower level of indebtedness tends to be conducive to economic development.

◆ **Indebtedness** is the condition of owing money. It is the overall debts owed to creditors by all individuals, businesses and the government of the country.

Economists believe that if an increase in aggregate demand necessitates public debt, then this is beneficial for the economy (see Chapter 17) as the country will experience economic growth

in the short run and possibly development in the long run. On the contrary, if a country continuously borrows money as loans from other nations and financial institutions, then the amount of debt accumulates in such a way that the country may risk defaulting on the loans owing to the accrued interest on the loans. This means, there is a high probability of not being able to repay the borrowed funds. Many low-income countries facing possible default on their loans are forced into borrowing money from another country or other financial institutions. Hence, the country falls into a spiral of ever-increasing indebtedness, which would clearly limit its ability to allocate funds to investment expenditure or funding projects related to the development of underprivileged people in the economy. Therefore, a mounting burden of debt beyond a certain sustainable level would jeopardize the development of the country.

◆ The **debt to GDP ratio** shows a country's national debt as a percentage of its GDP (national income).

Economists use the **debt to GDP ratio** to measure the degree of indebtedness in a country. The debt to GDP ratio shows the country's national debt expressed as a percentage of its national income (or GDP). A high debt to GDP ratio indicates that the value of the country's output of all goods and services is not enough to pay back debts without securing further debt finance. This is the situation found in many ELDCs (see Table 35.5). By contrast, MEDCs tend to have a low debt to GDP ratio indicating that these countries are financially secure and can afford to pay back their debts. However, high-income nations can also have a high debt to GDP ratio although they are better positioned to be able to afford these debts. In any case, a country with a debt to GDP ratio of more than 100 per cent will tend to find it more challenging to achieve economic development.

Table 35.5 Countries with highest debt to GDP ratio

Country	Debt to GDP ratio (%)
Japan	237.54
Venezuela	214.45
Sudan	177.87
Greece	174.15
Lebanon	157.81
Italy	133.43
Eritrea	127.34
Cape Verde	125.29
Mozambique	124.46
Portugal	119.46

Source: http://worldpopulationreview.com/countries/countries-by-national-debt/

Geography including landlocked countries

The economic growth and development of a country depends largely on its geographical location. Geography can influence the economic development of a country in numerous ways. For example, both theory and empirical evidence suggest that access to good transportation links, especially in terms of sea routes, is a major contributing factor for a country's economic prosperity. From the Middle Ages, cities and nations with ports and open sea routes enjoyed cheaper and faster ways to reach other countries for trade. Cities such as London (with the River Thames), Paris (River Seine), Amsterdam (Amstel River), Shanghai (Huangpu River) and Hong Kong (Victoria Harbour) have thrived from having access to rivers and seas to facilitate world trade. That has enabled these cities, which are not landlocked, to experience faster resource mobilization and higher rates of growth and development.

Figure 35.11 Shanghai is the wealthiest city in China

Another example is agricultural productivity, which can be heavily influenced by geography. Regions with high soil fertility enjoy higher agricultural output. By contrast, many people living on arid (unfertile) land areas struggle beyond subsistence farming for their livelihoods. Factors such as soil erosion, crop parasites, water availability, heat and humidity play havoc in determining agricultural output. Over-reliance on farming discourages communities to adopt urbanization and technological advancement. Some of the hottest and driest countries in the world also find that the climate in their locations can be a hindrance to growth and development, such as Somalia, Sudan, Oman, Iran, Algeria and Libya.

Tropical climates and endemic diseases

Location also dictates the climatic conditions of countries. Countries with tropical weather may have the finest beaches and the best tropical fruits, but they are also susceptible to the deadliest endemic diseases in the world. For example, malaria, which predominantly is a tropical disease spread by mosquitoes, kills approximately 655,000 people each year. Malaria requires a minimum 18°C to effectively transmit the disease. According to UNICEF, malaria is the third biggest killer of children globally, even though the disease is preventable and treatable. Warm weather conditions in the tropics make it ideal for bacterial infections. Other diseases like diphtheria and measles also thrive in tropical climates.

According to the World Health Organization, a child dies of malaria every 2 minutes. Children under the age of 5 are the most vulnerable group, and pregnant women are also at high risk. Sub-Saharan Africa is home to 93 per cent of malaria cases and 94 per cent of malaria deaths in the world. Governments continue to fight against malaria and other preventable diseases, but their efforts often fall short of what is required.

Diseases like malaria reduce labour productivity and change the age structure of the country. Population growth increases as a result to make up for a high death rate. As a result, these countries have fewer resources per capita to sustain their livelihoods. High population growth also tends to lead to a lower quality of life, limited access to education, basic amenities and

Figure 35.12 Malaria is caused by a parasite, spread through the bite of infected female mosquitos

healthcare. High fertility rates force women to look after their families thereby preventing them from joining or remaining in the workforce. Many women from tropical countries engage in unorganized (informal) sectors of industry, effectively contributing nothing to the national income of the country.

Theory of Knowledge

Do economically more developed countries have any moral obligation to help economically less developed countries in their pursuit of growth and development?

PAPER 1 EXAM PRACTICE QUESTION 35.2

Explain **two** economic barriers to growth and development. *[10 marks]*

Political and social barriers (AO2)

In addition to the economic barriers to development discussed above, political and social barriers also exist. A weak institutional framework, legal system, banking system and protection of property rights make it challenging for countries to achieve growth and development. In addition, there are constraints on systems and policies to tackle the problems of gender inequality, poor governance, corruption, unequal political power and status. These are serious political and social barriers to economic growth and economic development.

The IB Economics syllabus states four categories of political and social barriers to economic growth and/or development. These are: a weak institutional framework (legal system, ineffective taxation structures, banking system and property rights), gender inequality, lack of good governance (anti-corruption), and unequal political power and status.

Weak institutional framework

The **institutional framework** refers to the established systems, structures and contexts that shape economic behaviour in a country, such as cultural norms and the country's legal system (for example, security, police enforcement, law and order, protection of property rights and stringent copyright laws). It also involves creating and sustaining an effective financial and banking system that is conducive for economic growth and development.

◆ The **institutional framework** refers to the established systems, structures and contexts that shape economic behaviour in a country, such as cultural norms, the financial framework and the country's legal system.

An ineffective institutional framework, such as a poor law and order system that does not enforce its rules and regulations, deters overseas clients and firms from doing business in the country. Hence, an effective institutional framework is critical to countries achieving economic growth and development through internal and international trade. By contrast, a sound legal framework ensures there is a positive environment that fosters domestic and foreign investment. It also ensures that local authorities can work collaboratively and effectively with international legal bodies like the World Trade Organization (WTO), the International Labour Organization (ILO) and the World Intellectual Property Organization (WIPO) for smoother trade negotiations.

Four aspects of an effective institutional framework are explained below: the legal system, tax structures, banking system and the protection of property rights.

The legal system

The legal framework includes law and order of the country, but also includes corporate law, which facilitates both domestic and international trade. Each year, the World Bank compiles its Ease of

Doing Business rankings. A strong legal system contributes positively to these rankings, which help to guide investment decisions of multinational companies and other investors. Economies are ranked on their ease of doing business, from 1 to 190 (see Table 35.6). A high ranking means the legal framework of the country is more conducive to conducting business, including the establishment of new business start-ups. A range of measures are included, such as the typical length of time it takes to set up a business in the country, regulations regarding construction permits, laws concerning the purchase of commercial property, protection for investors, enforcement of contracts and corporation tax laws.

Table 35.6 Top and bottom 10 countries for Ease of Doing Business, 2019

Country	Rank	Country	Rank
New Zealand	1	Timor-Leste	181
Singapore	2	Chad	182
Hong Kong SAR, China	3	Congo, Dem. Rep.	183
Denmark	4	Central African Republic	184
South Korea	5	South Sudan	185
USA	6	Libya	186
Georgia	7	Yemen	187
United Kingdom	8	Venezuela	188
Norway	9	Eritrea	189
Sweden	10	Somalia	190

Source: Adapted from World Bank Doing Business Report, 2019 (https://www.doingbusiness.org/en/rankings)

Unsurprisingly, the top 10 countries for the Ease of Doing Business are all economically developed economies. An interesting observation is three Asian countries or regions appear in the top 10, along with seven countries from Europe. By contrast, the bottom 10 countries are represented by seven countries in Africa, two from Asia and one from South America.

Taxation structures

There is a clear correlation between the effectiveness of a country's tax structure and its level of economic growth and development. Higher tax rates can help to redistribute income and wealth but can also hinder economic growth if they are too high so as to cause disincentives to work and investment. Governments around the world use different approaches to determine the tax structure best suited to achieve their macroeconomic goals. This will determine the amount of money they have available to spend on improving the welfare of their citizens based on variables such as the demographics of the population, the magnitude of poverty in the country, and inequalities in income and wealth.

The existence of the informal economy is inversely related with the collection of tax revenue for the government. This explains why in ELDCs (low-income countries) only a handful of the population pay income tax. Essentially, corruption and a poor institutional framework encourage tax evasion. For example, less than 10 per cent of citizens in India (the world's second most populous country) pay income tax. Clearly, this is a significant limitation of any country's level of economic development.

The banking system

The banking system provides credit (borrowed funds) for individuals, households and firms to support their consumption and investment expenditure. The provision of credit along with other financial services such as deposits (for savers), insurance and other risk management services are vital to economic growth and development because without borrowing, the economy simply cannot invest in physical, human or natural capital. In many ELDCs, the lack of access to credit for investments in capital, innovations and skills development have impeded their ability to grow and develop. Billions of adults do not have a bank account or access to any form of formal credit (see Table 35.7).

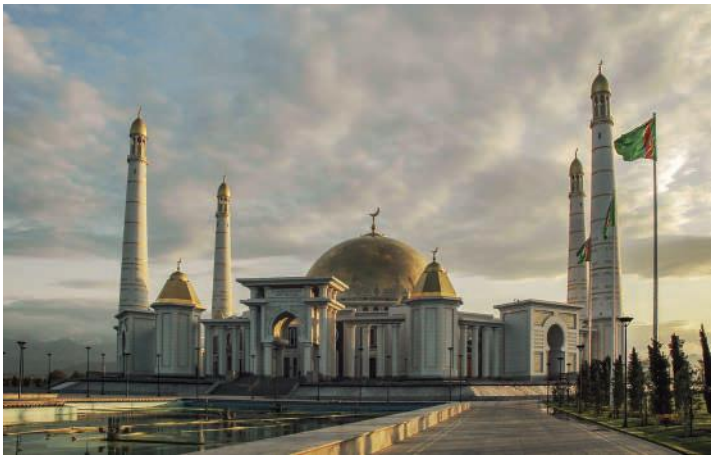


Figure 35.13 Less than 2 per cent of people in Turkmenistan have access to banks

Table 35.7 Countries with the highest and the lowest percentage of the population with a bank account

Bottom 10 countries	% of population with an account	Top 10 countries	% of population with an account
Turkmenistan	1.80	Slovenia	97.20
Central African Republic	3.30	Spain	97.60
Yemen, Rep.	6.50	Estonia	97.70
Niger	6.70	Switzerland	98.00
Guinea	7.00	Belgium	98.10
Burundi	7.10	Germany	98.80
Madagascar	8.60	Australia	98.90
Afghanistan	10.00	United Kingdom	98.90
Iraq	11.00	Canada	99.10
Tajikistan	11.50	Netherlands	99.30
Cameroon	12.20	New Zealand	99.50

Source: Business Insider (https://bit.ly/3bqoGvj)

Many low-income countries use **micro-credit schemes** in an attempt to address the issue of inaccessibility to credit from mainstream banks. Micro-credit refers to loans of exceedingly small amounts to individuals on low incomes in ELDCs for self-employment projects that generate income so they can care for themselves and their families. They help to promote an entrepreneurial culture, especially with women, so help to stimulate economic growth and development. Hence, an efficient and equitable banking system is essential for economic development.

◆ **Micro-credit** refers to loans of exceedingly small amounts to individuals on low incomes for self-employment projects that generate income so they can become self-reliant.

The concept of micro-credit financing was developed by Bangladeshi banker, economist and Nobel Peace Prize winner Muhammad Yunus (2006). The system grants small-sized loans to those (especially women) who cannot ordinarily borrow money from financial institutions. Commercial banks and non-government organizations (NGOs) are the main providers of micro-loans. Most banks will charge relatively high interest rates owing to the high risks of lending money to the poor. NGOs tend to charge lower interest rates, striving instead to stimulate development.

Figure 35.14 Micro-credit can help female entrepreneurs in ELDCs to become financially independent

The protection of property rights (IPRs)

Property rights are the entitlement to both tangible and intangible assets owned by an individual, organization or government, such as land rights. Intellectual property rights (IPRs) are non-tangible assets, such as copyrights, trademarks and patents. These rights cover a wide range of different industries such as cars, consumer durables, land, logos, inventions and commercial ideas.

A monitored and regulated framework for protecting property rights ensures individuals and firms have security over their own property, be they physical and/or intellectual. A well-designed system that enforces the protection of property laws would help firms and business owners feel assured of their investments in a country. By contrast, a loosely articulated and casually implemented system of protecting property rights would discourage investors, especially for international investments and foreign direct investments (FDI). This could lead to unnecessary expenditure in the form of compensation owing to lawsuits and give the country a very poor reputation, further deterring investors from the country.

Property rights offer the owner of the asset the right to:

- exclusive use of the economic good or IPR
- earn income from the asset
- transfer ownership of the asset to others
- enforce property rights, such as claims against others who infringe the property rights.

Common mistake

Beware of claiming that the use of micro-credit will get poor people out of poverty. Unfortunately, borrowing money is not that simple, not even for micro-loans. All lenders use a set of criteria to determine which individuals qualify for micro-credit, including the borrower's credit rating (past record of ability to repay debts), the value of the borrower's assets (to act as security should the borrower default on the loan), and how much debt the borrower can afford. Hence, not everyone will qualify for micro-credit.

◆ **Property rights** are the entitlement to both tangible and intangible assets owned by an individual, organization or government, such as land rights and intellectual property rights.

Figure 35.15 Protection of property rights is important for encouraging corporate investments

One of the biggest challenges in dealing with a lack of property rights, commonly found in ELDCs, is the exploitation of natural resources because the government, firm or individual finds it difficult to monitor or safeguard the asset. In microeconomics, the topic of exploitation of natural resources is covered under the topic of tragedy of the commons (see Chapter 11). Essentially, the lack of property rights and/or enforcement of property rights lead to problems of over-consumption, under-investment, and ultimately the depletion of a common-pool natural resource. Therefore, the lack of property rights in a country is a major political and social barrier to economic growth and development.

Gender inequality

In addition to a weak institutional framework, gender inequality is another political and social barrier to economic growth and development because it limits the quantity and quality of labour resources in the production process. Economists believe that reducing and removing gender inequalities and empowering women are vital steps to a country's development strategy and to reduce poverty in ELDCs.

This global view is supported by the United Nations Development Programme's Sustainable Development Goals (see Chapter 33) that focus on gender equality and the empowerment of all women and girls. The UNDP's Human Development Report, in 2016 (a year after launching the UNDP SDGs) stated: 'every year, 15 million girls under 18 get married – that is one child bride every two seconds. There are 32 countries where women follow different procedures than men to get passports, 18 countries where they have to have their husband's approval before they get a job.' Such inequalities are clearly detrimental to the economic growth and development of society as a whole.

Ignoring gender disparities is detrimental to people's well-being (to girls, women and their children/families). The empowerment of women and greater gender equality helps to reduce social and cultural discrimination against females. In some countries, women are not allowed to work with men, drive or vote, all of which hinders labour productivity and national output. Instead, the empowerment of women and girls can have a huge impact on their self-esteem and

Figure 35.16 Gender inequality and discrimination are harmful for growth and development

productivity. In the long-term, this has positive effects on maternal health and reduces child mortality (both Millennium Sustainable Goals of the United Nations). By contrast, countries that fail to empower women and promote gender equality, such as Iraq, Jordan and Pakistan, face the problems of lower productivity and slower economic growth, and clearly hinder human development as approximately half of the population are denied opportunities to improve their standard of living. Therefore, gender inequality is a hindrance to the economic and human development of all countries, but ELDCs in particular.

The World Development Report from the World Bank recommends four priorities to promote gender equality and to empower women:

1. Improve healthcare and education provision for women. These are two of the most significant investments for fostering human development in ELDCs in order to reduce excess female mortality and to close the education gap between men and women.
2. Improve access to economic opportunities for women, that is, increase employment opportunities for women to participate in the labour force. This also includes improving women's access to credit and other productive resources.
3. Increase women's voice and agency in the household and society. This means improving the ability of females to make effective economic choices and having freedom from the risk of violence.
4. Limit the reoccurrence of gender inequality across generations. For example, China's one-child policy from 1979 to 2016 was used to alleviate social, economic and environmental problems caused by rapid population growth. However, a consequence of this policy has been a major gender imbalance in favour of boys.

WISE ChoICES

Discuss the economic argument that the deliberate **choice** to discriminate against women in the workforce hinders **efficiency** and **economic well-being** for all of society.

Lack of good governance/corruption

A well-structured and well-enforced legal system and good governance are essential to create a framework and an environment conducive to internal and international trade and to prevent or limit corruption in the economy. Corruption is the dishonest or fraudulent conduct by government officials, firms and/or individuals and typically involves bribery and tax evasion. Corruption is deceitful so is detrimental to economic growth and development.

Economic activities and transactions that involve corruption and bribery escape being recorded in the official records of national income. Moreover, corruption causes a misallocation of resources as the result of a lack of good governance. Hence, a high degree of corruption makes it very unattractive for investors, which leads to a wasted opportunity for achieving economic growth and development.

Daniel Kaufmann and Cheryl Gray (1990), from the World Bank and IMF respectively, observed the following characteristics and outcomes of bribery and corruption:

- Bribery is more dominant in ELDCs and emerging (middle-income) countries, although there are disparities between regions of a country.
- Bribery encourages activities undertaken in informal sectors.
- Corruption leads to an increase in transactions costs and raises the level of insecurity in an economy.
- It leads to a decrease in taxes, thereby affecting a government's ability to provide public goods (and merit goods) in the country.
- Corruption tarnishes the legality and dependability of a state or nation.
- Corruption impedes economic growth and development in the long run.

Each year since 1995, the Transparency International publishes its **Corruption Perceptions Index** (CPI), a ranking of the level of public sector corruption in different countries. The CPI is calculated by surveys using 16 different opinion surveys and expert assessments from 12 different institutions (including the World Bank, Economist Intelligence Unit and World Economic Forum). The CPI's definition of corruption is the misuse of public power for private benefit. The CPI ranges from a scale of 0 (highly corrupt) to 100 (very clean).

◆ The **Corruption Perceptions Index** (CPI) is a ranking of the level of public sector corruption in different countries.

Table 35.8 The world's most honest and corrupt countries, 2020

Country	CPI	Country	CPI
New Zealand	87	Libya	18
Denmark	87	North Korea	17
Finland	86	Venezuela	16
Switzerland	85	Equatorial Guinea	16
Singapore	85	Sudan	16
Sweden	85	Afghanistan	16
Norway	84	Yemen	15
Netherlands	82	Syria	13
Luxembourg	80	South Sudan	12
Germany	80	Somalia	9

Source: https://www.transparency.org/cpi2019

A notable observation from the list of countries in Table 35.8 is that the top 10 countries, with the least perceived levels of corruption, are all economically developed (high-income) countries.

Good governance is a prerequisite for sustainable economic growth and development. It enables law and order, minimizes corruption and the misallocation of resources and establishes processes conducive to the overall well-being of individuals and societies (see Table 35.9). The World Bank's Worldwide Governance Indicators (WGI) use the following six measures or dimensions of good governance:

Table 35.9 Top 10 countries based on the Worldwide Governance Indicators

Country	Rank
Finland	1
New Zealand	2
Singapore	3
Sweden	4
Norway	5
Luxembourg	6
Netherlands	7
Switzerland	8
Germany	9
Liechtenstein	10

Source: The World Bank (https://info.worldbank.org/governance/wgi/)

- Voice and accountability
- Political stability and absence of violence
- Government effectiveness
- Regulatory quality
- Rule of law
- Control of corruption

Unequal political power and status

The fourth and final political and social barrier to economic growth and/or development is unequal political power and status. Historically, in literally all societies and civilizations, the rich have been in a position of power, whether it is in the context of economics, society or politics. However, exclusion and intolerance of the majority in a community prevents any economic progress reaching everyone, thereby constraining human development.

The United Nations Development Programme (UNDP) states that intolerance of others in all its forms – legal, economic, social or coercive – is harmful to human development. Unequal political power and status mean that millions of women, indigenous peoples and ethnic minorities continue to be excluded from global progress on human development. The most vulnerable groups of people include refugees and migrants, rural dwellers, people with disabilities and the LGBTI community.

American political philosopher, John Rawls (1921–2002) argued that a moderate inequality in political systems can be tolerable if the welfare of the poorest of the poor increases. He advocated that 'social and economic inequalities are to be arranged so that they are to the advantage of the worst off and should be attached to careers open to all'. Essentially, unequal political power and status thwart economic growth and development.

Figure 35.17 Unequal political power and status for immigrants and displaced people create barriers to development

PAPER 1 EXAM PRACTICE QUESTION 35.3

Explain **two** political and/or social barriers to growth and development. *[10 marks]*

Significance of different barriers to economic growth and/or economic development (AO3)

The economic, political and social barriers to growth and development that have been discussed in this chapter are of varied importance and depend on the context being examined (see case studies on the Democratic Republic of the Congo and Venezuela below). Countries will experience different degrees of these barriers, such as poor infrastructure, human capital inadequacies, over-dependency on the natural resources and agriculture, capital flight, poor governance, indebtedness, and other relevant factors may contribute in different ways as barriers to economic growth and development.

For example, poor infrastructure, a restrictive business environment and declining agricultural productivity all contribute to the lack of growth and development in Madagascar. For Burundi, the barriers are poor economic planning and a high degree of corruption, along with over-dependence on agricultural output. In both cases, the lack of opportunities for growth and development leave people in poverty.

Hence, evaluating the significance of different barriers to growth and development must be done in the context of the country under examination and their specific circumstances, needs and priorities. There are five domestic factors (those within the control of the domestic country) that contribute to the economic development of economically less developed countries:

- education and health
- the use of technology
- access to credit and micro-credit
- the empowerment of women
- income distribution.

These five domestic factors all have positive externalities of production and consumption for ELDCs, thereby improving growth and sustainable development. There are differences in the degree of economic growth and development of ELDCs for the following reasons:

- Resource endowment – Countries have different quantities and qualities of natural resources. For many ELDCs, the main exports consist of agricultural products, while some ELDCs such as Angola have oil and mineral reserves.
- History (colonial or otherwise) – Many ELDCs were former colonies of countries such as Spain and Britain. Colonization often meant wealthier countries extracting marketable resources from ELDCs rather than investing in the countries for economic development.
- Political systems – ELDCs have very different political systems, such as dictatorships found in Algeria, Burundi, Uganda and Venezuela, and democracies in Botswana, Ghana, Zambia and Peru.
- Political stability – This also varies from one ELDCs to another, including the extent to which political turmoil and corruption hinder economic development. Owing to the level of corruption and poor institutional framework typically found in ELDCs, political instability remains a major barrier to growth and development.

- Climate – Countries have varying weather conditions, which can directly impact on national output. Agricultural production is clearly dependent on the climate, but so are other economic activities such as the climate affecting productivity level. Torrential rainstorms and monsoons in South East Asia will hinder production in much the same way as severe hot weather warnings in the Middle East. It is the ELDCs that are most affected by this owing to their lack of diversification in production sectors.
- Population – While ELDCs tend to have large and/or growing populations (such as in Angola, Niger and Uganda), others do not (such as Haiti, Suriname, Syria and Tonga). ELDCs that have very fast-growing populations face greater challenges in achieving growth and development.

CASE STUDY

The Democratic Republic of the Congo (DRC)

The Democratic Republic of the Congo (DRC) has a rich reserve of mineral deposits, including coltan, cobalt, copper, diamonds, gold, tantalum, lithium and tin reserves. This makes the DRC one of the mineral richest countries. However, the DRC is one of the poorest countries in the world with one of the lowest sets of developmental indicators. Years of conflict, war and political upheaval have caused the country extreme poverty and deprived the DRC of achieving economic growth and development.

CASE STUDY

Venezuela

Venezuela is a leading producer of bauxite, coal, gold, iron ore and oil. The country's oil reserves are greater than those of the USA, Canada and Mexico combined. It also has the eighth largest reserve of natural gas, accounting for 2.7 per cent of the global supply. Moreover, Venezuela has the second-largest reserve of gold deposits. However, decades of extravagant government spending made the country solely dependent on its natural resources. Oil comprises 95 per cent of Venezuela's exports and 25 per cent of its gross domestic product (GDP). A drop in the global price of oil has caused the economy to collapse. Poor governance and weak policies have created further barriers to economic development for Venezuela.

Theory of Knowledge

Are the values on which development is based universal or do they depend on culture?

WISE ChoICES

Does the pursuit of greater **economic well-being** threaten **scarcity** in the long run?

PAPER 2 EXAM PRACTICE QUESTION 35.4

Somalia's implementation of fiscal policy has significantly improved, but challenges remain. Tax revenues have increased owing to tariffs. However, recurrent expenditures account for almost all of the government's spending, with capital expenditure accounting for just 3 per cent of total spending in 2016 and 2017. Current spending priorities focus on security and administrative services, which account for almost 90 per cent of total expenditure, thus limiting provisions for other economic and social services.

Figure 35.18 Somalia has the longest coastline on Africa's mainland

Credit to the private sector grew rapidly in 2018, and Somalia's financial sector continued to strengthen with improved regulation of the industry. The current account deficit declined slightly. The medium-term outlook is for a moderate increase in GDP growth. This outlook is subject to significant risks and depends on continued policy reforms. Key risks to medium-term growth prospects include security concerns, weather and climatic shocks, and political uncertainty (instability). Important challenges remain, however, and further delays to improving Somalia's relations with international financial institutions and access to critical external resources could jeopardize the country's economic development for its population of almost 15 million people. Somalia will also need to redirect government spending towards public investment, particularly in order to boost the nation's **human capital**.

a Define the term *human capital* indicated in bold in the text. *[2 marks]*

b Using information from the text/data and your knowledge of economics, examine the extent to which access to credit and appropriate government policies can contribute to economic development in Somalia. *[15 marks]*

Chapter summary

- The poverty cycle (or poverty trap) is a model that shows a combination of factors that perpetuate a vicious cycle of poverty. Intervention, from the government and/or private sector is imperative to break the poverty trap to prevent causing greater hardship for future generations.
- Low-income earners spend most, if not all, of their income on meeting their essential needs, so they have insufficient funds to invest in their future thus are stuck in poverty.
- There are many economic factors that act as a barrier to economic growth and economic development. The barriers specified in the IB Economics syllabus are: rising economic inequality, lack of access to infrastructure and appropriate technology, low levels of human capital (due to the lack of access to healthcare and education), dependence on primary sector production, lack of access to international markets, the informal economy, capital flight, indebtedness, geography (including landlocked countries), and tropical climates and endemic diseases.
- There are also political and social factors that can act as barriers to economic growth and development. The barriers stated in the IB Economics syllabus are: a weak institutional framework (weak legal system, ineffective taxation structures, poor banking system and lack of protection of property rights), gender inequality, lack of good governance/corruption, and unequal political power and status.
- The use of appropriate technology and access to infrastructure are key domestic factors determining the scope of economic development. Poor infrastructure and inappropriate technologies discourage multinational companies from investing in the country, resulting in missed opportunities for attracting foreign direct investment (FDI).
- A lack of access to good healthcare and education create further barriers to economic growth and development. A low level of education and healthcare provision reduces labour productivity and keeps people trapped in poverty.
- The natural and non-renewable resources used in the primary sector will eventually run out, so economies dependent on primary sector production will need to diversify. Primary sector production is also dependent on climate conditions and is susceptible to climate change and global warming.
- The lack of access to international markets for many ELDCs leads to a loss of potential revenue in terms of foreign currencies from exports. Foreign currencies are critical for financing a country's spending on imports on essential capital goods, so are considered as an important prerequisite to growth and development.
- The large informal economy in ELDCs creates some of the greatest challenges for development, such as corruption and the exploitation of child labour and other vulnerable groups in society.
- Capital flight is a large-scale withdrawal of assets and capital from a country due to political and/or economic instability in the country. This can have permanent damaging effects on ELDCs.
- All the economic, political and social factors discussed in this chapter contribute in different ways as barriers to economic growth and development for different countries, that all face different contexts and unique circumstances.

36 Economic growth and/or economic development strategies

Real-world issue

Why is economic development uneven?

Conceptual understandings (WISE ChoICES)

- Perceptions of the meanings of development and **equity** change over time and vary across cultures.
- Governments and other economic agents may intervene in an attempt to promote **economic well-being** and **equity** in societies.
- The pursuit of **sustainability** is subject to various constraints.

SYLLABUS CONTENT

By the end of the chapter, you should be able to understand:

- trade strategies to promote economic growth and/or economic development, namely import substitution, export promotion and economic integration (AO2)
- diversification strategies to promote economic growth and/or economic development (AO2)
- the role of social enterprises as a strategy to promote economic growth and/or economic development (AO2)
- market-based policies and strategies to promote economic growth and/or economic development, namely trade liberalization, privatization and deregulation (AO2)
- interventionist policies and strategies to promote economic growth and/or economic development, that is, redistribution policies including tax policies, transfer payments and minimum wages (AO2)
- the provision of merit goods to promote economic growth and/or economic development, including: education programmes, health programmes, and infrastructure (including energy, transport, telecommunications, clean water and sanitation) (AO2)
- inward foreign direct investment (FDI) as a strategy to promote economic growth and/or economic development, namely import substitution, export promotion and economic integration (AO2)
- foreign aid strategies to promote economic growth and/or economic development, namely humanitarian aid/development aid, debt relief, Official Development Assistance (ODA) and non-governmental organizations (NGOs) (AO2)
- multilateral development assistance strategies to promote economic growth and/or economic development, namely the role of the World Bank and the International Monetary Fund (AO2)
- the role of institutional change as a strategy to promote economic growth and/or economic development, namely improved access to banking (including microfinance and mobile banking), increasing women's empowerment, reducing corruption, property rights and land rights (AO2)
- draw relevant diagrams to show the impact of all the above strategies to achieve economic growth and/or development (using diagrams from other sections in the syllabus) (AO4)
- the strengths and limitations of strategies for promoting economic growth and economic development (AO3)
- the strengths and limitations of government intervention versus market-oriented approaches to achieving economic growth and economic development (AO3)
- the progress towards meeting selected Sustainable Development Goals in the context of two or more countries (AO3).

Strategies to promote economic growth and/or economic development (AO2, AO4)

Governments have adopted a variety of strategies to tackle barriers to economic growth and development, such as different trade strategies, market-based policies, interventionist policies, foreign aid and institutional changes. These strategies are used to achieve a higher standard of living for the people in the country. The following 10 strategies to promote economic growth and development are specified in the IB Economics syllabus:

- trade strategies
- diversification
- social enterprise
- market-based policies
- interventionist policies
- provision of merit goods
- inward foreign direct investment
- foreign aid
- multilateral development assistance
- institutional change.

Note: This final section of the IB Economics syllabus requires SL and HL students to be able to draw appropriate diagrams used in the other sections of the syllabus, such as AD–AS diagrams or diagrams using production possibility curves (PPC).

ATL ACTIVITY 36.1

Conduct individual research on a highly economically developed (high-income) country. Investigate one indicator of economic development where the country chosen is **not** performing well compared to the world average. Investigate the possible reasons why the country is not faring well on that performance indicator.

Trade strategies

Most economists argue there are potential net gains for countries that engage in international trade, enabling them to experience greater levels of economic growth and development (see Chapter 27). For example, export revenue for domestic producers helps to increase aggregate demand and increased employment opportunities for workers. However, international trade can cause problems for many economically least developed countries (ELDCs) and create further barriers to their economic development. Such barriers include:

- **Over-specialization** on a narrow range of products – Over-specialization refers to an individual or country focusing on producing only a small range of products for trade. Thus, over-specialization limits sustainable economic growth and development. Food and agricultural output account for a large proportion of total export earnings for most ELDCs owing to their over-specialization and lack of diversified industries.

- **Price volatility** of primary sector output (see Chapter 35) – Price volatility refers to unstable market prices for primary sector output, such as production in agriculture, fisheries, and mining. A major drawback of ELDCs specializing in the output of primary sector exports is that the prices of these products are highly volatile and often exhibit a downward trend. In many ELDCs, technological progress is relatively slow in primary sector production, thus limiting the ability of ELDCs to compete on an international scale. Supply of primary sector output tends to be inelastic (see Chapter 9), so producers cannot respond quickly to changes in the market. Again, this puts ELDCs in an unfavourable position. Consequences of a long-term decline in the average price of primary sector output include a fall in export earnings for ELDCs, a fall in domestic employment and a deteriorating current account on the balance of payments (see Chapter 32). In addition, ELDCs are more vulnerable to external shocks such as adverse weather conditions and natural disasters, thus making price predictions highly speculative. Such volatility in prices therefore creates uncertainty and obstructs growth and development.
- **Access to markets** – ELDCs lack the ability to access and compete in international markets against the dominance of large multinational companies from high-income countries. For example, France, the UK and USA subsidize their domestic agricultural industries, which means that farmers in ELDCs simply cannot compete in these overseas markets. In addition, ELDCs struggle to access international markets owing to trade protection measures (see Chapter 28) imposed by more economically developed countries (MEDCs) in regional trading blocs. For example, the European Union (EU) imposes a common external tariff on non-member countries including those classified as ELDCs.

In ELDCs, trade strategies enable primary sector output to dominate national income and employment. However, primary sector output tends to receive lower prices than secondary or tertiary sector output, thus worsening the conditions of trade for ELDCs in the long-term.

Figure 36.1 Access to new markets is an important growth and development strategy

Therefore, low-income countries that rely on the export of primary sector output tend to face a deterioration in their international competitiveness over the long-term. This means that the country will have to sell a greater volume of exports in order to finance a given amount of imports. In general, the gains from international trade favour more economically developed countries (MEDCs) owing to their better terms of trade and because of the highly unequal distribution of income in ELDCs. Ultimately, this leads to falling standards of living for citizens of the ELDC in the long run.

Generic trade strategies that countries can use to promote economic growth and/or development are import substitution, export promotion and economic integration.

Import substitution

Import substitution is an inward-looking strategy of economic growth and development that encourages domestic production and the purchase of domestic output through protectionist policies such as tariffs and quotas (see Chapter 28). Import substitution is a development strategy often used by governments to protect infant industries (see Chapter 29) from larger foreign producers until it achieves the necessary economies of scale to become competitive. Supporters of import substitution prefer to use trade protection policies.

◆ **Import substitution** is an inward-looking strategy of economic growth and development that encourages domestic production and the purchase of domestic output through protectionist policies such as tariffs and quotas.

Advantages of import substitution as a trade strategy to achieve growth and development include:

- An increase in domestic employment due to the increase in domestic production.
- Less dependence on imports gives the domestic economy more resilience from external shocks such as a global recession.
- Transportation costs, which affect the final price of goods and services, can be reduced as a result of fewer imports.

Disadvantages of import substitution include:

- The use of trade protection is detrimental to economic efficiency as it distorts the market forces of demand and supply for exports and imports.
- Consumers pay higher prices due to the reduction in imports. They also have less choice, so overall there is a loss of consumer surplus.
- A lack of exposure to international markets may reduce the competitiveness of domestic firms in the long run, thereby reducing the attractiveness of exports.
- Restrictive trade practices are likely to cause other countries to retaliate and impose their own barriers to international trade.
- Overall, there is likely to be a fall in national income as a result of the inefficiencies caused by an import substitution strategy.

Export promotion

Export promotion is an outward-looking trade strategy to achieve economic growth and development that focuses on greater international trade. Countries that adopt an export promotion strategy tend to benefit from greater international specialization and greater choice of goods and services (see Chapter 27). Furthermore, export promotion exposes domestic firms to foreign competition, possibly resulting in greater efficiencies, higher productivity and relatively lower production costs. Supporters of this trade strategy advocate freer and fairer international trade in order to achieve economic growth (see Figure 36.2) and economic development.

◆ **Export promotion** is an outward-looking trade strategy to achieve economic growth and development that focuses on greater international trade.

Advantages of export promotion as a trade strategy to achieve growth and development include:

- Using international specialization and comparative advantages (see Chapter 27) to create an export-led growth strategy to increase the productive capacity of the country. This is illustrated in Figure 36.2 where the export promotion strategy enables the country to benefit

from an increase in aggregate demand (shown by the shift from AD_1 to AD_2) and an increase in potential output (shown by the shift from $LRAS_1$ to $LRAS_2$). Hence, export promotion can increase the level of equilibrium national output (from Y_{f1} to Y_{f2}) without putting pressure on the average price level (P_L) in the long run.

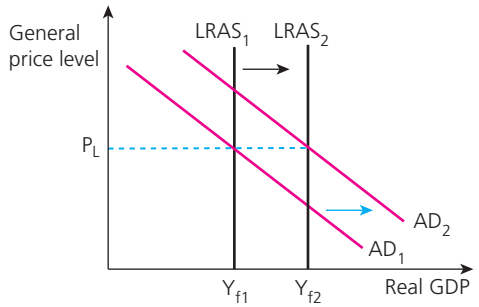


Figure 36.2 Potential gains from an export promotion strategy

- Export promotion enables domestic firms to access larger markets, creating opportunities to achieve greater economies of scale and larger profits.
- Being an outward-looking policy, export promotion brings efficiency such as the transfer of skills and technology, which helps the country to achieve productivity gains.
- Export promotion generates foreign currency earnings in order to pay for the country's import expenditure.

The disadvantages of export promotion include:

- Exporters in ELDCs will usually face stiff competition from other countries, and may struggle to compete against larger and well-established foreign multinational companies.
- Other countries, including high-income nations in North America and Europe, may use various protectionist measures to discourage export promotion from ELDCs. These measures include tariffs and non-tariff barriers (such as quotas, subsidies for their own domestic producers and administrative barriers).
- There is an opportunity cost of government expenditure to support an export promotion trade strategy. All government initiatives and schemes to support and encourage domestic industries will mean that other areas of government spending are sacrificed.

Economic integration

Economic integration is the process of countries becoming more interdependent and economically unified, mainly through the reduction and removal of barriers to trade, that is, through the partial or full abolition of tariff and non-tariff restrictions (see Chapter 30).

◆ **Economic integration** is the process of countries becoming more interdependent and economically unified.

Canadian economist Jacob Viner (1892–1970) coined the term economic integration in 1950, which he used when explaining how trade flows bring countries closer with the intent of greater choice and a reduction in the price of goods and services.

Advantages of economic integration as a trade strategy to achieve growth and development include:

- Greater flows of goods and services between countries, which create economic benefits for all trading partner countries.
- Increased choice of goods and services for consumers, and higher quality products due to international specialization and competition.
- Lowering or removing trade barriers ensures a reduction in prices and a gain in productivity due to a greater degree of competition and efficiency in international markets.
- An increase in investments between trading countries ensures more employment opportunity and sustainable growth opportunities.
- Easier access to international markets allows firms to reach to larger markets, thereby enabling them to take advantage of greater economies of scale and higher levels of profit.
- Improved political relationships between trading countries lead to more stable and harmonious trade negotiations that facilitate global peace and co-operation for the benefit of all members of society.

However economic integration can also have its disadvantages, especially for ELDCs because:

- A country may have to compromise and adjust its internal policies to align and adhere to the rules and policies set by other member states. The loss of autonomy in policymaking and decisions can lead to the domestic government having to adopt policies that are not directly beneficial for the country.
- Similarly, owing to the common rules and regulations of being a member state of a trade agreement or other formal arrangement of economic integration, the country may have to impose common trade restrictions on non-member nations, which can be detrimental as it causes the price of key imports to rise.

Countries can achieve economic integration in different steps or phases. The following categories are explained in Chapter 30, all of which can be adopted to achieve growth and development in the long run:

- Preferential trade agreements (bilateral, regional and multilateral through the World Trade Organization).
- Trading blocs (free trade areas/agreements, customs unions and common markets).
- Monetary union – although this is a very long-term and unrealistic target.

Diversification

Diversification is an economic growth and development strategy that involves countries broadening their supply of goods and services in export markets, rather than relying on a small range of primary sector output. As a strategy to achieve economic growth and development, an ELDC may wish to diversify its production away from the primary goods it has traditionally specialized in producing, opting instead to establish its manufacturing (secondary) industry.

◆ **Diversification** is an economic growth and development strategy that involves countries broadening their supply of goods and services in export markets.

The advantages of using diversification as a strategy to achieve economic growth and development include:

- Diversification helps low-income countries to overcome the problems of over-specialization (which tends to limit economic growth and development for many ELDCs) and creates new employment opportunities in the domestic economy.
- It can help ELDCs to reduce their vulnerability to fluctuating or falling prices in primary sector output.
- During a recession, when the spending power in the economy falls, a diversification strategy can help to create better balance in order to minimize losses during the economic downturn.
- It improves the economy's utilization of unused and scarce resources, which helps the ELDC to achieve better efficiency.
- Diversification can also help ELDCs to reduce their exposure to external supply-side shocks such as extreme weather conditions or natural disasters, which tend to affect primary sector industries the most.
- Diversification may be a sensible strategy for a smaller or land-locked economy so that the country is not overly dependent on a single source income from primary sector output.

Nevertheless, diversification carries potential disadvantages, including:

- The relatively high risks of failure because ELDCs lack the experience and expertise to diversify successfully.
- Diversification carries higher costs due to a broader range of products being produced by domestic firms.

- Diversification requires long-term plans and strategies to succeed, involving training and skill development of the nation's human resources and careful financial planning. This is both time consuming and expensive. The benefits of a diversification strategy, if successful, may not materialize until many years to come.

CASE STUDY

Azerbaijan's dependence on oil

Azerbaijan is a landlocked country located at the crossroads of Eastern Europe and Central Asia. The country completed its post-Soviet transition into a major oil-based economy, with much involvement from the government. The transition to oil production led to remarkable growth, reaching a staggering 26.4 per cent in 2005 before subsiding to 10.8 per cent and 9.3 per cent in 2008 and 2009 respectively following the global financial crisis. However, the real GDP growth rate for 2017 was –0.3 per cent, following the steep fall in world crude oil prices. Azerbaijan's export of oil and gas accounts for approximately 90 per cent of the country's total export earnings. Hence, the lack of diversification in Azerbaijan is a barrier to its development.

Figure 36.3 Baku, Azerbaijan

Social enterprise

A **social enterprise** (or **social business**) is defined as an organization that focuses on meeting specific social objectives, rather than primarily aiming to earn a profit for its owners. Instead, a social enterprise aims to maximize improvements in social and environmental well-being. An example is Oxfam International, a confederation of 20 charitable organizations, with operations in more than 90 countries. The organization strives to end the injustices that cause poverty around the world.

◆ A **social enterprise** (or **social business**) is defined as an organization that focuses on meeting specific social objectives.

Figure 36.4 Oxfam is an example of non-profit social enterprises

Social enterprises often apply strategies used by commercial for-profit organizations, but this is done primarily in pursuit of meeting social objectives such as offering support for the children, the elderly or unprivileged persons, and protection of the local environment (such as protecting local heritage, or developing sustainable tourism to support the local community).

However, not all social enterprises are charities, as social businesses *can* operate as for-profit organizations. Remember that social enterprises are businesses whose primary purpose is to meet social objectives for the common good. This does not mean that such organizations cannot earn profit, although the majority of the profit is reinvested back into the business. This enables them to continue operations in line with their agendas regarding social, environmental and social justice.

Figure 36.5 A social enterprise aims to maximize improvements in social and environmental well-being

PAPER 1 EXAM PRACTICE QUESTION 36.1

Explain two trade strategies used to promote economic growth and/or economic development. *[10 marks]*

Data from the Global Entrepreneurship Monitor (GEM) research project, the world's largest ongoing study of entrepreneurial dynamics, indicate that there is a positive relationship between the level of involvement of social enterprises in ELDCs and the level of economic development achieved during that period. Some of the benefits of using social enterprises in ELDCs are:

- It encourages volunteerism and civic commitment, that is, the use of volunteer labour from MEDCs and local people in community service projects to facilitate growth and development.
- Social enterprises will usually get government support, such as tax incentives and preferential support programmes to encourage widespread participation.
- It can attract more attention from the news media, thereby raising awareness of the support needed to help ELDCs to get out of poverty and to achieve sustainable human development.
- The services provided by social enterprises typically align with the ecological and environmental needs of the local community.

However, there are drawbacks to using social enterprises as a strategy to achieve economic growth and development, especially in ELDCs. These limitations include:

- Social enterprises tend to be small as the volume of their funding is relatively low. Improving social conditions in an ELDC is an extraordinarily complex and challenging task, which requires a long-term strategy and significant funding, which most social enterprises lack.
- As the profit generated from the social businesses often goes back to fund their ongoing expenditures (including salaries and wages for their staff), so this limits the funds that might be available to invest in the country.
- Social enterprises often fall short of expectations as there is a lack of commitment from volunteers (especially if there is no financial compensation for their efforts) and accountability within the organization.
- Given the preferential treatment and tax incentives, social enterprises need to maintain a more thorough accounting and reporting system as their activities are carefully audited and scrutinized to ensure compliance. This bureaucratic process can be time consuming and expensive for social businesses.

Theory of Knowledge

Would economists consider 'paid volunteers' (who work for charitable organizations) an oxymoron?

Theory of Knowledge

How can we know whether it is possible to operate a social enterprise in a profitable and sustainable way?

- Volunteers may not possess the required skills and technical expertise to carry out their duties. Furthermore, as there is usually a high turnover rate of volunteers, a continuous training programme is necessary, which can be expensive for the social enterprise.
- Sceptics argue that social enterprises are not a long-term solution because such organizations may lack the profit incentive to achieve development goals, and are involved only to gain a positive corporate image for their own financial gain.

CASE STUDY

MOOC facilitating improvements in human well-being

The term 'MOOC' was first used in 2008 to describe a new wave of open access, online courses, designed to provide learning to unlimited numbers of people for free. Pioneered by big-name institutions including Stanford University (Stanford Online), Harvard University (edX) and the Open University (FutureLearn), MOOCs usually have no entry requirements and they do not lead to formal qualifications. Yet they have become a huge success, with people around the world taking the opportunity to develop their knowledge without having to attend a university. Other examples of MOOC providers include Khan Academy and LinkedIn Learning.

Dr Sara Calvo and Dr Andrés Morales are experts in sustainability, entrepreneurship and innovation.

Figure 36.6 MOOC is a platform for open access online courses

They are the co-founders of Minca Ventures, a social enterprise that studies social enterprise around the world and offers sustainable education and consultancy services for sustainable businesses. In 2016, with financial backing from the British Council, they teamed up with Middlesex University Business School in London and the Jindal Centre for Social Innovation and Entrepreneurship in India to develop the content of a MOOC course to encourage more people to start up social enterprises.

Market-based policies

Market-based policies are dynamic outward-looking macroeconomic strategies to achieve economic growth and development via free market forces. They focus on liberating industries (from bureaucratic rules and regulations) and improving market incentives in order to increase the level of economic activity. Examples of market-based policies include policies to encourage competition and efficiency, and incentive-related policies to raise productivity and incentives to work. These policies focus on increasing the productive capacity of the economy by improving the supply-side of the economy (see Chapter 26). Such policies also create incentives to invest in the economy. Supporters of market-based policies argue that market forces allocate resources more efficiently than interventionist policies, thereby enhancing economic growth and development.

The IB Economics syllabus specifies three market-based policies to help achieve economic growth and development: trade liberalization, privatization and deregulation.

ATL ACTIVITY 36.2

With a partner, investigate the relationship between trade liberalization, consumer sovereignty and national sovereignty. To what extent are these interdependent?

Trade liberalization

Trade liberalization refers to growth and development policies that encourage free trade, including the free movement of capital flows, by removing barriers to international trade such as tariffs, quotas, subsidies and administrative barriers. The purpose of trade liberalization policies is to encourage more competition and greater efficiency in export and import markets. Improved trade relations and policies are needed to boost the level of international trade and inward foreign direct investment (FDI) in the economy. The International Monetary Fund (IMF) and World Trade Organization (WTO) believe that trade liberalization promotes economic growth, development and poverty reduction.

◆ **Trade liberalization** is the reduction or removal of trade barriers such as tariffs and quotas, thereby increasing competition, productivity and efficiency in international markets.

The advantages of using trade liberalization policies to achieve economic growth and development include:

- Trade liberalization encourages free and fair trade, which ensures a reduction in the costs of conducting international trade.
- A reduction in regulatory costs leads to an overall reduction in prices, which benefits citizens of the country.
- Freer and fairer international trade between different nations increases competition, productivity and efficiency. A key argument for trade liberalization is that market forces improve the global allocation of resources, thereby improving overall economic efficiency.
- It encourages foreign direct investments in the country. This would enable multinational corporations to operate within the ELDC, thus creating employment opportunities.

The disadvantages of using trade liberalization include:

- Intense competition from overseas firms, especially large multinationals, may make some industries no longer profitable for domestic firms to operate. Closure of local businesses will also cause job losses.
- Similarly, products from other countries may be dumped onto ELDCs, which are unable to retaliate. Dumping is the anti-competitive act of selling exports at below cost or deliberately lower than the prices in the local market to gain an unfair advantage (see Chapter 28).
- Freeing up trade and allowing the cheaper flow of goods and services in overseas markets may mean that some producers lack ethical and environmental compliance in order to cut their own costs. Hence, this can have detrimental impacts on quality and safety standards.

CASE STUDY

SEAT

SEAT (Sociedad Española de Automóviles de Turismo) is a Spanish automobile manufacturer with its head office in Martorell, Spain. It was founded in 1950 by a Spanish state-owned industrial company. It quickly became the largest supplier of cars in Spain, but struggled to remain profitable. In 1990, the Spanish government sold SEAT to the German Volkswagen Group in a privatization deal. SEAT became profitable just two years later.

Privatization

Privatization is a form of market-based policy to achieve growth and development by selling or transferring public sector assets to the private sector. The main reasons for privatization are to improve competition, efficiency and productivity in the economy. Privatization can encourage firms to become more efficient in order to survive and be profitable. Private sector firms, driven by financial motives, are argued to be more economically efficient than government officials running bureaucratic public sector enterprises.

◆ **Privatization** is the transfer of the ownership of government assets from the public sector to owners (shareholders) in the private sector. It is a form of supply-side policy that involves the sale or transfer of state-owned assets to the private sector, in order to improve competition, efficiency and productivity.

In ELDCs as a result of inefficiency and widespread corruption, many government-owned enterprises suffer huge accrued losses over many years. A transfer of ownership would not only help these governments offload the burden created by the losses and the ongoing costs of running these organizations but would also provide much-needed finance from privatization proceeds. This could then be used as part of government expenditure in other key sectors such as healthcare and education.

Other key advantages of privatization as a growth and development strategy include:

- Selling off state-owned assets cuts government expenditure on maintaining these assets or businesses, while raising one-off revenues from privatization proceeds for the government.

- The private sector, driven by incentives to work and invest, is more likely to achieve an optimal allocation of the nation's scarce resources.
- Private sector companies foster competition, which can help to reduce prices, improve the quality of the goods produced or the services provided, and increase choice for consumers.
- Privatized companies enjoy an environment free from political interferences and bureaucracy, which in turn would enable companies to operate more smoothly in their pursuit of profit. This helps to create income, wealth and jobs in the economy.

The main disadvantages of privatization as a strategy for achieving growth and development are:

- Privatized companies will typically cut jobs to remove inefficiencies, which can bring in unemployment in a society, at least in the short-term while the privatized firms make the necessary cuts to achieve financial stability.
- The price of the goods and services of a privatized company is more likely to rise as a result of the profit motive of the owners or shareholders. This will be harmful to consumers, especially those in ELDCs.
- Some privatized companies may emerge as a duopoly or monopoly with immense market power, resulting in the concentration of market power in the hands of a few. This again is detrimental to consumers as there is a lack of competition and choice.

Deregulation

Deregulation is the reduction or removal of statutory barriers to entry into certain industries, thus enabling the price mechanism to allocate resources more effectively. The purpose of deregulation is to allow markets to operate more competitively and efficiently. For example, getting rid of excessive rules and directives in a certain industry helps to promote beneficial competition (thereby increasing market supply) and greater productive efficiency. For example, deregulation of financial markets around the world in the 1990s has enabled commercial banks and other financial institutions to have greater flexibility in the types of financial products offered and the fees charged for their services. Having easier and more affordable access to finance is vital for sustainable growth and development.

◆ **Deregulation** is the reduction or removal of government rules and directives in an industry, with the intention of creating more competition and greater efficiency.

Figure 36.7 Deregulation helps to make markets operate more competitively and efficiently

The two main advantages of deregulation as a growth and development strategy are:

- Deregulation benefits trade, industry and the consumers as it limits the inefficiencies created by excessive government control and administrative procedures and processes.
- Similarly, as deregulation enables market forces to allocate resources, it creates a mechanism for quicker decision-making. By contrast, regulation requires bureaucratic processes that are also time consuming to administer.

PAPER 1 EXAM PRACTICE QUESTION 36.2

Explain **two** market-based strategies that can be used to promote economic growth and/or economic development. *[10 marks]*

However, there are disadvantages to the use of deregulation to achieve growth and development. These limitations include:

- Without sufficient regulation of markets and businesses, the economy is more exposed to the uncertainties and fluctuations in the global economy.
- Deregulation can often lead to an increase in deceptive schemes and corruptions. This is a particular concern in ELDCs where laws and regulations are often difficult to uphold. Without an effective legal framework, social and economic welfare can be jeopardized.
- Even with deregulation, the government needs to monitor business activity to ensure state laws are not contravened, such as child labour, monopolization of markets, corporate abuse and exploitation of the environment (such as economic activities that cause pollution to the land, air and water).

Interventionist policies

Interventionist policies are growth and development strategies used by a government to deal with market imperfections in order to improve the productive capacity of the economy. All interventionist policies rely on the country having an effective and functioning institutional framework (see Chapter 35).

Interventionist policies are used to correct market deficiencies, such as providing adequate housing, education and healthcare to ensure a minimum social safety net for all individuals and households. This is highly unlikely to occur in the absence of government intervention. Interventionist policies in the context of growth and development comprise redistribution policies, that is, progressive tax policies, transfer payments and minimum wage policies.

Tax policies

Tax policies are a key redistribution policy, used to redistribute income and wealth to protect and support the poorest members of society. A progressive taxation policy (see Chapter 23) can help a country to reduce income and wealth inequalities, which can ensure economic growth as low-income individuals and households have a higher marginal propensity to consume (see Chapter 25). This means that they are more likely to spend any extra income they receive, thereby increasing the level of aggregate demand in the economy. By contrast, the wealthiest members of society have a relatively low marginal propensity to consume because they have already reached their target wealth so are unlikely to spend a high proportion of any extra income they earn. Redistributing income and wealth in this way means a higher relative disposable income for the poor, thus improving their standard of living.

◆ **Tax policies** are a key redistribution policy, used to redistribute income and wealth to protect and support the poorest members of society.

Taxation policies can also be used to create incentives to work, that is, lower rates of income tax motivate people to find employment and lower rates of corporation tax encourage firms to invest. An effective tax system can therefore help nations to achieve sustainable economic growth and development. However, poor governance and corruption in many ELDCs often prevents governments from using tax revenues appropriately.

Transfer payments

Transfer payments are another means of interventionist policies used to facilitate greater economic growth and development. Transfer payments are financial assistance made to less affluent and less fortunate members of society, for which there is no corresponding change in economic activity (national output). Examples of transfer payments include:

- **State pensions** – This is an income support scheme for people who have reached the official retirement age of the country. State pensions are paid on a monthly basis to support retired people who may not have any other form of income.
- **Unemployment benefits** – These are transfer payments made to unemployed people while they have no other form of income to enable them to meet their basic human needs.
- **Child allowances** (also known as **child benefits**) – These are welfare payments made to parents or guardians of children and teenagers, to supplement their spending on essential items such as food, clothing, schooling, public transport and healthcare check-ups.

Transfer payments are helpful for improving the welfare of the most disadvantaged and marginalized members of the population, providing them with opportunities that would not exist with interventionist policies. However, they are very costly for any government, and the transfer payments are not made in return for any direct economic output or gain. A major limitation of this policy for achieving growth and development is that ELDC governments do not tend to have sufficient funds in their annual budgets to pay for all the necessary transfer payments needed to lift people out of poverty. With a large informal sector in most ELDCs, governments lack the tax revenues required to support vulnerable groups such as the unemployed, the retired and children.

WISE ChoICES

Given that all economies face limited resources, how do governments make the choice between the three types of transfer payments (state pensions, unemployment benefits and child allowances)? How does the concept of **choice** influence this decision?

Figure 36.8 Unemployment benefits are an example of transfer payments

Minimum wage policies

A third interventionist policy to achieve growth and development is the use of minimum wage policies to protect the welfare of the lowest paid workers. As the name suggests, a minimum wage is the lowest amount of money that employers are legally obliged to pay their workers per time period (usually per hour, but this can be per day, per week or per month in some cases). The policy is implemented and enforced by the government to help wage earners to receive a higher amount deemed to provide people with a sufficient living wage. This means that workers can afford basic necessities and have a greater chance of being able to save a proportion of their income. As explained in Chapter 16, a higher amount of savings can ensure investments in physical and human capital, such as educational, vocational training and healthcare provision, all of which are important in helping the country to achieve economic growth and development in the long run.

Figure 36.9 Cuba has one of the lowest minimum wages in the world

The effectiveness of minimum wage policies will depend on several factors, including:

- Compliance of all employers in the country, which would require the government to regulate and monitor the system. Heavy financial penalties may also be required to ensure all businesses conform to the law.
- The amount of pay relative to the cost of living and purchasing power parity. Countries like Australia, Luxembourg, New Zealand and Ireland have the highest level of minimum wages in the world, paying workers more than $12 per hour. Meanwhile, countries like Cuba, Georgia, Bangladesh and Tanzania have minimum wages of less than $0.10 per hour.
- The affordability to businesses as minimum wages will undoubtedly raise the cost of labour for businesses. If this results in unaffordable labour costs, then enforcement of the minimum wage policy will cause some unemployment in the economy.

However, free-market economists argue in favour of labour market reforms to help remove inefficiencies in the labour market. Abolishing minimum wages would be an example, as forcing employers to pay more than the equilibrium can create greater flexibility and productivity, and reduces labour market inefficiencies caused by the minimum wage. Nevertheless, this is a weak argument for those living and working in the poorest ELDCs – imagine the average worker in Cuba, Georgia, Bangladesh and Tanzania who would need to work at least 19 hours a day to earn enough money to match the international poverty line.

Provision of merit goods

Growth and development strategies also include the provision of **merit goods** to benefit all members of society, such as education programmes, healthcare programmes and infrastructure. Merit goods are products that have positive externalities of consumption, so their provision generates social benefits that exceed private benefits (see Chapter 11).

◆ **Merit goods** are products that create positive externalities when they are produced or consumed, that is, MSB > MPB.

Despite their social benefits, merit goods are under-provided and under-consumed in the free market (in the absence of government intervention), yet the consumption of such goods and services would enable positive benefits not only to individuals but also to the community as a whole. For example, without government intervention, the poorest members of society do not have access to education or healthcare services.

Governments can offer merit goods at a subsidized rate or completely free by direct provision to make sure the entire population of the country can consume these socially desirable products irrespective of the differences in the ability to pay for the good or service.

Common mistake

Merit goods can and are offered by both private and public sector organizations. It is incorrect to state that merit goods are provided by the public sector (only). For example, government hospitals and schools coexist with the private hospitals and private schools. Irrespective of whether an individual uses a private or public hospital or school, consumption of these merit good ensures a positive externality to the community as a whole.

Top tip!

The consumption of merit goods ensures significant positive spillover effects to the community, helping the country to achieve economic development in the long run. Make sure you are able to use a market failure diagram to show the impacts of merit goods (see Chapter 11).

Education programmes

Education programmes are a category of provision of merit goods by allocating a proportion of the government's annual budget for education (primary, secondary, tertiary and adult education) and training (such as skills development). This ensures there are spillover benefits to the country because the social benefits accrued from education and training surpass the private benefits in the long run. Educated and well-trained individuals are more likely to succeed in their own life by increasing their earning capacity. Hence, education improves the chances of people attaining a higher standard of living (but of course you know this very well as an IB student!). Education can therefore help individuals escape the poverty trap (see Chapter 35).

A highly educated society is also less prone to criminal activities and is more tolerant towards demographic differences (such as ethnicity and gender). An educated society tends to be more respectful towards public and private property, including intellectual property rights. An investment in higher education leads to improvement in research and development, which could be crucial to achieving an innovative and technologically superior economy. Hence, there are many reasons why spending on education programmes is such an important policy to reduce poverty and achieve economic development in the long run.

However, education is under-provided and under-consumed in ELDCs for several reasons:

- Parents may not be aware of the long-term benefits of education and/or they may not be able to afford the education. Even if education is provided free of charge, there are still costs involved, such as clothing/school uniform, transportation, stationery and other resources necessary for learning.
- In ELDCs, many children are discouraged from going to school so that they can join the workforce, mainly in the agricultural sector.
- Public spending on education can take up a significant amount of the government's budget (see Table 36.1). There is an opportunity cost to this expenditure, especially in military-dominant countries.
- Many ELDCs leave the responsibility of offering education in schools and universities to private sector enterprises. However, this makes education less accessible to most families.

Table 36.1 Education expenditure as proportion of GDP

Rank	Country	Education
1	American Samoa	14.6
2	Cuba	12.8
3	Marshall Islands	12.2
4	Kiribati	12.0
5	Solomon Islands	9.9
6	Botswana	9.6
7	Norway	8
8	Denmark	7.6
7	Palau	7.6
8	Iceland	7.5
9	Belize	7.4
10	Sierra Leone	7.1

Source: https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS

Theory of Knowledge

To what extent do you agree that the IB Diploma Programme is 'short-term pain for long-term gain'? How does time as a concept influence your view of this? How do you optimize the amount of effort or 'pain' to expend in return for proportionate gains?

Figure 36.10 Knowledge can change over time

Health programmes

The provision of merit goods includes investments in health programmes, which generates numerous benefits to the society. A good healthcare system enables the workforce of the country to be healthy and productive. This helps to minimize absenteeism from work, which reduces the number of working days lost due to illness. Hence, good healthcare provision prevents a loss in productivity and potential output.

A good preventive healthcare system, such as universal access to vaccinations, can prevent the spread of contagious diseases and pandemics. By contrast, the outbreak of a major contagious disease could harm economic growth and development. Preventive medicines and vaccinations generate more social benefits than personal benefit. Government spending on healthcare also includes social marketing and promotional campaigns to raise awareness of health-related diseases, such as preventive measures against malaria or HIV/AIDS. Of course, this all comes at a financial and opportunity cost to the government (see Table 36.2). Countries that are able to afford to spend more than 10 per cent of their GDP on healthcare tend to come from Western European and Scandinavian countries, and the USA and Japan.

Table 36.2 Healthcare spending as a percentage of GDP

Country	Health spending (% of GDP)
USA	16.9
Switzerland	12.2
Germany	11.2
France	11.2
Sweden	11.0
Japan	10.9
Canada	10.7
Denmark	10.5
Belgium	10.4
Austria	10.3

Source: **https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS**

Theory of Knowledge

Given the moral arguments for planning and supporting healthcare and education programmes, how do policymakers know how best to distribute public sector funds between these areas of government expenditure?

Infrastructure

Infrastructure refers to the physical structures and facilities required for the efficient running of a country, such as transportation networks, telecommunications networks, social infrastructure and public utilities. Proper infrastructure is also needed to encourage inward foreign direct investment (FDI) to support economic growth and development.

The provision of merit goods also includes government spending on appropriate infrastructure. In the context of development economics, this is particularly relevant in the context of energy, transport, telecommunications, and clean water and sanitation.

Figure 36.11 Investment in infrastructure is vital for growth and development

Energy

Access to energy sources helps a country to progress from an agricultural society to an industrialized one, with its own manufacturing base. Access to energy is a prerequisite to industrial production. A country lacking access to energy sources would have to import these from energy-rich countries, depleting the essential foreign currency reserves. Many countries that lack ecological deposits of crude oil and natural gas have to look beyond these exhaustible energy sources and some are able to invest in environmentally friendly green energies. Green and clean energy sources enable a country to provide affordable and reliable energy in the long run, which facilitates economic sustainability and development.

Transport

Railways, roads, shipping and aviation are common classifications of modes of transportation. Development of infrastructure enables a country to move goods and services quickly for production, consumption and international trade. An effective transportation system includes good use of public transport systems, which would help to promote environmental sustainability, reduce congestion and improve social welfare. The economic costs of congestion extend beyond the private costs for the individual driver (in terms of increased used of fuel, maintenance costs of the vehicle and time lost stuck in traffic) as there are also external and opportunity costs of the journey (such as impact on those using public transport who are delayed, air pollution and lost productivity to the economy).

Railways provide a relatively cheap and reliable mode of transport for passengers and freight. In particular, heavy industries depend on the railways for the transportation of raw materials and semi-finished materials and components. Manufacturers may also rely on railways and ports to transport their finished goods to wholesalers or retailers. Highly valuable products (such as diamonds and jewellery) or perishable products being exported overseas may require air freight, so investments in airports are also important for international trade. However, the investment in transportation infrastructure requires significant funding from the government and transportation companies. This remains a huge challenge for many ELDCs owing to their relatively low national income.

Telecommunications

A good telecommunications system is a positive agent of economic growth and development. Investment in telecommunication networks, such as a stable broadband and wireless internet infrastructure, helps to foster effective communications, facilitates education (online learning),

Figure 36.12 Traffic jams are a common occurrence in Bangkok, Thailand

strengthens healthcare provision and enhances the banking system (see Chapter 35) through mobile banking. A country with good telecommunications infrastructure is more likely to attract foreign direct investment, which would create more jobs and increase per capita income.

In 2015, Tom Goodwin, an executive at the French media group Havas, stated the following:

- Airbnb – the world's largest accommodation provider owns no real estate.
- Alibaba – the world's most valuable retailer has no inventory.
- Facebook – the world's most popular media network creates no content.
- Uber – the world's largest taxi company owns no vehicles.

What these market-leading companies all have in common is the reliance on internet telecommunications technologies, enabling them to act as connectors or intermediaries between consumers and producers, through the use of their online and mobile apps (applications). However, investments in telecommunications infrastructure remains a huge barrier to growth and development in ELDCs owing to the sheer financial costs of such investments.

Clean water and sanitation

A final aspect of infrastructure needed for sustained growth and development is clean water and sanitation. This is part of the UNDP's Sustainable Development Goals, to ensure availability and sustainable management of sanitation and water for all by 2030 (Goal 6). Part of this goal is to ensure that everyone has a safe, working toilet. According to the Norwegian Refugee Council (www.nrc.no), around 30 per cent of the world's people do not have access to toilets. This lack of access deprives people of good health, safety and dignity. To put this into perspective, it was reported back in 2013 that there are more people on the planet with a mobile phone than there are people with flushing toilets. The latest findings from the United Nations (2018) show that this is still the case.

Clean drinking water and sanitation is a basic need of any society. Access to clean drinking water and sanitation helps communities to live a healthy life. Contaminated drinking water and a lack of sanitation often leads to the spread of infectious diseases that harm entire communities in terms of lost productivity, lower national output, increased government spending to support the sick and the loss of lives. According to data from the UN, bad sanitation facilities kill around 280,000 people each year (or 32 people every hour). For displaced people (those forced to flee their community or country), having access to clean water and appropriate sanitation facilities are among the most urgent of all human needs.

Figure 36.13 Clean drinking water helps communities to enjoy healthier lives

WISE ChoICES

Government **intervention** and market-based approaches are essential to achieve **economic well-being**. Policymakers face a strategic **choice** between these two options. How can governments determine which approach is more appropriate for their country?

PAPER 1 EXAM PRACTICE QUESTION 36.3

Explain how the provision of merit goods improves economic growth and/or economic development. *[10 marks]*

Inward foreign direct investment (FDI)

Foreign direct investment (FDI) refers to the long-term capital expenditure of multinational companies in overseas markets, such as Apple's manufacturing plants in China. Inward FDI means the country in question attracts FDI funds from foreign multinationals – for example, China attracts inward FDI from Apple, a US company. By contrast, outward FDI occurs when the country in question uses FDI in a foreign country, such as Japan's outward FDI in Europe and the USA as Toyota, Nissan and Honda all have production plants in the European Union and North America.

◆ **Foreign direct investment** (FDI) is spending by multinational corporations (MNCs) abroad.

Foreign direct investment is the money devoted by multinational companies (MNCs) to business operations abroad, such as building production facilities in overseas markets or acquiring (taking over) a foreign company operating abroad. Hence, FDI expenditure can help to improve economic growth and development in ELDCs. Globalization and the promotion of freer international trade from the World Trade Organization (WTO) have encouraged a significant increase in foreign direct investment in ELDCs. For example, China's investment projects in Latin America and sub-Saharan Africa have helped to create jobs and boost economic growth in these regions.

More economically developed countries (MEDCs) account for the vast majority of the world's FDI. This is because the largest multinational corporations are from MEDCs. MNCs have production plants and/or service operations in overseas countries. Many MNCs have grown by using joint ventures and strategic alliances with firms in ELDCs. This has resulted in a large increase in inward FDI in ELDCs.

FDI expenditure is highly dependent on changes in the state of the world economy. During economic booms, FDI tends to rise significantly as firms believe that they can increase their profits by expanding overseas. There are a number of reasons why multinational companies are eager to invest in this form of FDI, although this will have varying impacts on the growth and development in ELDCs:

- Cheaper production costs – Many MNCs have operations in low- and middle-income countries in order to exploit lower costs of production. For example, it is far cheaper for Nike to hire workers in Indonesia, China and Vietnam than in the USA.
- Economies of scale – Similarly, by operating on a larger scale owing to a larger customer base, MNCs are able to enjoy greater economies of scale, thereby reducing their unit costs of production. By operating in overseas markets, MNCs are also able to spread their risks (risk-bearing economies of scale).
- Access to natural resources – MNCs are keen to expand into ELDCs that are well-endowed in natural resources that can be exploited for production.
- Increased sales revenue – Access to fast-growing economies with large populations such as Indonesia, Nigeria, Bangladesh, Vietnam and Ethiopia present enormous growth opportunities for MNCs. Hence, MNCs are keen to establish operations in these overseas markets.
- Avoid trade barriers – By locating within foreign countries, MNCs are able to avoid trade barriers such as tariffs, quotas and administrative obstacles. For example, German carmaker Volkswagen has production plants in India, Indonesia and Nigeria.
- Logistical reasons – MNCs locate overseas to reduce delivery times and costs to customers in ELDCs and emerging markets. For example, US power tools maker Black & Decker has production units in China, so their products sold in Asia do not have to be exported from the USA.

- Financial incentives – In order to attract inward FDI, the governments of ELDCs often offer MNCs incentives to locate in their country, such as tax rebates, grants, subsidies and cheaper rents.

Inward FDI is generally beneficial to all countries, but particularly to ELDCs and developing countries for the following reasons:

- Inward FDI is a major source of national income for ELDCs. The funds from FDI are far more stable than financial aid, and with long-term benefits for ELDCs. In the long run, inward FDI helps to shift the ELDC's long run aggregate supply and aggregate demand curves to the right, thereby increasing national income without putting undue pressure on the general price level.
- Higher national income, resulting from inward FDI, can help ELDCs to close their savings gap. As the level of savings in the ELDC increases, more funds become available for investment expenditure in the economy, with long-term benefits to the country.
- The additional profits generated from the investments of MNCs in an ELDC contribute to the country's tax revenues.
- Inward FDI allows the transfer of knowledge, technology and more efficient work practices (collectively known as *skills transfer*) from MEDCs to ELDCs.
- Inward FDI and direct rivalry from foreign MNCs can force domestic producers in the host country to become more efficient and competitive. The presence of MNCs in ELDCs also provides domestic households and firms with a wider range of choice, and at more competitive prices.
- Inward FDI by foreign MNCs provides employment opportunities in the ELDC or host country. This helps to increase consumption expenditure of domestic individuals and households, thereby improving their general standard of living.
- The (potential) presence of MNCs often encourages governments in ELDCs to invest in infrastructure; large MNCs may also help to provide some of the funding for this. Such investments help to benefit the country as a whole.

However, inward FDI in ELDCs also has its disadvantages, such as:

- Although inward FDI helps to create employment opportunities in ELDCs, MNCs often bring in their own management teams while locally hired employees are low-skilled workers on minimal pay. This reduces the benefits of skills transfer to ELDCs.
- The lenient regulatory framework in ELDCs means MNCs often exploit their position by:
  - disregarding issues of health and safety at work, such as imposing long working hours on employees who have to tolerate poor working conditions
  - ignoring the external costs of their activities in ELDCs, such as pollution of the natural environment (land, air and sea) and depletion of non-renewable resources.
- Inward FDI from large MNCs makes domestic rivals less competitive due to the advantages enjoyed by the MNCs, including economies of scale, technological know-how and tax rebates from the ELDC government. In extreme cases, this can lead to the collapse of domestic firms.
- In most cases, the profits generated by MNCs are repatriated to their home country rather than reinvested to further improve the facilities and infrastructure in ELDCs. Thus, critics argue that MNCs exploit and profit from ELDCs without really giving much back.

Top tip!

Economists assume that economic agents (households and firms) act rationally, that is, in their own best interest. This suggests that MNCs spend money on FDI in order to increase their own revenues and profits, rather than to promote economic development in ELDCs. After all, companies operate to satisfy the financial interests of their shareholders (owners) rather than to create social welfare.

- Some MNCs are criticized for having too much power when negotiating tax breaks and other financial incentives with ELDC governments. Many ELDCs have limited bargaining power as large MNCs have the option to redirect their FDI to neighbouring ELDCs that are willing to offer better terms and conditions.
- Critics also argue that the growing presence and power of large MNCs creates a loss in the cultural, political and economic identity of ELDCs in favour of economically developed nations, like the consumption of fast foods, soft drinks and expensive branded goods.

PAPER 2 EXAM PRACTICE QUESTION 36.4

Prior to the political crisis, Egypt was an attractive Middle Eastern market for **inward foreign direct investment** (FDI), receiving abundant liquidity from the Gulf States. FDI decreased sharply as a result of the global economic crisis and following the socio-political revolution of 2011 grew at a steady rate until 2017 and 2018 (see Table 36.3). Egypt remains the largest FDI recipient in Africa, with a stock of $116.4 billion (46.6 per cent of GDP). The UK is by far the largest investor in Egypt. FDI is concentrated in the oil and gas industry (around two-thirds of total investments, especially after the discoveries of offshore gas reserves in the country's western desert and in the Zohr offshore field, believed to be the largest-ever gas discovery in Egypt and the Mediterranean). This is followed by FDI in the construction, manufacturing, real estate and financial services sectors.

Table 36.3 Key economic indicators for Egypt

Key economic indicator	2018	2017	2016
Inward FDI ($ million)	8,107	7,409	6,798
GDP per capita ($)	3,525	2,441	2,549
Inflation rate (%)	13.81	29.50	20.86

a Define the term *inward foreign direct investment* (FDI). *[2 marks]*

b Using a poverty cycle diagram, explain how the fall in inward foreign direct investment (FDI) in Egypt between 2016 and 2018 might lead to a deterioration in economic development. *[4 marks]*

Source: Adapted from **https://santandertrade.com/en/portal/establish-overseas/egypt/foreign-investment**

Common mistake

It is a common mistake to think that MNCs exploit workers by paying them low wages. While wages paid to workers in ELDCs are comparatively lower than in the MNC's home country, they are generally higher than wages paid by local firms in the host country. It would be unreasonable to expect MNCs to pay workers in Suriname, Pakistan or Nicaragua the same wages as workers in Germany, Norway, the UK and the USA owing to the different costs of living in these countries.

Foreign aid

Foreign aid refers to assistance in the form of finance, goods and/or services granted to ELDCs for the purpose of economic development. Foreign aid is concessional and non-commercial, that is, it is a gift from the donor country rather than a loan for financial gain. Such assistance helps to make up for the shortcomings of the free market that fails to support ELDCs during times of economic need, such as emergency humanitarian aid following a major natural disaster in the country.

◆ **Foreign aid** is financial assistance from donor countries to ELDCs for the purpose of economic development. Aid is given on concessional and non-commercial terms, rather than for financial gain.

Foreign aid to ELDCs can be granted by governments of donor countries or by non-governmental organizations (NGOs), such as Oxfam and the Bill & Melinda Gates Foundation. Donations by private individuals are usually made through a charity or NGO. NGOs operate independently of any form of government, often pursuing their own specific aims to improve social well-being. The motives for giving foreign aid include:

- Humanitarian motives – Foreign aid is provided for humanitarian reasons, such as emergency relief to help those most vulnerable after a major natural disaster or civil war. It is also part of the UN's Sustainable Development Goals (SDGs) such as eradicating famine and improving maternal health for the benefit of all humanity.
- Economic motives – Donors give foreign aid to ELDCs because it may be in their financial interest to do so as this builds better economic ties with the recipient countries. For example, tied aid provides economic benefits to the donor country as the foreign aid must be spent on goods and/or services from the donor country.

Figure 36.14 Humanitarian aid is an example of foreign aid for ELDCs

- Political motives – Some European countries such as the UK and France provide foreign aid to their former colonies. Historically, the USA has provided foreign aid to support capitalism and free market practices in ELDCs. Japan has given foreign aid to Nicaragua to influence its vote on the banning of whaling.

The generic advantages of foreign aid include:

- Without foreign aid, many of the world's poorest countries would struggle to ever get out of the poverty cycle. Thus, there are humanitarian, economic and political benefits to providing foreign aid.
- Foreign aid can be used to increase production and productivity in ELDCs, thus helping them to achieve economic growth.
- Foreign aid can help to reduce or eradicate extreme poverty (the first of the UN's Sustainable Development Goals).
- As a form of injection into the circular flow of income, foreign aid can help to reduce inequalities and unemployment in the ELDC.

However, there are limitations to relying on foreign aid. These include:

- Foreign aid can create economic dependence on donors from MEDCs. This does not help ELDCs to be financially independent or to develop in the long run.
- Corrupt governments in ELDCs often misuse foreign aid rather than passing on the funds for local projects that would benefit local communities and industries.
- Foreign aid in the form of concessionary loans incurs interest repayments, so this adds to the financial burdens of many ELDCs. This is exacerbated if ELDCs experience a fall in their exchange rate, because debts have to be repaid in foreign currencies.
- ELDCs tend to prefer more favourable terms of international trade to help their economic development rather than outright foreign aid. For example, food aid and tied aid do little for the long-term prosperity of an ELDC.
- Most economists argue that economic development should be based on promoting free international trade, such as increasing the exports of ELDCs, instead of relying on foreign aid. This is because foreign aid does not necessarily lead to growth or development, whereas trade is more likely to do so.
- More often than not, foreign aid is insufficient to really help ELDCs develop their economies. The aftershock of the global financial crisis of 2008 and the coronavirus pandemic also led to a decline in donations from MEDCs (as a percentage of their GDP).
- The type of foreign aid given is not always appropriate. For example, aid might promote capital-intensive projects rather than labour-intensive output that is better suited for ELDCs. Foreign aid might also involve political interferences rather than direct economic benefits.

The IB Economics syllabus specifies four categories of foreign aid: humanitarian aid/development aid, debt relief, Official Development Assistance (ODA), and support from non-governmental organizations (NGOs).

Humanitarian aid/development aid

Humanitarian aid (or **development aid**) is foreign aid used to help recipient countries to achieve their economic development objectives. These aims include eradicating extreme poverty, improving education, reducing child mortality and improving the overall standards of living in the country. Humanitarian aid refers to altruistic aid, typically given to save lives and maintain human dignity in response to violence, natural disasters and national emergencies. It consists of food aid, medical aid and emergency relief aid (such as aid for reconstruction work following a natural disaster).

◆ **Humanitarian aid** (or **development aid**) is foreign aid used to help recipient countries to achieve their economic development objectives and to improve their overall standards of living.

Development aid can be provided by individual countries, by NGOs, or by multilateral organizations such as the World Bank and Save the Children. Humanitarian aid includes the following:

- **Grants** – These are non-repayable financial assistance provided by governments, that is, they do not have to be repaid. Most grants are given to fund specific projects.
- **Concessional long-term loans** – Also known as **soft loans**, these are loans with highly favourable terms and conditions such as low interest rates and long repayment periods.
- **Project aid** – As the name suggests, this is foreign aid for specific developmental projects, such as financial support for schools (education) and hospitals and sanitation (healthcare).
- **Programme aid** – This is financial aid given to support a specific industry, such as funding of the financial sector or the education sector to get more primary-age children into schools.
- **Conditional aid** – This is foreign aid in the form of capital, grants or loans on the condition that the recipient nation uses this only for policy reforms and structural adjustments in the ELDC.
- **Tied aid** – The recipient country needs to spend the foreign aid on buying products from the donor country.

Top tip!

Tied aid is often criticized as the ELDC needs to buy goods and services from the donor country even if this might not be appropriate or efficient. It is often perceived as an indirect subsidy for suppliers in the donor country. Hence, tied aid can harm the competitiveness of firms from the ELDC. Nevertheless, a valid reason for tied aid is that it ensures that foreign aid is used for appropriate purposes only, while also benefiting the donor country.

Debt relief

Debt relief is a form of foreign aid involving partial or total remission of foreign debts, especially those owed by low-income countries and ELDCs to external creditors from MEDCs. Hence, debt relief is sometimes referred to as **debt forgiveness**.

◆ **Debt relief** (or **debt forgiveness**) is the partial or total remission of foreign debts, especially those owed by low-income countries and ELDCs.

Figure 36.15 Debts can be a major burden on growth and development

Governments can incur two types of debt. **Internal debt** is the money owed by a country to domestic lenders, such as private commercial banks, because the government has a budget deficit (that is, government expenditure exceeds government revenues). By contrast, **external debt** (or **foreign debt**) is money owed to foreign creditors (lenders), such as foreign commercial banks, foreign governments and international financial institutions such as the World Bank and the International Monetary Fund (IMF).

Foreign debts are often incurred by countries with weak institutional frameworks (including the domestic monetary system) and poor infrastructure. Hence, they tend to borrow from foreign creditors to develop their economies. Foreign debt is a potentially serious matter for many ELDCs because the burden of interest payments and strict loan terms have often resulted in huge opportunity costs, such as cutting expenditure on education and healthcare in order to repay these loans.

It is not uncommon for ELDCs to be heavily indebted because relatively high interest rates are charged by creditors and/or domestic inflation reduces the value of their currency (foreign debts have to be repaid in foreign currencies). This makes their debt financing become increasingly unsustainable. The IMF and World Bank have identified a group of 39 nations classified as **heavily indebted poor countries** (HIPCs) with high levels of poverty and unaffordable debts, which makes them eligible for special assistance from these international organizations.

◆ A **heavily indebted poor country** (HIPC) is a low-income nation with a huge outstanding debt, making it eligible for special financial assistance from the IMF and the World Bank.

The problems of foreign debt can be traced back to the global oil crisis of 1973 when oil prices increased by 500 per cent due to the Arab–Israeli War (or the Yom Kippur War). This created a huge surplus of profit for exporters of oil, deposited in banks, with the money being lent to ELDCs. The subsequent oil crisis of 1979 caused a worldwide recession, with ELDCs unable to export enough of their commodities to pay off their debts. As a result, many ELDCs defaulted on their loans. Foreign debt can cause both economic and social instability. For example, the global financial crisis caused huge economic problems even for MEDCs, such as unemployment reaching 28 per cent in both Spain and Greece. Ultimately, high levels of foreign debts can make ELDCs even poorer.

Debt relief can give ELDCs and HIPCs a lifeline, granting these countries with more resources to spend on human capital or improving infrastructure in the country. However, there are two main economic limitations of debt relief:

- Corruption in the country and misuse of funds often limit the value and effectiveness of debt relief programmes. In too many cases, ELDCs incur huge debts because concessionary loans and financial aid are misused by corrupt leaders, for example, to finance military spending.
- Recipient countries may assume that similar relief will be offered in the future, encouraging them to be less financially prudent and going further into debt.

Some countries have become so heavily indebted that they have had to reschedule their debt repayments to banks and other lenders. **Debt rescheduling** means renegotiating the length of time to repay existing loans, often leading to further borrowing and escalating debts. HIPCs are also highly vulnerable to external shocks that are beyond their control, such as an oil crisis (which causes inflation). These supply-side shocks add to the soaring debts of HIPCs. This results in debts incurred by some HIPCs exceeding the government revenue from taxpayers, thus causing a **debt trap**, that is, they are unable to ever repay their mounting debts. This reduces their credit rating (financial status), thus making future lending and foreign direct investment in these countries less attractive.

The mounting debt burden faced by HIPCs has led to international pressures for creditors to cancel the debts (known as debt forgiveness) in an attempt to restart or improve the economic development of HIPCs. However, this can cause some HIPCs to become complacent (or reckless) as they are protected from their irresponsible behaviour, so may continue to be irresponsible in the future.

Common mistake

Not all debt is detrimental to the well-being of ELDCs. This depends on the level of affordability of the debt. Borrowing money to fund structural changes in the pursuit of economic development can be beneficial to ELDCs – it is when the foreign debt is unaffordable and poorly managed that problems arise.

◆ **Debt rescheduling** means renegotiating the length of time to repay the existing loans of highly indebted nations.

ATL ACTIVITY 36.3

As a student of the IB and economics, do you support debt relief for heavily indebted poor countries? Which traits of the IB Learner Profile would influence your answer the most?

Official Development Assistance (ODA)

A third category of foreign aid is **Official Development Assistance** (ODA), which is foreign financial assistance from donor governments, rather than from NGOs or not-for-profit organizations, specifically for development purposes. Foreign aid through ODA can be provided bilaterally (from donor government to recipient government) or channelled through a multilateral development agency (such as the United Nations or the World Bank). ODA includes grant and concessional loans offered at preferential interest rates.

◆ **Official Development Assistance** (ODA) is foreign aid from donor governments, rather than from NGOs or non-profit organizations, for development purposes.

The United Nations has advocated that economically developed countries should spend 0.7 per cent of their gross domestic product (GDP) on ODA to help countries eradicate poverty and become self-sufficient. Thirty member countries of the Development Assistance Committee (DAC) of the Organization for Economic Co-operation and Development (OECD) are the key contributors to this programme. The DAC members have used ODA to support ELDCs since 1969. Table 36.4 indicates the top 15 contributing countries to ODA, in absolute terms and as a percentage of their GDP. Notice that only seven countries have been able to commit to the UN's recommendation of contributing 0.7 per cent of their GDP to ODA.

Table 36.4 Top contributing countries to ODA

Country	ODA ($bn)
USA	31.08
United Kingdom	18.70
Germany	17.78
Japan	9.32
France	9.32
Sweden	7.09
Netherlands	5.81
UAE	4.39
Canada	4.29
Norway	4.28

Country	ODA (% of GDP)
Sweden	1.40
UAE	1.09
Norway	1.05
Luxembourg	0.93
Denmark	0.85
Netherlands	0.76
United Kingdom	0.71
Finland	0.56
Turkey	0.54
Switzerland	0.52

Source: **http://www.oecd.org/dac/stats/ODA-2015-detailed-summary.pdf**

Non-governmental organizations (NGOs)

Non-government organizations (NGOs) are an important agent in the provision of foreign aid. NGOs have provided more foreign aid than many OECD donor countries, raising tens of billions of dollars each year. UNDP data show there are eight dominant NGOs contributing to foreign aid: World Vision International, Oxfam International, Save the Children International, Plan International, Médecins Sans Frontières, CARE International, CARITAS International and ActionAid International. Many international and local NGOs are active in ELDCs and HIPCs, carrying out major humanitarian work from the foreign aid raised in donor countries.

International, national and the local NGOs help to increase the capacity to access funds offered by governments, local communities and other international organizations. Many of the larger international NGOs have a larger reach and influence with governments around the world, with a trained and specialized workforce. They are often supported or endorsed by politicians, celebrities and academics, which enables these NGOs to have greater influence and access to foreign aid.

Some NGOs, such as Oxfam and World Vision, are directly involved in planning, implementing and managing developmental programmes and the humanitarian needs in ELDCs and HIPCs. Some of these NGOs have also undertaken their own research to collect data for assessing and analysing the underlying problems faced in low-income countries, in order to lobby governments for a better future and to create greater opportunities for growth and development.

Top tip!

'Give a man a fish, and you feed him for a day; show him how to catch fish, and you feed him for a lifetime' – Anne Isabella Thackeray Ritchie (1837–1919). This is a useful quote to remember when discussing the 'Aid versus Trade' debate in economics regarding growth and development.

Top tip!

When evaluating foreign aid as a strategy for economic growth and development, consider the following questions as a framework:

- What are the types of foreign aid being given?
- What are the underlying motives behind giving foreign aid?
- Is the foreign aid being given for the right (economic) reasons?

Multilateral development assistance

Multilateral development assistance (MDA) is another strategy used by governments to achieve economic growth and development. MDA occurs when a donor country allocates financial aid to an intermediary organization (such as the World Bank, the IMF or the United Nations), which then distributes these donations to recipient countries. These multilateral institutions are international organizations made up of member governments around the world, who pool their resources together thereby enabling large-scale development programmes to be funded.

◆ **Multilateral development assistance** is financial support delivered through international institutions such as the World Bank and the International Monetary Fund (IMF).

Multilateral development assistance is generally seen as a less political form of foreign aid than bilateral aid and tied aid as MDA encourages international co-operation rather than focusing on the financial interests of donor countries. Multilateral development assistance often takes the form of non-concessionary loans, that is, lending that incurs interest and repayment periods determined by market forces. However, this differs from commercial bank lending as the loans are specifically for development purposes.

The World Bank

The **World Bank** is the international organization that lends money to ELDCs for economic development projects and programmes for structural change. The majority of MDA finance is for physical capital projects, such as irrigation systems, road networks, schools, hospitals and transportation systems. The World Bank was set up in 1944 (at the same time as the IMF) to provide foreign development assistance (concessionary and non-concessionary lending) to low- and middle-income countries to reduce poverty and improve standards of living. The World Bank is made up of two international financial institutions:

- The **International Bank for Reconstruction and Development** (IBRD) – The IBRD provides loans to middle-income countries (not ELDCs), so the funds are not technically classified as foreign aid but development assistance. Most of the lending from the World Bank is made by the IBRD. In 2018, the largest recipients of World Bank loans were India ($859 million) and China ($370 million). China remains the largest borrower of IBRD, having borrowed more than $60 billion since its membership in 1980.
- The **International Development Association** (IDA) – The IDA is the part of the World Bank that supports the world's poorest countries. Member countries of the IDA offer concessionary, interest-free loans of up to 30 years to low-income countries in order to reduce poverty. In 2017, the IDA committed to $19.5 billion in MDA finance for ELDCs.

The World Bank has been criticized for offering most of its lending to middle-income countries rather than to ELDCs that face conditions of extreme poverty. For example, Egypt, Indonesia and Iraq borrowed a total of $14.57 billion from the World Bank between 2016 and 2018, which is about 2.6 times larger than the GDP of Togo and almost 4 times higher than the GDP of Sierra-Leone (both HIPCs). Furthermore, critics point out that the requirements for eligibility of funding from the World Bank disqualifies many poor and heavily indebted poor countries.

Figure 36.16 The World Bank Headquarters in Washington DC

The International Monetary Fund (IMF)

The **International Monetary Fund** (IMF) is an international multilateral financial institution set up in 1944 by 29 member countries. Today, there are 188 members of the IMF. Its goal is to oversee the international financial system and to promote global monetary co-operation. Thus, the IMF helps to facilitate sustainable economic growth and development. This includes providing multilateral development assistance and making short-term non-concessional loans to countries that experience difficulties making their international payments.

Traditionally, most of the lending from the IMF went to ELDCs. However, the global financial crisis of 2008 saw a significant amount of lending to MEDCs with their huge balance of payments problems, such as Portugal, Italy, Ireland, Greece and Spain. However, the governance of the IMF is often criticized because the wealthier member states have a greater share of the voting rights. Thus, critics argue that the welfare of ELDCs and HIPCs and their development priorities are often overlooked.

Common mistake

Many students tend to think of the World Bank as a single institution; it is not, as the World Bank comprises the International Bank for Reconstruction and Development (IBRD) and the International Development Association (IDA), with different foci.

Institutional change

The final strategy mentioned in the IB Economics syllabus for achieving growth and development is institutional change. This covers the following five aspects of institutional change: improved access to banking (including microfinance and mobile banking), increasing women's empowerment, reducing corruption, extending and protecting property rights, and uploading land rights.

Improved access to banking, including microfinance and mobile banking

A key institutional change needed to facilitate growth and development in ELDCs is improved access to the banking system, including access to finance. Improved access to banking will enable more individuals and firms to have access to credit (borrowed funds). This is important because without access to credit, ELDCs do not have the necessary money to invest in the economy, which hinders innovation and skills development of the workforce. Money borrowed from the banking system can be used by firms for investment purposes to increase production and productive capacity. In turn, this would help the economy to have a higher level of employment, less inequality in income and less poverty.

While commercial banks all across the world offer credit to individuals and businesses, this is not readily available in most ELDCs. People living and working in rural parts of the country, where the poorest communities tend to be, may not have any access to any banks. These communities are unable to break the poverty cycle because they do not have any opportunity to access finance to engage in economic activities of their own.

This was a major reason for the introduction of microfinance schemes, which were originally developed in Bangladesh to address the inaccessibility of banking services for the poorest members of society from commercial banks. **Microfinance** refers to small sums of borrowed funds by individuals in ELDCs for self-employment purposes so that they can generate income for themselves and their families. These sums help to promote an entrepreneurial culture, thereby stimulating economic growth and development. The idea of microfinance was developed by Bangladeshi banker, economist and Nobel Peace Prize winner Muhammad Yunus (2006). The World Bank and the United Nations are supporters of microfinance as a form of economic and human development, especially as it can be used to empower women.

◆ **Microfinance** refers to small sums of borrowed funds by individuals in ELDCs for self-employment purposes so that they can generate income for themselves and their families.

Advantages of microfinance as a strategy to achieve growth and development include:

- The system advances small-sized loans to people, especially women, who cannot ordinarily borrow money from commercial financial institutions.
- They can simultaneously enhance gender equality and income distribution, by empowering women and supporting female entrepreneurs.
- Those who qualify for microfinance schemes are required to attend classes on financial management. Thus, such schemes have educational and social benefits to the economy.

The main limitations of microfinance schemes include the following arguments:

- Most providers of microfinance will charge relatively high interest rates owing to the high risks of lending money to the poor who lack any financial assets as security.
- The lack of property rights in ELDCs is a major source of poverty because individuals have little, if any, collateral (financial security on their assets) to borrow money to fund development projects. This also reduces the credit rating or credit worthiness of the borrower.
- They do not create job opportunities on a large enough scale to reduce the overall level of unemployment or to increase the ELDC's productive capacity.

- Self-employed micro-entrepreneurs in ELDCs often have unstable incomes yet endure a huge degree of risk, so many people who qualify for microfinance choose to take jobs at factories instead as there is greater financial stability. This clearly defeats the purpose of microfinance schemes.
- Micro-entrepreneurs tend to have limited qualifications, skills, training and experience (as they lack opportunities and access to education). Hence, there is a high chance that the funds are misused or used ineffectively, thus leading to further business failures.

Mobile banking is also gaining its popularity in many developing countries in Asia, Africa and South America. **Mobile banking** is a service provided by a bank or other financial institution that allows its customers to conduct financial transactions remotely using a mobile device, such as a smartphone, tablet computer or laptop.

◆ **Mobile banking** is a service provided by a bank or other financial institution that allows its customers to conduct financial transactions remotely using a mobile device, such as a smartphone, tablet computer or laptop.

Mobile banking started with basic services, using Short Message Service (SMS) banking services, where the enrolled customer used to get information about certain transactions through SMS. Since then, many banks have taken advantage of high-speed internet technologies to offer a full range of banking services to the masses who use mobile devices, thereby enabling many more people to access banking services without having to physically go to a bank to do so. Owners of a basic mobile phone with internet connection can generally access the following mobile banking services:

- account information, such as mini-statements, account balances, transaction alerts, business loans and interest rate charges
- banking transactions, such as fund transfers to third parties, online bill and tax payments, and foreign currency exchange
- investment information, such as portfolio management support and direct buying/selling of stocks and shares
- online support from the banks regarding their banking services, including credit facilities.

Figure 36.17 Mobile banking

CASE STUDY

M-Pesa offering financial opportunities in Kenya

M-Pesa (M for mobile, *pesa* is Swahili for money) is a mobile phone-based money transfer, financing and microfinancing service. It was launched in 2007 by Vodafone for Safaricom and Vodacom, the largest mobile network operators in Kenya and Tanzania. It has since expanded to Afghanistan, South Africa, India, Romania and Albania. M-Pesa allows users to deposit, withdraw and transfer money as well as pay for goods and services easily with a mobile device.

As a mobile banking service, M-Pesa does not have any branches. M-Pesa customers can deposit and withdraw money from a network of agents that includes airtime resellers and retail outlets acting as banking agents. With lower operating costs, it can reach out to a greater number of people. Users are charged a small fee for sending and withdrawing money using the service. The service has been praised for giving millions of people access to the formal financial systems and for reducing crime in otherwise largely cash-based societies. Within the first five years of its operations, M-Pesa had registered about 17 million accounts in Kenya, and about 7 million accounts in Tanzania.

Figure 36.18 Microfinance is used to reach more people

Increasing women's empowerment

A secondary institutional change strategy to achieve growth and development is increasing the empowerment of women. Gender inequality is a significant barrier to economic and human development because it limits the quantity and quality of labour resources in the production process and deprives women of the economic opportunities and financial independence. Economists believe that empowering women is a vital development strategy to reducing poverty in ELDCs. Ignoring gender disparities is detrimental to the well-being of women and their children, and to the economic and human development of all countries, but ELDCs in particular.

The empowerment of women (gender equality) helps to end social and cultural discrimination against females. For example, in some countries, women are not allowed to work with men, vote or even to drive, which hinders labour productivity and national output. The empowerment of women can also have a huge impact on their self-esteem and mental well-being. In the long-term, this has positive effects on maternal health and reduces child mortality.

The World Development Report from the World Bank recommends four priorities to promote gender equality and to empower women:

1. Reduce excess female mortality and close education gaps where they remain. Improving healthcare provision for women and educating females are two of the most significant investments for fostering human development, especially in ELDCs.
2. Improving access for women to participate in economic activity, such as increasing opportunities for females to partake in the workforce and improving women's access to credit and other productive resources.

3 Increasing women's voice and agency in the household and society, thereby improving the ability of females to make effective choices and to transform their choices into desired outcomes. This includes freedom from the risk of violence.
4 Limiting the reproduction of gender inequality across generations to achieve sustainable human and economic development.

Countries that fail to empower women and promote gender equality, such as Ghana and Rwanda, generally face the problem of lower productivity, slower economic growth and hindered human development as approximately half of their population are denied of the opportunities to improve their standard of living.

Reducing corruption

Corruption is the act of being dishonest or fraudulent by those in a position of power and authority. This will typically involve bribery, embezzlement, deceit, manipulation and abuse of power. A well-structured and well-enforced legal system and good governance are essential to creating an institutional framework and an environment conducive to national and international trade. Dishonest or illegal conduct by government officials is immoral, inefficient, and detrimental to economic growth and development. Bribery, for example, is considered to be part of the informal or unofficial sector of the economy so would not be counted in the national income of a country. Moreover, corruption adds financial burdens to producers and consumers.

World Bank economists Daniel Kaufmann and Cheryl W Gray (1990) observed the following characteristics and outcomes of bribery and corruption:

- Bribery is more dominant in ELDCs and emerging countries (middle-income countries), although there are regional disparities.
- Corruption increases insecurity and low confidence in an economy.
- Bribery encourages activities done in informal (unofficial) sectors, thus limiting the country's reported gross domestic product (GDP).
- Bribery and corruption reduce tax revenues collected by the government, which negatively impacts the government's ability to raise money to provide public and merit goods.
- Corruption tarnishes the legality and dependability of a state, thereby putting off investors and inward foreign direct investment.
- Corruption impedes economic growth and development in the long run.

Since 1995, Transparency International has compiled its Corruption Perceptions Index (CPI) the world's leading indicator of public sector corruption. Countries are ranked on a scale from 0 (highly corrupt) to 100 (very clean/honest). The lower the CPI score, the more corrupt the country is perceived to be. Table 35.8 shows a clear correlation between the CPI and the level of economic development of a country.

Top tip!

Good governance applies to anyone in a position of responsibility with decision-making power, not just government officials, but also lawmakers, the military, educators, scientific researchers and religious leaders.

ATL ACTIVITY 36.4

To gain a better understanding of the concept of corruption, investigate the following:

1. What is corruption?
2. What are the social and economic costs of corruption?
3. How might corruption in your country affect you personally?
4. What are the main global solutions for tackling corruption?
5. Which **three** of the WISE ChoICES concepts do you think best apply to the economics of corruption? Justify your reasons for these.

Be prepared to share your findings to the above tasks. This may be a good starting point for your research: https://www.transparency.org/en/what-is-corruption

Common mistake

Students often claim that eradicating corruption is a prerequisite to economic development. It is unlikely that corruption can be eradicated (it exists in economically developed countries too), but development occurs when corruption is reduced.

Property rights

Property rights give the owner of an asset the exclusive authority to determine how the resource is used, whether the asset is owned by an individual, group, firm or government. They are the legal entitlement to both physical and intangible assets of their owners. Examples include physical properties (such as cars, consumer durables and houses) and intellectual property rights (such as copyrights, trademarks and patents).

◆ **Property rights** are the entitlement to both tangible and intangible assets owned by an individual, organization or government, such as land rights and intellectual property rights.

■ **Figure 36.19** Property rights protect the legal entitlements of the owners

The proper planning, implementation and enforcement of property rights would help business owners and governments to feel assured of their investments in a country. By contrast, the loose articulation of property rights that are casually enforced would discourage domestic investors and overseas firms in the form of inward foreign direct investment (FDI).

There are four main types or categories of property rights of an economic good or service:

- the right to use the asset
- the right to earn income from the asset
- the right to transfer the good or service to others
- the right to enforce property rights of the asset.

One of the biggest challenges for many ELDCs in dealing with property rights is the common exploitation of natural resources, as governments often find it difficult to monitor or safeguard these properties. The exploitation of common pool resources was covered in Chapter 11 in the discussion of the tragedy of the commons. This is the problem of a lack of property rights, which leads to over-consumption, under-investment and ultimately depletion of the natural resource. Hence, the lack of property rights creates a major barrier to economic growth and development.

Land rights

The final example of institutional change is the extension and enforcement of land rights. **Land rights** refer to the ability of individuals to obtain, use and hold land at their will. For example, landowners (those with legal land rights) can allow third parties to access their land for which they earn rental income. Land ownership is an important source of security, income and wealth. For many people in ELDCs, land rights are vital in giving them a basic means of living. Research findings from several United Nations agencies indicate a direct correlation between the absence of land rights and poverty and hunger in a country. A higher quality of life can also be achieved by providing, extending and enforcing land rights to poor people in rural areas of the country.

◆ **Land rights** refer to the ability of individuals to obtain, use and hold land at their will.

Figure 36.20 Land rights are an important source of security, income and wealth

Globally, there has been an increased focus on land rights, as governments across the world engage with land reform and policies for dealing with sensitive issues such as the distribution of land. However, in many ELDCs, political and cultural differences and dysfunctional legal and institutional facilities thwart their ability to take advantage of land reform laws.

Another important issue related to land rights is the land ownership rights of indigenous people on indigenous land. Governments and policymakers often struggle to find ways to recognize and protect indigenous and community land rights. This is an important matter for the indigenous population as they are dependent on primary resources such as hunting, fishing and consuming forest resources. It is also important to their well-being owing to sensitivities regarding religious connections of the land and self-identify. Essentially, the lack of land rights or their poor management is a significant source of underdevelopment, conflict and environmental degradation in many communities.

Strengths and limitations of strategies for promoting economic growth and economic development (AO3)

The 10 strategies discussed earlier in this chapter, used to achieve economic growth and development, all have their relative strengths and limitations. Governments and policymakers use different combinations of these strategies to promote economic growth and development. However, the strengths and weaknesses of the strategies are dependent on different factors, including the context of the country, budgetary constraints, political and social influences, and the degree of good governance.

In addition, what works for one country at a particular point in time may not necessarily be the case for other nations. For example, incentives provided for exporters in Ethiopia in the manufacturing sector were insufficient to motivate the private sector to engage in an export-driven growth strategy, making export promotion an ineffective strategy. By contrast, the US government's initiative Export Promotion Programs, launched in 1993, has been hugely successful in helping many firms to engage in export promotion. The programme involves the US government assisting domestic firms with their export-driven strategies by providing grants, subsidies, tax breaks and information about the importing countries. Some commentators argue such assistance from governments of MEDCs creates unfair advantages and prevents firms from ELDCs being able to thrive or survive.

Another example of context being important in the evaluation of strategies to achieve growth and development is population size. In general, the larger the population, the more challenging it is for a government's policies to have the desired impact on everyone in the country. For instance, the highly debated one-child policy used in China between 1979 and 2015 was empirically proven to aid economic development in the world's most populous country but caused a major imbalance in favour of boys. In 2016, when the policy was abolished, there were 33.59 million more males than females in China. Although the policy was deemed to be effectively implemented and enforced owing to the prevailing economic system in China, it would be extremely difficult and inappropriate for many free-market and democratic countries.

Figure 36.21 China's one-child policy has caused a large imbalance in the gender ratio

Other contextualized examples of growth and development strategies, which work for some countries but certainly not all, include:

- The Philippines relies on the empowerment of women to work outside of the country, which enables them to send back valuable remittances to the country. The remittances can then be used in the Philippines for consumption expenditure by households and investment expenditure in physical and social capital.
- Highly industrialized countries like Taiwan and South Korea use their technical expertise to gain a competitive advantage in the production of technologically advanced products and services.
- Gulf countries (such as the United Arab Emirates), Indonesia and Malaysia used the principles of Islamic banking to fund projects related to economic development with considerable success.
- Good governance of growth and development plans and finances have enabled Uruguay, Bhutan and Botswana to uphold developmental initiatives, whereas Yemen, Venezuela and Libya have squandered resources set aside for developmental initiatives. The lack of transparency in government affairs and accountability has prevented these countries from achieving their growth and development targets.

Top tip!

When evaluating the effectiveness of strategies used to achieve growth and development, it can be useful to discuss the relative effectiveness between interventionist-based policies and market-based policies, written in the context of real-world examples.

Strengths and limitations of government intervention versus market-oriented approaches to achieving economic growth and economic development (AO3)

The strengths of government interventionist policies to achieve economic growth and development include:

- The provision of essential infrastructure – Without government intervention, there would be a lack of **infrastructure** (the physical structures required for the effective operation of society). Examples include road networks, railways, ports, telecommunications networks, sewage structures and flood control systems.
- Investment in **human capital** – The private sector is unlikely to provide sufficient investments in human capital through education and training, especially in ELDCs. Thus, interventionist policies are required to encourage more provision of such merit goods.
- Establishment of a **stable economy** – Development requires government intervention to provide a safe and stable economic environment to protect the interests of all members of society by using interventionist demand- and supply-side policies. A stable macroeconomy is also conducive to attracting inward foreign direct investment to further improve growth and development.
- Provision of a **social safety net** – Interventionist policies through direct government provision and a social welfare system help to ensure that all members of society have access to basic necessities, thus preventing absolute poverty in the economy.
- Interventionist policies can be used to tackle **income and wealth inequalities**, which hinder the development and prosperity of ELDCs. For example, cultural and historical contexts in many countries mean that women are not given the same opportunities as men. Thus, intervention is necessary to correct such disparities.
- Intervention is also required when a country faces a major **emergency or disaster**, such as civil war or the outbreak of an infectious pandemic. Without intervention, the productive capacity of the country will decline along with a fall in FDI, employment and the overall standards of living.

However, there are limitations of government interventionist policies to achieve economic growth and development. These include:

- Excessive **bureaucracy** – This refers to administrative systems, formal structures, rules and regulations that govern economic activity. The administrative procedures of interventionist policies in ELDCs often lead to economic inefficiencies rather than economic growth and development.
- **Poor planning** – Political instability and conflict, commonly found in many ELDCs, can cause major delays in production, thus limiting the effectiveness of interventionist policies to achieve economic growth and development. The lack of market signals and incentives (the market forces of demand and supply) mean that government planning is often unrealistic, thereby limiting the opportunities for growth and development.

Top tip!

While many ELDCs share common characteristics, it can be rather naïve to assume that *all* ELDCs have the same features and that there are no exceptions to these generalizations. For example, although GDP per capita may be low in ELDCs, there can certainly be extremely wealthy individuals in the country.

- **Corruption** – Dishonest and deceitful governments that misuse sources of public sector finances limit the effectiveness of macroeconomic policies intended to promote growth and development. Corruption among public sector officials reduces trust between individuals, firms and governments, thereby acting as a deterrent to inward FDI.

The strengths of market-oriented policies to achieve economic growth and development include:

- **Efficiency** – The key benefit of market-oriented growth and development policies, such as deregulation and privatization, is that resources are allocated more efficiently than through government intervention, which distorts market forces. This helps to increase the level of economic activity.
- **Competitiveness** – Market-based strategies such as labour market reforms create incentives to work and invest in the economy. These policies therefore help to improve labour market flexibility and productivity, resulting in a more internationally competitive labour force.
- **Economic growth** – The profit motive in free markets encourages business owners to work hard and to take entrepreneurial risks, such as expenditure on investment and innovation. Thus, market-oriented policies have a positive impact on economic growth and development in the long run.
- **Benefits of free trade** – Free and fairer trade policies can lead to many benefits, such as increased consumer choice, lower prices and improved quality. It also enables firms to sell to more customers, beyond the borders of their country. This inevitably contributes to greater profits, employment, economic growth and development. By contrast, interventionist policies such as tariffs and quotas often cause trading partners to retaliate, which reduces choice but forces up prices of exports and imports.
- **Investment opportunities** – The liberalization of trade and capital flows through market-based policies helps to reduce barriers to international trade. This is an important factor in attracting inward foreign direct investment, that is, the capital expenditure of multinational corporations in overseas economies. Again, this helps the economy to develop over time.

PAPER 1 EXAM PRACTICE QUESTION 36.5

Using real-world examples, discuss the strengths and limitations of government intervention and market-based approaches to achieving economic growth and/or development. *[15 marks]*

However, there are limitations of using market-oriented policies to achieve economic growth and development. These include:

- **Market failures** – The inability of market-oriented policies to deal with market failures is their main weakness. For example, negative production and consumption externalities are not dealt with (see Chapter 11). ELDCs also lack sufficient funds to provide merit goods such as education and healthcare, so there is a strong case for government intervention to correct such market failures.
- The development of a **dual economy** – This occurs when two distinct economic sectors exist within a country, with different levels of development. It is common in ELDCs to have a low-income agricultural sector catering for local demand of the majority of the population and another manufacturing sector for export-driven international markets. This creates growing disparities between the two sectors in terms of productivity and profits.

- **Income and wealth inequalities** – The advantages of economic growth and development do not always automatically trickle down to benefit the poorest members of society, so government intervention is required to tackle the problems of income and wealth inequalities. Such inequalities (see Chapter 23) can cause major issues, such as poverty, thereby limiting the effectiveness of market-based strategies to achieve economic growth and development.

Owing to the relative advantages and disadvantages of interventionist policies and market-based policies, a complementary approach may be the best way for countries to achieve economic growth and development. Evidence suggests that neither extremes work well in the real world.

PAPER 2 EXAM PRACTICE QUESTION 36.6

When Mozambique gained independence in 1975, it was one of the world's poorest countries. Socialist policies, economic mismanagement, and a brutal civil war from 1977 to 1992 further impoverished the country. In 1987, the government embarked on a series of macroeconomic reforms designed to stabilize the economy. These steps, combined with donor assistance and with political stability, propelled the country's GDP, in **purchasing power parity** terms, from $4 billion in 1993 to about $37 billion in 2017.

Mozambique's once substantial **foreign debt** was reduced through debt forgiveness and debt rescheduling under the IMF's Heavily Indebted Poor Countries (HIPC) initiative. However, in 2016, information surfaced revealing that the government was responsible for over $2 billion in government-backed loans secured between 2012 and 2014 by state-owned defence and security companies without parliamentary approval or national budget inclusion. This prompted the IMF and international donors to halt direct budget support to Mozambique.

Table 36.5 GDP composition (2017)

Agriculture	Industry	Services
23.9%	19.3%	56.8%

Table 36.6 Labour force by occupation (2015)

Agriculture	Industry	Services
74.4%	3.9%	21.7%

Table 36.7 Key economic indicators (2016 and 2017)

Indicators	2016	2017
Inflation	15.3%	19.2%
Export in billion ($bn)	3.3	4.7
External debt ($bn)	10.48	10.91

a Define the following terms:

i *Purchasing power parity (PPP).* *[2 marks]*

ii *Foreign debt.* *[2 marks]*

b Using information from the text/data and your knowledge of economics, discuss methods that the government of Mozambique might employ to achieve greater economic growth and/or economic development. *[15 marks]*

Progress towards meeting selected Sustainable Development Goals in the context of two or more countries (AO3)

The Sustainable Development Goals were announced by the United Nations Development Programme in 2015 (see Chapter 33), detailing the 17 goals to be met by all member states by 2030. The progress and review processes are reflected in the annual SDG Progress Report, prepared by the Secretary-General.

The final requirement of the IB Economics syllabus is that you need to be able to discuss and evaluate the progress towards meeting selected Sustainable Development Goals in the context of two or more countries. A broad summary of the global situation is outlined below, taken from the Sustainable Development Goals Report 2019.

Goal 1: No poverty

The Sustainable Development Goals Report 2019 indicates that the world is not on track to end poverty by 2030. Of the world's population, 55 per cent has no access to social protection. Forecasts suggest that without significant shifts in policy, extreme poverty will still be in double digits in sub-Saharan Africa by 2030. Natural disasters continue to cause enormous suffering; between 1998 and 2017, direct economic losses from natural disasters were estimated to be almost $3 trillion, making ELDCs and emerging countries even poorer.

Goal 2: Zero hunger

The number of people suffering from hunger has been on the rise since 2014. Diseases continue to affect the growth and cognitive development of millions of children. The number of people not getting adequate food is rising. Almost 820 million people were undernourished in 2017; the same figure was observed in 2010.

Goal 3: Ensure healthy lives and promote well-being for all at all ages

There has been satisfactory progress in improving the health of millions of people. Maternal and child mortality rates have been reduced, life expectancy continues to increase globally, and the fight against some infectious diseases has made steady progress. However, outbreaks of measles and diphtheria have resulted in many preventable deaths.

Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all

Millions of children are still out of school, and not all who do attend are learning. More than half of all children and adolescents worldwide are not meeting minimum proficiency standards in reading and mathematics. Disparities in educational opportunities and outcomes are found across all regions, while sub-Saharan Africa and parts of Central and Southern Asia lag behind. As a result, many students are not adequately prepared to participate in an overly complex global economy. Around 750 million adults still cannot read and write a simple statement; two-thirds of those adults are women.

Goal 5: Achieve gender equality and empower all women and girls

Although discriminatory laws and social norms remain pervasive, many women and girls continue to suffer from harmful practices including all forms of violence against females. However, fewer girls are forced into early marriage, more women are serving in parliament and

positions of leadership, and laws are being reformed to advance gender equality. Women and girls perform a disproportionate share of unpaid care and domestic work. At home, at work and in political life, many women continue to be denied decision-making power.

Goal 6: Ensure availability and sustainable management of water and sanitation for all

Donors have increased their aid commitments to the water sector. Most countries have recognized the need for better co-ordination of the use of water resources. Most rivers in Africa, Asia and Latin America are more polluted now than they were in the 1990s. An estimated 50 to 70 per cent of the world's natural wetland areas have been lost over the last 100 years. While substantial progress has been made in increasing access to clean drinking water and sanitation, billions of people (mostly in rural areas) still lack these essential services.

Goal 7: Ensure access to affordable, reliable, sustainable and modern energy for all

The world is making progress towards this goal with encouraging signs that energy is becoming more sustainable and widely available. Access to electricity in poorer countries has begun to accelerate, while energy efficiency continues to improve, and renewable energy is making impressive gains in the electricity sector. However, around 1.3 billion people do not have access to electricity, yet reaching the unserved will require increased efforts.

Goal 8: Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all

Many ELDCs have achieved economic growth but the UNDP's 7 per cent target rate of growth is still out of reach. Labour productivity is rising, although wide disparities exist between different regions. Although the global unemployment rate dropped steadily between 2015 and 2019, it remains high in some ELDCs and among youth workers. The Great Lockdown caused by the worldwide outbreak of the coronavirus in 2019–20 halted growth for the world economy, derailing the progress made for this goal. Even in the USA, which saw the highest number of deaths from the Covid-19 global pandemic, a record 22 million people were made unemployed in a single month (April 2020), which took the unemployment rate beyond 20 per cent.

Goal 9: Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation

ELDCs need to accelerate the development of their manufacturing sector if they are to meet the 2030 target, In particular, they need to increase their investment in scientific research and innovation. Despite recent progress, industrialization in ELDCs is still too slow to meet the target by 2030. Small-scale industries in the poorest countries lack the financial services they need to grow and develop.

Goal 10: Reduce inequality within and among countries

Income inequality continues to rise in many parts of the world, even as the poorest members of the population in most countries experience income growth. Greater effort is needed to reduce inequalities, including those related to labour market access and trade. All countries alike can benefit from policies promoting equality and inclusivity in pursuit of growth and development.

Goal 11: Make cities and human settlements inclusive, safe, resilient and sustainable

Since 2007, more than half the world's population has been living in cities, and that share is projected to rise to 60 per cent by 2030. Rapid urbanization and population growth are outpacing the construction of adequate and affordable housing. Access to public transportation is increasing, but faster progress is needed in developing and rural regions.

Goal 12: Ensure sustainable consumption and production patterns

Material footprint (the total amount of raw materials extracted from the planet to meet the economy's final consumption) is increasing across the world. The lifestyles of many people in the wealthiest nations are heavily dependent on resources extracted from poorer countries. The unfavourable developments in energy consumption reported for SDG 7 have resulted in a deterioration of the overall assessment of SDG 12. This means countries are not doing enough to ensure responsible consumption and production.

Goal 13: Take urgent action to combat climate change and its impacts

According to the UNDP, climate change is the most defining issue of our time and the most significant challenge to sustainable development. The compounding effects of climate change are speeding up, leaving even less time for people, firms and governments to act. Unprecedented changes in all aspects of society will be required to avoid the worst effects of climate change. Climate-related financial flows have increased but are small compared to the scale of the problem and still overshadowed by investments in fossil fuels.

Goal 14: Conserve and sustainably use the oceans, seas and marine resources for sustainable development

Land-based pollutants and marine debris threaten coastal habitats and ecosystems. An increase in acidification is also threatening marine life and hampering the ocean's role in moderating climate change. However, improvements in water quality are being made. Protection of marine areas has doubled since 2010, but more must be done to safeguard key biodiversity areas.

Goal 15: Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss

The loss of forests is slowing but continues at an alarming rate; 1 million plant and animal species are at risk of extinction and an estimated 20 per cent of the Earth's land area was degraded between 2000 and 2015. Land degradation is affecting one-fifth of the Earth's land area and impairing the lives of 1 billion people. Progress in the protection of key biodiversity areas needs to accelerate if SDG 15 is to be met by 2030.

Figure 36.22 Climate change is the most significant challenge to sustainable development

CASE STUDY

Australia's bushfire crises

The 2019–20 Australian bushfire crisis was one of the worst bushfires in history and caused widespread damage to each state and territory. Bushfires, caused by climate change, lasted from September 2019 to March 2020, that is, for eight consecutive months. The fires burned an estimated 18.6 million hectares (46 million acres or 186,000 square kilometres) of land, destroying buildings, habitats and lives. An estimated 1 billion animals were killed, and the air quality dropped to hazardous levels, NASA estimating that 306 million tonnes of CO_2 had been emitted into the atmosphere.

Figure 36.23 Bushfires in Kakadu National Park, Northern Territory, Australia

Goal 16: Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels

Realizing the goal of peaceful, just and inclusive societies is still a long way off. In recent years, no substantial advances have been made in ELDCs towards ending violence, promoting the rule of law, strengthening institutions at all levels, or increasing access to justice. Millions of people have been deprived of their security, human rights and economic opportunities, while attacks on human rights activists and journalists are holding back development. Even in MEDCs, deaths due to homicide or terrorist attacks have kept governments on high alert.

Goal 17: Strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development

Global support for implementing the SDGs is gaining momentum, but major challenges remain. A growing share of the global population has access to the internet and mobile banking for ELDCs has been established, yet a huge digital divide persists. Personal remittances are at an all-time high, but Official Development Assistance (ODA) is declining and very few countries contribute the UN-recommended 0.7 per cent of their GDP to ODA (see Table 36.4). However, the demand- and supply-side shocks caused by the coronavirus pandemic have disrupted progress in meeting this goal as countries work on interventionist measures to help their respective economies recover from the most damaging economic crisis in the last century.

ATL ACTIVITY 36.5

For any two countries of your choice, investigate the progress made in meeting the 17 Sustainable Development Goals. Identify the goals in which these countries have shown the most and least improvement so far. Examine the reasons for this. You may find it beneficial to work in small groups.

WISE ChoICES

With reference to the Sustainable Development Goals, evaluate the view that **change** will bring about a greater level of **economic well-being**.

Chapter summary

- Governments have adopted a variety of strategies to tackle the barriers to economic growth and development in order to achieve a higher standard of living for the people of their countries.
- These strategies include the use of trade strategies, diversification, social enterprises, market-based policies, interventionist policies, provision of merit goods, inward foreign direct investment, foreign aid, multilateral development assistance and institutional changes.
- There are potential gains from international trade that can help countries to experience economic development, such as export revenue for domestic producers and increased employment opportunities for workers.
- Trade strategies include import substitution (inward development strategy), export promotion (outward development strategy), and economic integration (through a complete or partial removal of tariff and non-tariff restrictions on international trade).
- Diversification helps to overcome the problems of over-specialization (which tends to limit economic growth and development for many ELDCs) and creates new employment opportunities.
- A social enterprise aims to maximize improvements in human and environmental well-being, such as offering support for vulnerable groups such as children, the elderly or unprivileged persons.
- Market-based policies focus on increasing the productive capacity of the economy to achieve economic development. These policies include trade liberalization, privatization and the deregulation of markets.
- Interventionist policies are used to correct market deficiencies, by redistributing resources, incomes and wealth. The government can use redistributive policies including progressive tax policies, transfer payments and minimum wages to ensure a social safety net for all members of society.
- The provision of merit goods is another strategy used to ensure the entire population of the country can consume these socially beneficial goods and services, irrespective of the differences in the ability to pay. The provision of merit goods includes education programmes, health programmes and infrastructure programmes.
- The government's investment in infrastructure includes energy, transport, telecommunications, clean water and sanitation in order to sustain a decent standard of living and to attract inward foreign direct investment.
- Globalization and the promotion of freer international trade from the World Trade Organization (WTO) have also encouraged a significant increase in FDI in ELDCs.
- Foreign aid makes up for the shortcomings of the free market that fails to help ELDCs during times of need, such as emergency relief aid following a natural disaster or civil war.
- Foreign aid to ELDCs can be granted by governments of donor countries or by non-governmental organizations (NGOs), such as Oxfam and World Vision.
- Foreign aid comprises humanitarian aid/development aid, debt relief, Official Development Assistance (ODA) and intermediation from non-governmental organizations (NGOs).
- Multilateral institutions, such as the World Bank and the IMF are international organizations made up of member governments around the world. These member states pool resources together to fund large-scale development programmes.
- Institutional change is a final growth and development strategy. This consists of five areas of change: improved access to banking (including microfinance and mobile banking), increasing women's empowerment (gender equality), reducing corruption, property rights and land rights.
- Governments and policymakers use different strategies to promote economic growth and development. However, the strengths and weaknesses of these strategies are dependent on different factors, for instance context, the economic system in the country, the time when the policy was implemented and the degree of good governance.
- Government intervention and market-oriented approaches have their respective advantages and disadvantages to achieving economic growth and development. Governments use both types of strategy to achieve their desired macroeconomic development goals.
- Although most governments across the world have adopted the UNDP's 17 SDGs in 2015, commitment has varied between countries and regions. At the time of writing, it is unlikely that any of the SDGs will be met by the 2030 target year.

Glossary

Abnormal profit refers to profit that is greater than normal profit, meaning that a firm is earning more profit than in its next best alternative.

Absolute advantage occurs when a country can produce more of a good or service than another country using the same amount of resources (or being able to produce the same amount of a good or service using fewer resources).

Absolute poverty exists when people are deprived of basic human needs for survival. Often these people are living below the *international poverty line* or their country's own *national poverty line*.

Actual output refers to the current, rather than potential, level of real GDP in the economy. It is represented by a movement towards the country's PPC as more resources are employed.

Administrative barriers are the application of bureaucratic standards and regulations imposed on foreign firms in order to protect domestic firms and consumers.

An **ad valorem tax** imposes a percentage tax on the value of a good or service. Examples include property taxes, tariffs (taxes on imports) and sales taxes.

Adverse selection is a form of opportunistic behaviour that refers to the undesired decisions or results that occur when buyers and sellers have access to asymmetric information.

Aggregate demand (AD) is the value of all goods and services demanded in the economy, per time period.

Aggregate supply (AS) is the total amount of output of goods and services that firms within an economy are willing and able to supply at a given time and at an overall price level.

Allocative efficiency is the socially optimal situation that occurs when resources are distributed in such a way that consumers and producers get the maximum possible benefit, that is, no one can be made better off without making someone else worse off.

Anchoring is a cognitive bias that influences how people view a product by comparing it to something else.

Ancillary services are firms that exist to support firms in an industry. These are often services such as legal, packaging or shipping services.

Anti-competitive measures occur when firms engage in practices that either prevent other firms from joining their market, or aim to make it hard for existing firms to remain in the market.

Anti-monopoly regulation are laws that control or limit the restrictive practices and market power of dominant firms in an industry.

Appreciation is a sustained increase in the value of one currency in terms of another under a floating exchange rate system.

Asymmetric information refers to missing, unbalanced or incorrect information that exists when one economic agent has more information than the other in an economic transaction.

Austerity measures refer to cutbacks in fiscal spending in order to repay national debt. This is achieved by reducing government spending and raising taxation.

Automatic stabilizers are aspects of fiscal policy that naturally reduce fluctuations in the level of economic activity, thereby stabilizing the rate of growth in real national output.

Availability bias refers to the ease with which an idea or event can be recalled easily from memory. This causes people to overestimate the likelihood or probability of the event or idea ever happening, thereby distorting rational decision-making.

Average costs (AC) are the unit costs of production, that is, the cost of producing one unit of output.

Average fixed costs (AFC) are the production costs that do not change with the level of output.

The **average rate of tax** refers to the amount of tax paid compared to the amount of income earned, that is, the total tax paid divided by the total income for an individual, or $T \div Y$. (HL only)

Average revenue (AR) refers to the price received from the sale of a good or service.

Average variable costs (AVC) are production costs incurred directly per unit of output for a particular good or service.

The **balance of payments** is a financial record of a country's transactions with the rest of the world, usually over one year.

It includes the country's trade in goods and services with other countries.

The **balance of trade** (or **trade balance**) is the difference between a country's total export earnings and its total import expenditure on both goods and services.

Barriers to trade are obstacles to international trade, imposed by a government to safeguard national interests by reducing the competitiveness of foreign firms.

The **basic economic problem** refers to how best to allocate the economy's scarce resources to best meet its unlimited wants. It is about addressing the fundamental questions in economics, that is, *what*, *how* and *for whom* production should take place.

Benign deflation is generally unharmful deflation as the economy is able to produce more without an increase in the general price level (shown by an outwards shift of the SRAS curve).

The **Better Life Index** (BLI) is an alternative to national income as a measure of well-being. The BLI is based on 11 topics identified by the OECD to be essential in terms of material living conditions and quality of life.

A **bilateral trade agreement** is a contractual trade arrangement between two countries, usually by mutual agreement to reduce or remove barriers to trade.

A **boom** is a phase in the business cycle when the level of economic activity rises, caused by an increase in aggregate demand.

Bounded rationality theory states that rationality is limited. This means that people's cognitive decision-making capacity is limited owing to barriers such as information failure or the time available to make the decision.

Bounded self-control means people may lack the self-control to think independently in order to make a rational choice, as they conform to social norms and group preferences, which may not match their own preferences.

Bounded selfishness means that people are not necessarily selfish and only act in self-interest, but often behave reciprocally and altruistically. People do not always pursue their own pure self-interest.

Break even means a firm does not make a profit or a loss. This occurs when price is equal to average cost, or total revenue equals total costs.

A **budget deficit** exists when the value of government spending exceeds government revenue per time period, that is, $G > T$.

The **business cycle** (or **trade cycle**) is a model that illustrates the fluctuations in a country's level of economic activity over time. It determines the long-term trend of economic growth in the economy.

The **capital account** records the different forms of capital inflows and outflows of a country during a given time period, namely capital transfers and transactions in non-produced, non-financial assets.

Capital expenditure refers to long-term items of government spending (public sector investments) that boost the economy's productive capacity.

Capital flight is the withdrawal of money, assets and resources from a country due to economic and political uncertainties, thereby hindering investor confidence and economic development.

Capital transfers are payments associated with debt-forgiveness, project-aid for developmental purposes and sums of money/assets brought into an economy by immigrants (inflow, credit) or those leaving with emigrants (outflow, debit).

A **carbon tax** is an indirect tax imposed on producers and consumers for their carbon emissions in order to reduce greenhouse gas emissions and the use of carbon fossil fuels.

A **cartel** is an agreement between oligopolistic firms in the same industry to collude in fixing prices or to restrict the level of output in the market, thereby effectively acting as a monopolist.

The **central bank** is the monetary authority responsible for regulating the country's financial system and implementing monetary policy, including regulation of commercial banks and the country's money supply.

Ceteris paribus is a Latin phrase meaning 'all other factors remaining constant' or 'all else unchanged'. It is used by economists to explain cause and effects of economic variables.

A **change in demand** refers to a shift in the demand curve for a product, due to a change in non-price factors affecting demand.

A **change in quantity demanded** refers to a movement along a demand curve, that is, a change in quantity demanded when price changes.

A **choice architect** is an individual or organization who organizes the context in which people make decisions, based on the simplified design of different choices being presented to them.

Choice architecture refers to the way choices are presented to members of society and how different designs affect the choices made. Effective choice architecture helps people avoid making poor choices and irrational decisions.

A **circular economy** describes an economic system in which raw materials, components and other resources are used sustainably to generate output.

The **circular flow of income model** is a macroeconomic tool used to explain how activity and national income are determined.

The **claimant count** is a measure of the number of people who are out of work, actively seeking employment but claiming unemployment benefits.

Classical economics was the main economic school of thought during the eighteenth and nineteenth centuries, which focused on self-regulating markets to allocate resources efficiently.

A **closed economy** is part of the circular flow of income model comprising only domestic economic decision makers, that is, households, firms and the government.

Collective self-governance refers to voluntary communal actions to tackle the problems of negative externalities and the problems associated with the exploitation of common pool resources.

Collusion is where two or more firms act together to reduce competition in a market.

Collusive oligopoly is the agreement between two or more oligopolistic firms to limit competition by restrictive trade practices; for example, price fixing or collectively limiting output.

Command and control (CAC) regulation and legislation refers to the direct rules or laws governing an activity or industry, stating what is permitted and what is illegal.

Common access resources (or **common pool resources**) are rivalrous but non-excludable goods or services, that is, they are not owned by a private individual or firm; for example, public beaches, country parks and commercial fishing.

A **common external tariff** means that all members of a customs union impose the same (common) trade barriers to non-member countries.

A **common market** (or **single market**) is the most integrated trading bloc, consisting of a customs union that allows the free movement of factors of production between member countries.

Community surplus (also known as **social surplus**) is the sum of consumer and producer surplus at a given market price and output, thereby maximizing economic welfare.

Comparative advantage exists when a country can produce a given amount of output at a lower opportunity cost than another country, that is, it gives up fewer resources than other countries in producing a certain good or service.

Competitive supply means the output of one product prevents or limits the production of alternative products, owing to competing resources.

Complements (or **complementary goods**) are products that are jointly demanded; for example, torch and batteries, or printers and ink cartridges.

A **composite indicator** is a statistical method that combines single indicators of economic development into a combined index such as the Human Development Index.

The **concentration ratio** measures the degree of market power in an industry by adding the combined market share of the largest few firms.

Conflicting macroeconomic objectives means that it is not always possible for a government to simultaneously achieve all of its macroeconomic objectives owing to potential trade-offs.

Constant prices refer to the values of real GDP and real GNI as they have been adjusted for inflation over time.

The **consumer price index** (CPI) is a weighted index used to calculate the change in prices of a representative basket of goods and services consumed by the average household. It is used to measure the inflation rate in the economy.

Consumer rationality refers to the assumption that individuals use rational calculations or considerations to make sensible choices that align with their own best interests. It assumes people are driven by self-interest and incentives.

Consumer surplus refers to the gain or benefit to buyers who can purchase a product at a price lower than what they are willing and able to pay for the product. It is calculated using the formula $CS = WTP - P$.

Contracting out is an approach to government intervention in response to public goods by paying a specialized third-party firm with the expertise to provide a public good, such as refuse collection services.

A **contraction** in the quantity supplied is caused by a lower price for the product, thus causing a downward movement along the supply curve.

Contractionary fiscal policy is used to reduce the level of economic activity by decreasing government spending and/or raising taxes to limit consumption (C) and investment (I).

Contractionary monetary policy aims to slow down economic activity usually by increasing interest rates to cut consumption and investment expenditure in the economy.

A **contraction in demand** occurs when there is a fall in the quantity demanded for a product due to an increase in its price.

Corporate income tax (or **corporation tax**) is a direct tax imposed by a national or local authority on the profits of businesses, including foreign firms operating in the country.

Corporate social responsibility (CSR) is about businesses considering the impact of their operations on the society as a whole welfare in a positive and ethical way.

The **Corruption Perceptions Index** (CPI) is a ranking of the level of public sector corruption in different countries.

Cost-push inflation refers to inflation caused by higher costs of production, such as higher labour costs or skyrocketing rents, thereby forcing up average prices.

Costs of production (see total costs) are the aggregate amount of production costs spent on the output of a given good or service. It is the sum of variable and fixed costs.

A **crawling peg** is a form of fixed exchange rate system in which a currency is permitted to fluctuate within predetermined bands of exchange rates.

Credit creation refers to the process by which commercial banks create money from deposits from savers and use these funds as loans to borrowers.

Credit items are payments received from consumers, firms and institutions or governments located outside of the economy.

Credit rating refers to the creditworthiness of a borrower based on the likely ability to repay the debt. It is what lenders look at before they approve any new credit or loans to the customer.

Crowding out occurs when increased government borrowing causes interest rates to rise, thereby reducing private sector investment expenditure.

Currency liquidity refers to the availability of currency to a market. Some governments choose to reduce access to their currency, which can lead to lower liquidity.

The **current account** is a record of all trade flows (exports and imports of goods and services), income flows and income transfers between countries by individuals, firms and governments.

The **current account balance** is the sum total of all items listed in the current account. The balance can be in deficit (when M > X), in surplus (X > M) or zero (X = M).

Current expenditure refers to government spending on goods and services consumed within the (current) year. It is sometimes referred to as **consumption expenditure**.

Current prices refer to the actual prices of goods and services at the time of collecting the data, which are used to report nominal GDP and nominal GNI.

Current transfers are the inflows and outflows of payments that are not made in exchange for anything (e.g. official development assistance, grants, concessionary loans and donations) between an economy and the rest of the world.

A **customs union** consists of member countries in a trading bloc, which engage in free trade with each other but impose a common external tariff when trading with non-member states.

Cyclical unemployment (or **demand deficient unemployment**) is a type or cause of unemployment, resulting from a lack of demand in the economy affecting most, if not all, industries.

Debit items are payments made to consumers, firms and institutions or governments located outside of the economy.

Debt relief (or **debt forgiveness**) is the partial or total remission of foreign debts, especially those owed by low-income countries and ELDCs.

Debt rescheduling means renegotiating the length of time to repay the existing loans of highly indebted nations.

Debt servicing costs are the expenses of financing the national debt, that is, repaying the principal amount borrowed plus interest costs.

The **debt to GDP ratio** shows a country's national debt as a percentage of its GDP (national income).

A **decrease in demand** refers to a leftwards shift of the entire demand curve for a product, caused by unfavourable changes in non-price factors that affect demand.

A **deficit** on an account exists when the total value of debit items exceeds the total value of credit items, over a given period of time.

Deflation refers to the persistent fall in the average price level in an economy over time, that is, lower prices in general.

A **deflationary gap** (also called a **recessionary gap** or **negative output gap**) occurs when macroeconomic equilibrium is below the full employment equilibrium, that is, AD is insufficient to create full employment.

A **deflationary spiral** is a circular process of downward pressure on the general price level caused by a weak economy, which leads to lower wages and confidence levels that in turn cause prices to fall further.

Demand refers to the willingness and ability of customers to buy a good or service at given prices over a period of time.

The **demand for money** refers to the desire of households and firms to hold money (rather than saving it) in order to finance spending and investment.

Demand-pull inflation is inflation caused by higher levels of spending or aggregate demand (AD) in the economy, which drives up the general price level.

A **demand-side policy** refers to any government strategy or plan to influence the level of aggregate demand, such as reducing interest rates to reduce the costs of borrowing money to finance household consumption expenditure (C) and corporate investments (I).

Demerit goods are products that create negative spillover effects (or negative externalities) to others in society from their production and consumption, that is, MSC > MPC.

Depreciation is a sustained decrease in the value of one currency in terms of another under a floating exchange rate system.

Deregulation is the reduction or removal of government rules and directives in an industry, with the intention of creating more competition and greater efficiency.

Desertification is a category of land degradation caused by natural processes or human-induced activities that result in fertile land becoming desert; that is, the over-exploitation of soil and land through human activity.

Devaluation occurs when the price of a currency operating in a fixed exchange rate system is officially and deliberately lowered.

Direct provision occurs when the government directly provides or supplies certain goods and services deemed to be in the best interest of the public.

Direct taxes are levies or charges imposed on personal income and wealth, or on the profits of businesses. They are imposed directly on the person or firm liable to pay the tax.

Discouraged workers are people who have stopped searching for jobs because they have become disheartened having been rejected by prospective employers too many times.

Diseconomies of scale are inefficiencies caused by an increase in the scale of production, leading to increasing long-run average costs of production.

Disinflation refers to a fall in the rate of inflation, that is, prices are still rising, but at a slower pace, rather than a fall in the average price level.

Diversification is an economic growth and development strategy that involves countries broadening their supply of goods and services in export markets.

Dumping is the sale of goods and services by foreign firms at a price lower than the cost of production.

Economic agents are economic decision makers who interact for the purpose of production and consumption. They comprise households, firms and the government.

Economically least developed countries (ELDCs) are low-income countries facing severe structural barriers to sustainable economic development. They have low levels of human capital and are highly vulnerable to economic and environmental shocks.

Economic costs are the explicit and implicit costs of all resources used by a firm in the production process.

Economic goods are the resources and products that are limited in supply and have economic value (utility) to society. The vast majority of products are economic goods.

Economic growth is a key macroeconomic objective that measures the increase in the level of national output over a specified period of time. It can be measured by the annual percentage change in the economy's real gross domestic product (GDP).

Economic integration is the process of countries becoming more interdependent and economically unified.

Economic methodology refers to the processes, practices and principles of economics by using models, theories and assumptions that underlie economic reasoning.

Economic profit (or **abnormal profit**) exists when total revenue exceeds the economic costs of a transaction, thus creating incentives for firms to produce.

Economic sustainability refers to the optimal use of scarce resources in such a way to ensure future generations are not disadvantaged in favour of today's generation.

An **economic system** describes the way in which an economy is organized and run, including alternative views on how resources are best allocated.

Economic thought refers to a historical account of the different economic ideas, beliefs and principles.

Economics is an academic discipline under the category of social sciences, examining how resource allocation can best meet the needs and wants of individuals and societies.

Economies of scale are lower average costs brought about by an increase in the long run scale of production.

Education indicators are a category of single indicators of measuring economic development by using education-related determinants of the quality of life, such as literacy rates and the mean years of schooling.

Elasticity of demand is a measure of how the quantity demanded of a product changes due to a change in a factor that affects demand, such a change in the price of the product or a change in real disposable income.

Embargoes are a form of administrative barrier that involves the use of bans on trade with a certain country, often due to political and/or economic disputes.

Empirical evidence refers to first-hand data and information acquired by observation or experimentation of certain behaviours and patterns.

Employment refers to the use of factors of production in the production process. The term is usually applied to the use of labour resources.

Energy indicators are a category of single indicators of measuring economic development by using factors that create affordable, reliable, sustainable and modern energy for all citizens.

The **Engel curve** is used to demonstrate the relationship between income and the quantity demanded. Normal goods (services and luxuries) have a positive slope, whereas inferior goods have a negative slope along an Engel curve.

Environmental indicators are a category of single indicators of measuring economic development by using environmental issues that influence or determine the quality of life, such as global warming and climate change.

Environmental sustainability refers to the responsible use of the planet's natural resources so that future generations are not compromised on their use of these natural resources.

Equality is the concept concerned with everyone being equal and should have equal recognition. It is about social fairness and collectivism. Equality means there is parity in income (earnings) between individuals, that is, everyone is paid or receives the same so no inequalities exist.

Equilibrium is the condition that holds when a market is cleared of any shortages or surplus. Hence, it occurs at the price where the quantity demanded for a product is equal to the quantity supplied.

Equilibrium price (or **market clearing price**) is the point where the demand for a product matches the supply of the product, in a given period of time.

Equity is the concept concerned with economic fairness in the distribution of resources. Equity means economic fairness, such as those with higher levels of qualifications, skills and experience being paid more, that is, the existence of justified inequalities.

Errors and omissions is a balancing item in the balance of payments used to represent statistical discrepancies when compiling the accounts.

Excess demand occurs when the price is set below the equilibrium, that is, a **shortage** exists.

Excess supply refers to a disequilibrium situation where the price of a product is set above the equilibrium price, thus creating a **surplus** in the market.

Exchange controls are a form of administrative barrier involving restrictions on the quantity of foreign exchange that can be bought or sold by domestic residents.

Exchange rates are the value of one currency expressed in terms of another currency.

Expansion in the quantity supplied is caused by a higher price for the product, thus causing an upward movement along the supply curve.

Expansionary fiscal policy is the use of increased government spending and/or reduced taxes in order to stimulate the level of economic activity.

Expansionary monetary policy aims to boost economic activity by expanding the money supply, achieved mainly by lowering interest rates, to stimulate aggregate demand, thereby closing a deflationary gap.

An **expansion in demand** occurs when there is an increase in the quantity demanded for a product following a fall in its price.

The **expenditure method** measures economic activity by calculating the value of total spending on newly produced goods and services during the year, comprising consumption (C), investment (I), government (G), and net exports (X – M).

Expenditure reducing policies are measures designed to cut a current account deficit by lowering disposable income to limit aggregate demand and import expenditure in particular. (HL only)

Expenditure switching policies are measures intended to cut a current account deficit by encouraging households and firms to buy domestically produced goods and services rather than imported alternatives. (HL only)

Explicit costs are the identifiable and therefore accountable costs related to the output of a product. Examples include wages, raw material costs, utility bills and rent.

Export competitiveness refers to the ability of domestic firms to export products successfully to foreign markets owing to their ability to compete in overseas markets.

Export promotion is an outward-looking trade strategy to achieve economic growth and development that focuses on greater international trade.

Export subsidies are forms of financial assistance that are targeted at protecting specific export-orientated firms.

The **external balance** refers to value of a country's export earnings being roughly equal to the value of its import expenditure.

External economies of scale are lower average costs brought about by an increase in the long-run scale of production of the industry.

Externalities (or **spillover effects**) are the external costs or benefits of an economic transaction, causing the market to fail to achieve the social optimum level of production or consumption.

Factor endowment refers to the quantity and quality of factors of production available in a country, such as natural resources and human capital. Countries well endowed in natural resources tend to have a comparative advantage in products using such resources.

Factors of production are the four categories of resources that are required to produce any good or service, namely land, labour, capital and enterprise.

The **female participation rate** measures the proportion of women who are economically active in the labour force.

The **financial account** is a record of the transactions that relate to the change in ownership of assets, that is, cross-border investments. These include foreign direct investment, portfolio investment, reserve assets and official borrowing.

Finite resources are part of the central problem of economics referring to factors of production being limited in supply.

Fiscal policy is the use of taxation and government expenditure policies to influence the level of economic activity and macroeconomic objectives.

Fixed costs, such as insurance premiums and advertising costs, are expenses that do not change with the level of output.

A **fixed exchange rate system** exists when the central bank (or monetary authority) buys and sells foreign currencies to ensure the value of its currency stays at a single, predetermined rate.

A **floating exchange rate** is where the value of a currency is determined by the demand for and supply of the currency in the foreign exchange market.

Foreign aid is financial assistance from donor countries to ELDCs for the purpose of economic development. Aid is given on concessional and non-commercial terms, rather than for financial gain.

Foreign currency reserves (or **reserve assets**) are stocks of foreign currencies held by a central bank, usually to influence the value of its currency.

Foreign direct investment (FDI) is spending by multinational corporations (MNCs) abroad.

Formal collusion is where there is an agreement between firms to reduce competition in the market.

Formal sector employment refers to officially recorded employment, with workers paying income taxes and contributing to the country's official GDP.

Free goods are resources or products that are unlimited in supply, so their output has no opportunity cost. Examples include desert sand, air, rainwater and seawater.

The **free market economy** is an economic system that relies on the market forces of demand and supply to allocate scarce resources in the economy.

The **free rider problem** occurs when people have access to (use or benefit from) a good or service without having to pay for it. As a result, the good or service will be under-provided or not provided at all in a free market.

Free trade occurs when firms can sell their goods and services in foreign markets without government intervention.

A **free trade area** (FTA) is a type of trading bloc between member states that agree to trade freely with each other but can impose separate trade restrictions with non-member countries.

Frictional unemployment (or **transitional unemployment**) is a type of unemployment that exists when people are temporarily unemployed as they actively search for a new job or are waiting to start a new job.

The **full employment level of output** exists at the point where unemployment is at its natural rate, so there will be no unemployment due to a lack of aggregate demand.

Game theory is an economic model that attempts to explain the nature of strategic interdependence in oligopolistic markets by considering the actions of competitors when making a decision, based on probable outcomes.

A **GDP deflator** (or **price deflator**) is used to convert GDP (or GNI) at current prices to GDP (or GNI) at constant prices by removing the impact of inflation on the value of national income.

GDP/GNI per capita (or **GDP per head**) is the average income in an economy, i.e. the total GDP/GNP divided by the population size.

The **gender gap** refers to the difference between women and men in terms of social, political, intellectual, cultural and economic activities, attitudes and opportunities.

The **Gender Inequality Index** (GII) is a composite measure of development by calculating gender disparities through three dimensions, namely reproductive health, empowerment and labour market participation.

The **Gini coefficient** is a statistical tool that measures income or wealth inequality, with the outcome ranging from 0 (complete equality) to 1 (total inequality).

Globalization is a multidimensional process of greater interaction and integration of goods, services, cultures, languages, social behaviours, business operations and governments worldwide.

Goods are physical (tangible) items made in the production process. Examples include furniture, cars, toothpaste and stationery.

The **government budget** is a country's planned (forecasted) revenues and its planned expenditures over a given period of time (usually one year).

Government debt (or **national debt**) refers to the sum of all accumulated government budget deficits from previous years, that is, it is the total amount of money owed by the government to its domestic and foreign creditors.

Government failure arises when the cost of attempting to prevent or correct free market imperfections or distortions turn out to be greater than the social costs of the original market failure itself.

Green GDP adjusts the country's GDP to take account of the value of environmental degradation, damage and destruction associated with economic growth.

Gross domestic product (GDP) measures the total monetary value of all goods and services produced within a country, per time period (usually one year). It measures GDP in terms of current prices, that is, the nominal prices at the time of measurement, instead of taking any changes in the price level into account.

Gross national income (GNI) is a measure of the total value of the country's income, generated from its factors of production, regardless of their location. Hence, GNI = GDP plus income from abroad minus income sent abroad.

The **Happiness Index** is an alternative to national income as a measure of well-being, by considering how information technology, governance and social norms influence communities and their level of well-being.

The **Happy Planet Index** (HPI) is a measure of sustainable human well-being, that is, how individuals and countries are able to achieve long, happy and sustainable lives.

Health indicators are a category of single indicators of measuring economic development by using health-related determinants of the quality of life, such as life expectancy and the under-five mortality rate.

A **heavily indebted poor country** (HIPC) is a low-income nation with a huge outstanding debt, making it eligible for special financial assistance from the IMF and the World Bank.

Herfindahl Index is a measure of market concentration that gives greater weighting to the market power of larger firms by squaring the value of their market share.

Hidden unemployment (or **disguised unemployment**) refers to people who are without a job, but calculations of unemployment and reporting of official unemployment statistics do not include these people, such as discouraged workers.

Homogenous products are goods and services that are identical to each other (even in consumer experience).

Human capital is the accumulation of skills, knowledge and experience of the workforce, viewed in terms of their value or cost.

The **Human Development Index** (HDI) is a composite indicator as a measure of economic development, comprising real income, life expectancy and educational attainment.

Humanitarian aid (or **development aid**) is foreign aid used to help recipient countries to achieve their economic development objectives and to improve their overall standards of living.

Hyperinflation (or **runaway inflation**) refers to extortionately high and uncontrollable rates of inflation that causes serious macroeconomic problems.

A **hypothesis** is an assumption, notion or educated guess made before research has been conducted.

Imperfect competition occurs when firms compete with each other on price and product. This is a level of competition that exists between perfect competition and monopoly.

Imperfect information occurs when people have inaccurate, incomplete or unreliable information, so decision-making is not optimal.

Implicit costs are the opportunity costs of the output, i.e. the income from the best alternative that is foregone.

Import substitution is an inward-looking strategy of economic growth and development that encourages domestic production and the purchase of domestic output through protectionist policies such as tariffs and quotas.

Incentive-related policies are a form of market-based supply-side policies that create greater incentives to work, namely by cutting personal income tax and business taxes.

The **incentives function** is an aspect of the price mechanism important in allocating resources. As price changes, the mechanism provides an incentive for producers and consumers to change their behaviour in order to maximize their benefits.

Income is a flow concept referring to the money a person receives from the production process; for example, wages and salaries.

The **income effect** is one of three factors accounting for the downward sloping demand curve. As the price of a product falls, consumers' real incomes increase so they are able to buy more goods and services at lower prices.

Income elastic demand occurs when the percentage change in the quantity demanded of a product is greater than the percentage change in consumers' real income, that is, PED > 1. This tends apply to the demand for services and luxury products.

Income elasticity of demand (YED) measures the degree of responsiveness of quantity demanded following a change in the real income of consumers.

Income inelastic demand occurs when the percentage change in the quantity demanded of a product is less than the percentage change in consumers' real income, that is, PED < 1. This applies to the demand for necessities.

Income inequality refers to income being unevenly distributed in a community or country.

The **income method** measures economic activity by calculating the value of all factor incomes earned in the economy, that is, it is the sum of wages, rent, interest and profits.

An **increase in demand** refers to a rightwards shift of the entire demand curve for a product, caused by favourable changes in non-price factors that affect demand.

Indebtedness is the condition of owing money. It is the overall debts owed to creditors by all individuals, businesses and the government of the country.

An **indirect tax** is a payment taken indirectly from the consumer's income, through their expenditure on goods and services.

Industrial policies are interventionist supply-side policies aimed at encouraging the development and growth of the manufacturing sector of an economy in order to promote growth and employment.

The **Inequality-adjusted Human Development Index** (IHDI) is a composite measure of the average level of human development by accounting for inequalities in society.

Inequality of opportunity means that some individuals and societies have less or no access to employment and consumption opportunities so are deprived of economic prosperity. It is a major cause of poverty and economic inequality.

Infant industries are often new or unestablished industries that have the potential to achieve a comparative advantage but are too underdeveloped to compete with foreign firms.

Inferior goods are products with a negative income elasticity of demand. In other words, the demand for such products tends to fall when consumers' real incomes rise.

Inflation is the sustained (continual) increase in the general (average) price level, over time. It is commonly measured by using a consumer price index (CPI).

An **inflationary gap** (or **positive output gap**) exists when an economy's actual real GDP exceeds the potential output at the full employment level of national output.

Inflation targeting refers to the practice of central banks that aim for a specific rate of inflation and use monetary policy to achieve this goal.

The **informal economy** (or **parallel markets**) can be defined as the various economic activities that are not officially recorded as part of a country's gross domestic product (GDP).

The **informal sector** of the economy refers to various economic activities that are not officially recorded as part of a country's Gross National Product (GNP) or Gross Domestic Product (GDP).

Infrastructure is the physical and organizational structures and facilities of an economy necessary for its operations, such as transportation networks, telecommunications networks, power supplies and waste disposal systems.

Injections (J) put money into the circular flow of the income. They comprise government spending (G), export earnings (X) and investment expenditure (I).

The **institutional framework** refers to the established systems, structures and contexts that shape economic behaviour in a country, such as cultural norms, the financial framework and the country's legal system.

The **interest rate** is the price of money. Expressed as a percentage measure, interest rates represent the cost of borrowing money or the return for savers.

Internal economies of scale are lower average costs brought about by an increase in the long-run scale of production of the individual firm.

The International Monetary Fund (IMF) is a worldwide multilateral financial institution set up in 1944 with the aims of fostering global monetary co-operation and promoting financial stability. This helps to facilitate international trade, promote sustainable economic growth, and reduce poverty around the world.

An **international poverty line** is the minimum threshold level of income that a person must earn or have access to in order to meet the basic needs necessary for human survival.

International trade is the exchange of capital, goods and services across nation states. It involves the sale of exports (goods and services sold to overseas buyers) and imports (foreign goods and services bought by domestic households and firms).

Interventionist supply-side policies are the deliberate attempts by a government to influence aggregate supply and the productive capacity of the economy.

The **invisible hand** is Adam Smith's metaphor describing how each individual's behaviour and decision-making ends up benefiting society as a whole because they are incentivized to act in their own self-interest.

Inward foreign direct investment refers to foreign multinational companies expanding their operations in the domestic economy.

The **J-curve effect** is an economic model that shows the impact on a country's balance of trade following a currency depreciation, namely that the trade balance worsens before it gets better. (HL only)

Joint supply refers to the supply of a product that results in the output of at least one by-product.

Keynesian economics refers to an interventionist approach to macroeconomic policy by advocating increased government expenditure and lower taxes to stimulate aggregate demand.

The **Keynesian multiplier** shows that any increase in the value of injections into the circular flow of an economy results in a proportionately larger increase in aggregate demand.

The **labour force** refers to people of working age who are in employment (including the self-employed) plus those who are seeking work, that is, those who are available for work at current real wage rates.

Labour market policies (or **labour market reforms**) are government policies designed to create greater flexibility and efficiency in the labour market.

A **labour union** (or **trade union**) is an organization that aims to protect the interests of its worker members, that is, the terms and conditions of employment, such as workers' pay and benefits.

Laissez-faire is the approach to economics that allows transactions to take place without government interference. It is a French term that loosely translates as 'let it be'.

Land rights refer to the ability of individuals to obtain, use and hold land at their will.

The **law of demand** refers to the inverse relationship between the price of a good or service and its quantity demanded, that is, the quantity demanded falls as price rises, while the quantity demanded rises if prices fall.

The **law of diminishing marginal returns** states that by employing additional variable factors of production, the marginal returns will eventually decline. This is due to at least one of the factors of production being fixed in the short run.

The **law of diminishing marginal utility** states that as individuals consume more of a product, the satisfaction gained from each additional unit of consumption declines. Hence, customers would only purchase more units at lower prices.

The **law of supply** is a general rule in economics stating there is a positive relationship between the quantity supplied of a product and its price, ceteris paribus.

Leakages or **withdrawals** (W) take money out of the circular flow of the income. They comprise savings (S), taxation (T) and imports (M), that is, $W = S + T + M$.

Legislation (in the context of asymmetric information) refers to laws stipulated by the government as a response to tackling imperfect information in markets, such as requiring health warnings on cigarette packets.

Limit pricing is a strategy where an incumbent firm adopts a price that is below the estimated costs of a potential new entrant.

The **liquidity trap** exists when the minimum lending rate cannot be cut any further, as it is close to or at zero, thereby making traditional expansionary monetary policy ineffective.

Logic refers to rationality and reasoning, rather than emotions or beliefs, in explaining economic phenomena and policymaking.

The **long run** is a period of time when no factors of production are fixed, hence, all factor inputs are variable.

The **long run aggregate supply** (LRAS) curve, according to the new classical model, is the maximum level of real GDP at the full employment level of national output. It shows that aggregate supply is independent of the price level, that is, LRAS is perfectly price inelastic.

The **long run average cost** (LRAC) represents the cost per unit at different levels of output when all factors of production are variable. It is equal to the average fixed cost (AFC) plus the average variable cost (AVC).

The **long run Phillips curve** (LRPC) shows that in the long run, there is only a single rate of unemployment (the natural rate) that is consistent with a stable inflation rate. Hence, the LRPC is vertical, that is, there is no trade-off between low inflation and low unemployment.

The **Lorenz curve** is a graphical representation of income distribution in a country, based on the income or wealth accounted for by each decile or quintile of the population.

A **loss** occurs when a firm experiences negative economic profit, that is, its total costs of production exceed its total revenue, that is, TC > TR.

Luxuries are superior goods and services with a YED value of greater than one, that is, their demand is highly income elastic. An increase in real consumer incomes leads to a proportionally greater increase in the demand for luxuries.

Macroeconomic equilibrium exists when aggregate demand is equal to aggregate supply in the economy, that is, AD = AS.

Macroeconomics is the area of economics focused on the operations of the economy as a whole, rather than the behaviour of economic agents in individual markets.

Malign deflation is generally harmful deflation due to a decline in aggregate demand for goods and services in the economy (shown by an inwards shift of the AD curve, and hence, a fall in real GDP).

A **managed exchange rate** is a system where the government or central monetary authority intervenes periodically in the foreign exchange market to influence the exchange rate, when deemed necessary to maintain certainty and confidence in the economy.

Marginal cost (MC) is the cost of producing an extra unit of output. It is calculated by dividing the change in total costs (ΔTC) by the change in the level of output (ΔQ), that is, $MC = \Delta TC/\Delta Q$.

Marginal private benefit (MPB) is the additional value enjoyed by households and firms from consuming or producing an extra unit of a good or service.

Marginal private cost (MPC) is the additional expense of production for firms or the extra amount paid by customers for producing or consuming an extra unit of a good or service.

The **marginal rate of transformation** (or the **rate of substitution**) is the gradient of the PCC, showing the opportunity cost between the two products in question.

Marginal revenue (MR) is the extra revenue received from the sale of an extra unit of output, that is, $MR = \Delta TR \div \Delta Q$.

Marginal social benefit (MSB) is the total gains to society from an extra unit of production or consumption of a particular product, that is, the sum of marginal private benefits and external benefits to others in society.

Marginal social cost (MSC) is the total expenses to society from an extra unit of production or consumption of a particular product. It includes both private costs and the external costs to others in society.

The **marginal tax rate / marginal rate of tax** (MTR) refers to the percentage of direct tax paid on the last dollar of an individual's income, i.e. the change in tax rate paid from a given change in income or $\Delta T \div \Delta Y$.

Marginal utility refers to the benefit or satisfaction gained from consuming an additional unit of a good or service.

A **market** is any place where transactions take place between buyers and sellers. For example, shares are traded in a stock market and currencies are traded on the foreign exchange market.

Market access refers to the ability of a country to enter foreign markets to sell its goods and services.

Market-based supply-side policies focus on freeing up markets and improving market incentives in order to increase aggregate supply, thereby improving incentives to increase investments and productivity and incentives to work.

Market concentration measures the extent to which sales revenue in an industry is dominated by one or more of the largest firms.

Market demand refers to the sum of all individual demand for a particular good or service at each given price level. It is found by adding up all individual demand at separate price levels.

The **market demand curve** is the sum of all individual demand for a product, that is, the aggregate of individual demand at each price level.

Market disequilibrium occurs when the quantity demanded for a product is either higher or lower than the quantity supplied in the market, that is, there is either a shortage (excess demand) or surplus (excess supply).

Market equilibrium occurs when the quantity demanded for a product is equal to the quantity supplied of the product, that is, there are no shortages or surpluses.

Market failure refers to any situation when the price mechanism (the free market forces of demand and supply) allocates scarce resources in an inefficient way.

Market-oriented policies are dynamic outward-looking macroeconomic policies used to stimulate economic growth and development via market forces, e.g. using anti-monopoly regulation to encourage competition and efficiency.

Market power refers to the ability of a firm to manipulate the price of a product, usually above the perfectly competitive level.

Market share refers to a firm's portion of the total value of sales revenue in a particular industry. It is calculated by using the formula: (Firm's total sales revenue / Industry's total sales revenue) × 100.

Market structure refers to the categorization of firms in a particular industry, based on their level of market power; for example, the number and size of firms in the industry, the nature of barriers to entry, and the degree or intensity of competition.

Market supply refers to the sum of all individual supply of a product at each price level. It is found by adding up all individual supply of producers at each price level.

The **Marshall-Lerner condition** states that a devaluation or depreciation of the currency will work in rectifying a balance of trade deficit only if the sum of the price elasticity of demand for exports and imports is greater than 1 (that is, price elastic). (HL only)

Marxism is an approach to macroeconomic policy that focuses on meeting the needs and values of the masses, rather than for the privilege of a minority of capitalists.

Menu costs are a cost of inflation because with prices continually rising, firms incur ongoing costs of changing and updating price lists and price tags.

Mergers occur when two firms voluntarily decide to combine into one, larger, firm.

Merit goods are products that create positive externalities when they are produced or consumed, that is, MSB > MPB.

Micro-credit refers to loans of exceedingly small amounts to individuals on low incomes for self-employment projects that generate income so they can become self-reliant.

Microeconomics is the area of economics focused on the actions of individuals and firms in particular markets rather than the operations of the economy as a whole.

Microfinance refers to small sums of borrowed funds by individuals in ELDCs for self-employment purposes so that they can generate income for themselves and their families.

The **Millennium Development Goals** (MDGs) of the United Nations committed member countries to combat poverty, hunger, disease, illiteracy, environmental degradation and discrimination against women. The MDGs were scheduled to be achieved by 2015.

The **minimum efficient scale** occurs at the point when all economies of scale have been exploited, that is, shown at the lowest point of the LRAC curve.

The **Minimum Income Standard** (MIS) measures the lowest amount of income needed for what members of the public in the country think is acceptable to be able to live in a socially appropriate way.

The **minimum lending rate** (MLR) is the official rate of interest charged by the central bank on loans to commercial banks. The central bank will not lend money below this rate.

The **minimum reserve ratio** (MRR) is the lowest amount that commercial banks are required to keep as reserves in the central bank thereby limiting how much they can lend and how much credit they can create.

A **minimum wage** is the lowest rate of pay that firms must pay to their employees per time period, as stipulated by the government.

A **mixed economy** is an economic system that features aspects of both a planned and market economic system, with some resources being owned and controlled by private individuals and firms with others being owned and controlled by the government.

Mobile banking is a service provided by a bank or other financial institution that allows its customers to conduct financial transactions remotely using a mobile device, such as a smartphone, tablet computer or laptop.

The **mobility of factors of production** refers to the level of ease and cost of factor substitution in the production process. The easier it is to substitute factors of production, the more price elastic supply tends to be.

A **model** is a hypothesis that has been repeatedly tested and proven or rejected and can be used to explain the real world.

Monetarists believe that monetary policy is the most powerful and effective macroeconomic stabilization policy to influence the overall level of economic activity.

Monetary policy is the use of interest rates and the money supply to influence the level of aggregate demand and economic activity, in order to achieve the country's macroeconomic objectives.

Monetary union refers to the monetary system in a common market that requires the convergence of monetary policy that is governed by a common central bank.

The **money multiplier** is a formula used to calculate by how much an initial deposit (savings at a bank) increases the money supply. The formula is: Money multiplier = 1 ÷ Reserve ratio.

Money supply is the amount of money in circulation within the economy at a particular point in time, as determined by the central bank. The money supply comprises legal tender (banknotes and coins), loans, credit, bank deposits and central bank reserves.

A **monopoly** is a market structure where there is a single supplier of a particular good or service, thus having the power to influence the market supply and price.

A **moral hazard** is a situation where a party protected from risk (due to having superior information) behaves differently than if they were fully exposed to the risk. This potentially imposes costs on the party that has inferior information.

Movement along a demand curve is caused by price changes only. A fall in price causes quantity demanded to expand while an increase in price causes quantity demanded to contract.

Movements along a supply curve occur when the price of the product changes, causing a change in the quantity supplied, rather than a shift in the supply curve.

The **Multidimensional Poverty Index** (MPI) is a composite measure of poverty that identifies numerous deficiencies of individuals and households based on three dimensions or indicators of poverty – health, education and standards of living.

Multilateral development assistance (MDA) is financial support delivered through international institutions such as the World Bank and the International Monetary Fund (IMF).

A **multilateral trade agreement** is a legally binding preferential trade agreement between more than two countries and/or trade blocs, under the guidelines of the WTO.

Multinational corporations (MNCs) are businesses that operate in two or more countries, i.e. they have operations outside of their home country. For example, Japan's Toyota, Nissan and Honda all have production plants in the UK and USA.

National income can be defined as the value of all the goods and services produced in a country during a year, thereby indicating the level of economic activity in the country.

Nationalization is the purchase of privately owned assets or industries by the government. This usually occurs when the government takes control of an industry previously in the private sector in order to run it in the best interest of the public.

The **National Minimum Wage** (NMW) is a price floor in the labour market, requiring employers to pay at least this amount to their employees. They are not permitted to pay less than this.

The **national poverty line** is the level of income below which a person in the country cannot meet their basic minimum needs in terms of shelter, nutritional level and clothing.

Natural monopoly occurs when only one firm can operate in a market profitably owing to natural barriers to entry and economies of scale.

The **natural rate of unemployment** (NRU) is the level of unemployment at the full employment equilibrium level of national output, comprising frictional, seasonal and structural unemployment.

Necessities are essential products that consumers buy in order to meet their needs. They have a positive PED value that is less than 1, that is, the proportionate change in quantity demanded is less than the proportionate change in real incomes.

Needs are the goods and services deemed to be essential for human survival. Examples include nutritional food, clean water, clothing, shelter (housing), and access to healthcare and education.

Negative economic growth occurs when the level of economic activity declines, that is, there is a fall in real GDP. This is associated with a recession in the business cycle.

Negative economic profit occurs when a firm makes a loss, that is, its total costs of production exceed its total revenue.

Negative externalities (or **external costs**) are expenses incurred by third parties in an economic transaction for which no compensation is paid.

A **negative multiplier effect** occurs when an initial leakage from the circular flow of income leads to a greater than proportionate fall in final real GDP.

Negative real interest rates occur when the inflation rate exceeds the nominal interest rate, meaning the nominal returns on savings are completely diminished by the effects of higher prices in the economy.

Net exports, the value of exports minus the value of imports, are part of the aggregate demand equation (and GDP).

Net factor income from abroad is the net income gained from other nations, that is, the difference between a country's income earned from abroad and the income earned by foreign individuals and firms in the domestic economy.

The **new classical counter revolution** (NCCR) favoured supply-side macroeconomic policies over interventionist arguments, such as privatization of state-owned enterprises and deregulation of markets.

Nominal economic growth is the annual rate of change in the monetary value of GDP. This is expressed in terms of current prices, that is, the money value of GDP at the current point in time.

The **nominal interest rate** is the actual rate that is agreed between a bank and the customer, that is, it is the rate borrowers pay on their loans or the return savers receive on their cash deposits.

Nominal GDP measures national output using current market prices, that is, the value of GDP at the time of measurement. It measures the monetary value (or face value) of all goods and services produced within a country for a given period of time, usually a year.

Non-collusive oligopoly exists where large firms in the industry act strategically by competing independently, taking into account the likely or possible actions of rival firms.

Non-discrimination means a WTO member country cannot discriminate between its trading partners in terms of their goods, services or people.

Non-excludability is one of the two characteristics of pure public goods, where it is not possible to exclude someone from using a good or service even if they have not paid for it, because it is not practical to charge a price.

Non-price competition refers to all forms of competition that are not related to price such as advertising, branding, consumer experience and product.

Non-price determinants of demand are the various factors other than the price of the good or service that affect the demand for the product.

Non-price determinants of supply are the various factors other than the price of a good or service that affect the supply of the product.

Non-renewable resources are natural resources that cannot be readily replaced by natural means at a quick enough pace to keep up with the level of consumption.

Non-rivalrous is one of the two characteristics of pure public goods, where the consumption of the good or service by one person does not reduce the amount available for others.

Non-tariff barriers are restrictions to international trade that do not involve a tax or duty (tariff), such as quotas, export subsidies and administrative barriers.

Normal goods are products that customers tend to buy more of as their real income level increases. This includes both necessities and luxury goods and services.

Normal profit exists when a firm earns the same amount of profit as it would from its next best alternative option. The firm would therefore remain in its current market as it would not gain more profit from operating in any other market.

Normative economics considers people's varying opinions and beliefs about what should be (or what ought to be). Such claims are subjective and expressed as value judgements.

Nudge theory is the practice of influencing the choices that people make. Nudges are created by choice architects using small prompts or tweaks to alter social and economic behaviour, but without taking away the power for people to choose.

Official borrowing occurs when governments borrow from creditors or institutions outside of their economy. Examples of this can be loans from the IMF for development purposes, or borrowing from foreign banks and financial institutions.

Official Development Assistance (ODA) is foreign aid from donor governments, rather than from NGOs or non-profit organizations.

Oligopoly is a market structure in which a few large firms dominate the industry, with each firm having a high degree of market power.

The **open economy** is part of the circular flow of income model comprising domestic and foreign economic decision makers, that is, households, firms, the government and the foreign sector (which accounts for exports and imports).

Open market operations (OMOs) refer to the buying and selling of government securities by a country's central bank.

Opportunistic behaviour means that one party can take advantage of the opportunity that the other party lacks information, resulting in adverse selection or moral hazard.

Opportunity cost refers to costs of an economic decision measured in terms of the best alternative choice forgone. Owing to scarcity, there is always an opportunity cost when making an economic decision (with the exception of *free goods*).

The **output method** measures economic activity by calculating the actual value of all final (finished) goods and services produced within the economy each year. This method of measuring the level of economic activity is also referred to as **national output** (O).

Outward foreign direct investment refers to multinational companies from the domestic economy expanding their operations in overseas markets.

An **overvalued currency** occurs when the value of a currency is above its equilibrium value in the long run.

Pareto efficiency can be seen on a PPC and occurs when it is impossible to reallocate resources in such a way to make one party better off without making someone else worse off.

The **peak** of the business cycle occurs when economic activity is at its highest level.

A **pegged exchange rate** is a form of managed exchange rate that usually refers to the actions of a central bank to keep the value of its currency within a particular range against another currency.

Perfect competition is a market structure where there is intensive competition, with no individual firm being large enough to have any market power to influence the price or quantity traded.

Perfect information refers to equal and easy access of information about alternative products available on the market, such as prices and product specifications. Rationality assumes that people use all the available information to make their decision.

Perfectly price elastic demand is an extreme situation of PED, where an increase in price will result in a firm losing all of its quantity demanded (until it reaches zero). A reduction in price will mean that all of the demand in the market will go to this firm. Quantity demanded will change by an infinite amount (until the firm has all of the market demand). Therefore, the value of PED = ∞.

Perfectly price elastic supply refers to a theoretical situation when the value of PES = infinity, that is, firms can supply any amount at the given price level due to constant unit costs of production and no limits to productive capacity.

Perfectly price inelastic demand is also an extreme situation of PED, where a given percentage change in price has no impact whatsoever on the quantity demanded. In this case, the value of PED = 0.

Perfectly price inelastic supply refers to a theoretical situation when the value of PES = 0, that is, an increase in price has no impact on the quantity supplied as there is no spare capacity to raise output.

Personal income tax is a direct tax that governments impose on income generated by an individual within their legal jurisdiction.

A **Pigouvian tax** is a government levy on the sale of certain goods and services used in response to negative externalities and CPRs by taxing buyers and sellers of the product so they pay for the negative externalities of production and consumption.

A **planned economy** is an economic system where the government (or public sector) allocates scarce resources.

Portfolio investment is the purchase of financial investments abroad, such as the purchase of stocks, shares and bonds of overseas firms and governments. It is part of the financial account that refers to the stock of investment assets, such as equities (stocks and shares) and government debt (securities, bonds and treasury bills).

Positive economics is the study of economics that is provable, that is, factual statements about the economy or statements of 'what is' rather than 'what ought to be'.

Positive externalities (or **external benefits**) are the benefits enjoyed by a third party not directly involved in an economic transaction.

Potential output refers to the possible level of real GDP of an economy, as shown on its production possibility curve, if all resources are used efficiently. It is a long run phenomenon represented by an outwards shift of the PPC.

Poverty is the condition of an individual, household, community or country being extremely poor, that is, not being able to meet their basic human needs.

The **poverty trap** (or **poverty cycle**) is a vicious cycle of poverty and deprivation causing even greater poverty, from one generation to the next.

A **preferential trade agreement** (PTA) is a trade treaty between two or more countries, giving special or favourable terms and conditions of trade to member countries.

A **price ceiling** (also known as a **maximum price**) occurs when the government sets a price *below* the market equilibrium price to encourage output and consumption.

Price competition is the use of pricing strategies to compete in an industry.

Price controls are government regulations establishing a maximum or minimum price to be charged for certain goods and services.

Price elastic demand describes demand for a product that is relatively responsive to changes in the product's price, usually due to close substitutes being widely available. In this case, the value of PED > 1.

Price elastic supply occurs when firms can quite easily increase the quantity supplied if there is an increase in the price of the product. As PES > 1, the change in price leads to a greater percentage change in quantity supplied.

Price elasticity of demand (PED) measures the degree of responsiveness of quantity demanded for a product following a change in its own price. The formula to calculate this is PED = $\%\Delta Q_D \div \%\Delta P$.

Price elasticity of supply (PES) is a measure of how responsive the supply of a product is to a change in the price of the product. It is calculated using the formula PES = $\%\Delta Qs \div \%\Delta P$.

A **price floor** (also known as a **minimum price**) is the imposition of a price guarantee set *above* the market equilibrium price to encourage supply of a certain good or service.

Price inelastic demand describes demand for a product that is relatively unresponsive to changes in the product's price, mainly due to the lack of close substitutes being available. In this case, the value of PED < 1.

Price inelastic supply occurs when supply is relatively fixed in the short run, so firms cannot easily increase quantity supplied in a short period of time following a rise in the market price. This means the value of PES < 1.

A **price maker** (or **price setter**) is a single or dominant firm that has significant market power enabling it to manipulate its prices as it has significant control over market supply.

The **price mechanism** refers the interactions between buyers and sellers (the forces of demand and supply) in order to allocate resources, thereby determining production and consumption choices.

Price rigidity refers to the tendency of prices to remain unchanged in non-collusive oligopoly.

Price stability occurs when the general price level remains largely constant due to a low and stable rate of inflation in the economy.

Price takers are firms that have no market power and are therefore unable to influence their price.

The **primary sector** of the economy involves industries in the primary stage of production, such as the extraction and collection of natural resources. Examples include agricultural farming, forestry, hunting, fishing, oil exploration and mining.

Private benefits are the gains of production and consumption enjoyed by an individual firm or person.

Private costs of production and consumption are the expenses incurred by an individual firm or person.

Private goods are both excludable (those who do not pay cannot benefit from the product) and rivalrous (one person's consumption of the product reduces the amount available to others).

The **private sector** is the sector of the economy where private firms and individuals produce goods and services.

Privatization is the transfer of the ownership of government assets from the public sector to owners (shareholders) in the private sector. It is a form of supply-side policy that involves the sale or transfer of state-owned assets to the private sector, in order to improve competition, efficiency and productivity.

Producer surplus refers to the gain or benefit to firms who receive a price that is higher than what they are willing and able to supply.

The **production possibility curve** (PPC) or **production possibility frontier** (PPF) is a diagrammatic representation of the maximum combination of two products that a country can produce, given the efficient use of all its resources, at any moment in time.

Productive efficiency occurs at the output level where average costs are minimized, i.e. where MC = AC. Hence, there is no wastage or unemployment of scarce resources.

Profit is the positive difference between a firm's total revenues (TC) and its total costs (TC). It is calculated using the formula: TR – TC.

Profit maximization is assumed to be the key goal of firms operating in the private sector. The largest positive difference between total revenue and total costs occurs at the output level where MC = MR.

A **progressive tax** charges a higher *percentage* tax as an individual's income rises, that is, those who earn more pay a greater proportion of their income in tax.

Property rights are the entitlement to both tangible and intangible assets owned by an individual, organization or government, such as land rights and intellectual property rights.

A **proportional tax** (or a **flat rate tax**) charges the same percentage rate of tax, irrespective of the taxpayer's level of income, wealth or profits.

Provision of information means the government provides additional information about goods and services, or requires firms to do so, in order to help buyers make more informed decisions.

Public goods are collective consumption goods provided by the government that are non-rivalrous and non-excludable.

The **public sector** is the sector of the economy where the government produces or supplies certain goods and services.

Purchasing power parity (PPP) refers to the exchange rate that enables residents to purchase a common basket of goods and services in different countries.

Quantitative easing (QE) is a monetary policy tool that injects money directly into the economy via the central bank purchasing bonds (government-backed securities with a promise to pay later).

Quantity demanded refers to the amount of a good or service that customers are willing and able to purchase at each price level. The higher the price, the lower the quantity demanded, ceteris paribus.

Quantity supplied is the amount of a good or service that firms are willing to supply at a given price over a period of time.

Quotas are quantitative limits on the importation of a good into a country.

Rational consumer choice refers to the decision-making process based on the assumption that people make choices that result in the optimal level of benefits (utility). It is the basis of most introductory and mainstream economic theories.

The **rationing function** is an aspect of the price mechanism that serves to limit or preserve resources, that is, higher prices lower the quantity demanded thereby helping to preserve (ration) the good or service.

Real economic growth adjusts nominal economic growth to take account of inflation (that is, changes in the general price level as measured by changes in consumer prices).

Real GDP per capita measures the value of the annual GDP per person, that is, the real gross domestic product of a country divided by its population size.

Real gross domestic product (**real GDP**) measures the value of national output adjusted for inflation to reflect the true value of goods and services produced in a given year.

The **real interest rate** accounts for the impact of inflation on the return to savers and the cost of debts to borrowers. The formula is: Real interest rate = Nominal interest rate – Inflation rate.

A **recession** is a phase in the business cycle that occurs when there is a fall in gross domestic product (GDP) for two consecutive quarters.

A **recovery** occurs when GDP starts to rise after the trough in the business cycle, eventually leading to economic growth.

Refutation is the act of a statement or theory being proved to be wrong or false by the empirical evidence.

A **regressive tax** occurs when those on higher incomes pay less tax as a proportion of their income.

Regulation (in the context of asymmetric information) refers to the act of monitoring and controlling the activity of firms, such as advertising information required to be legal, decent, honest and truthful.

Relative poverty in a country refers to individuals and households who are unable to earn the minimum amount of income needed for them to maintain the average standard of living in the community or country.

Remittances refer to the movement of money when nationals working abroad send money back to their home country.

Renewable resources are factors of production that can regenerate and replenish, such as water, air, solar energy and cultivated plants.

Research and development (R&D) is the process of business activities to improve, introduce and innovate products, processes and procedures.

Reserve assets are currencies or other assets, such as gold, that are readily transferable for international transactions and balancing international payments. They are recorded in the financial account of the balance of payments.

Retaliation (or **trade retaliation**) refers to the actions taken by a country in response to trade restrictions being imposed on it from other countries.

Revaluation occurs when the price of a currency operating in a fixed exchange rate system is officially and deliberately increased.

Revenue is the money received from the sale of a firm's output of goods and services.

Revenue maximization occurs when a firm produces at an output where marginal revenue is equal to zero (MR = 0).

Rule of thumb is a type of bias referring to making decisions in a practical and approximate way without having to be exact. It is about judging a situation based on experience and sticking to a default choice.

Sales revenue is the sum of money received from the sale of a good or service. It is calculated by the formula: P × Q, where P = price and Q = quantity traded.

Satisficing is a business objective that aims for a satisfactory or adequate level or profit. This is because profit maximization is likely to incur too many sacrifices.

The **savings ratio** is the amount of money people keep aside for future use expressed as a percentage of total disposable income.

Say's law states that the ability to purchase a product depends on the ability to produce or supply, thereby generating income. Essentially, this means that supply can create its own demand.

Scarcity refers the finite resources (limited in supply) of an economy relative to the unlimited needs and wants of individuals and societies.

Screening is a strategy undertaken by parties with less or inferior information to tackle the problem of adverse selection, with the aim of maximizing their own return or utility.

Seasonal unemployment is a type or cause of unemployment due to fluctuations and periodic changes in the level of demand for a certain good or service during different times of the year.

Services are intangible (non-physical) products such as haircuts, bus journeys, telephone calls, and internet access.

A **shift of the demand curve** occurs when there is a change in any non-price factor that affects the demand for a product. The entire demand curve shifts inwards if there are unfavourable changes affecting demand, and vice versa.

Shifts of the supply curve occur when there is a change in non-price factors that affect supply, resulting in changes in supply. This is represented by a shift of the entire supply curve of the product.

Shoe leather costs are a cost of inflation, as customers spend more time searching for the best deals, that is, it represents a greater opportunity cost for customers in terms of time and effort needed.

The **short run** is a period of time when at least one factor of production is fixed (such as capital), whereas other factor inputs are variable (such as labour).

The **short run aggregate supply** (SRAS) curve shows the total planned national output at different price levels, ceteris paribus.

The **short run average cost** {SRAC) represents the cost per unit at different levels of output when at least one factor of production is fixed, such as capital resources. It is equal to the average fixed cost (AFC) plus the short run average variable cost (SRAVC).

The **short run Phillips curve** (SRPC) shows a potential trade-off between pursuing low unemployment and low inflation as macroeconomic objectives.

Signalling is a strategy undertaken by parties with more or superior information to tackle the problem of adverse selection with the aim of maximizing their own return or utility.

The **signalling function** is an aspect of the price mechanism that signifies to producers and consumers where resources are required (in markets where prices increase) and where they are not (in market where prices fall).

A **single indicator** refers to a statistical measure of economic development that uses one particular gauge, such as literacy rates, income per capita or life expectancy.

A **slump** (or **trough**) occurs at the bottom of a recession in the business cycle when aggregate demand remains low.

Social benefits are the total benefits of consumption or production to society, that is, the sum of private benefits and external benefits.

Social costs are the total costs of consumption or production to society, that is, the sum of private costs and external costs.

A **social enterprise** (or **social business**) is defined as an organization that focuses on meeting specific social objectives.

Social science is the study of human interactions and relationships between individuals and societies.

Social surplus is the sum of producer and consumer surplus in a specific market. It is used to demonstrate efficiency in a market.

Social sustainability refers to the ability of an economy to develop social processes and structures that enable its current population to live optimally and to support the ability of future populations to do the same.

A **specific tax** (also known as a **per unit tax**) imposes a fixed amount of tax on each product that is sold; for example, taxes on cigarettes, air passenger tax, electronic road pricing and highway tolls.

Speculation occurs when a financial asset, such as the domestic currency or a foreign currency, is purchased in the hope or anticipation that the resale value will be higher.

Stagflation occurs when there is rising inflation but falling national output and rising unemployment.

The **standard of living** in an economy indicates the extent to which people have the ability to lead a life that includes, but is not limited to, personal wealth and material comfort derived from having access to goods and services.

Structural unemployment is a type or cause of unemployment that occurs when there is a decline in the demand for goods and services produced in a certain industry.

Subsidies are a form of financial assistance from the government to domestic firms by lowering their costs of production in order to help the firms compete against foreign imports. Subsidies encourage output, reduce the price of certain products, or keep down the cost of living for the domestic country's citizens.

Substitutes are products in competitive demand, that is, they can be used instead of each other, such as Coca-Cola or Pepsi and tea or coffee.

The **substitution effect** states that as the price of a product falls, more people can buy the product, so choose this over rival products, that is, it causes consumers to replace higher priced products with lower priced ones.

Supply is the amount of a good or service that firms are *willing* and *able* to provide at any particular price, per time period.

The **supply of money** refers to the total amount of money circulating in the economy at any point in time.

Supply-side policies are the long-term government strategies used to increase the productive capacity of the economy by improving the quality and/or quantity of factors of production.

A **surplus** is created when the supply of a product exceeds its demand because the price is set higher than the market equilibrium price. A surplus on an account exists when the total value of credit items exceeds the total value of debit items, over a given period of time.

Sustainable development is economic development that meets the needs of the present generations without compromising the ability of future generations to meet their own needs.

The **Sustainable Development Goals** (SDGs) of the United Nations Development Programme consist of 17 international development targets aimed at achieving peace and prosperity by all UN member countries by 2030.

Sustainability is a key concept referring to an economy achieving the economic goals that allow an increase in living standards without jeopardizing the needs and wants of future generations.

Tacit collusion occurs when two or more oligopolistic firms implicitly (or informally) agree to use restrictive trade practices, e.g. price cutting or price leadership.

Tariffs are a specific tax on imported goods and services.

A **tax** refers to a government levy on income directly or indirectly through expenditure. Therefore, taxes can be categorized as direct or indirect tax.

Tax policies are a key redistribution policy, used to redistribute income and wealth to protect and support the poorest members of society.

Taxable income is the amount of income that an individual or firm is liable to pay tax on, after taxable allowances have been accounted for.

The **tax threshold** (or **tax allowance**) refers to the amount of income that workers can earn per year before they become liable for tax.

A **theory** is a broad generalization used to explain situations or scenarios already supported by economic evidence and data from economic models.

Time lags refer to the variable delays between changes to government policy and their influence on the level of aggregate demand, national output and the general price level (rate of inflation).

Total costs (TC) are the aggregate amount of production costs spent on the output of a given good or service. It is the sum of variable and fixed costs.

Total revenue (TR) is the overall amount of money received by a firm from selling its output.

Total variable costs (TVC) are the overall production costs incurred directly from the output of a particular good or service. TVC continually rise with higher levels of output, such as labour costs and raw material costs.

Tradable permits (or **cap and trade schemes**) are government-regulated emissions trading schemes that limit pollution in an industry to a more socially efficient level. Efficient firms can sell any excess permits that they do not need to use.

Trade creation occurs when trade shifts from higher-cost producers outside of a trading bloc to lower-cost producers within the bloc, due to the removal of trade barriers.

Trade diversion occurs when trade shifts from lower-cost producers outside of a trading bloc to higher-cost producers within the bloc under the terms and conditions of the trade bloc agreements.

Trade liberalization is the reduction or removal of trade barriers such as tariffs and quotas, thereby increasing competition, productivity and efficiency in international markets.

Trade protection (or **protectionism**) is the use of barriers to trade to safeguard an economy from excessive international trade and foreign competition.

A **trade war** is an economic conflict that results from extreme measures of trade protection between two or more countries, by imposing reciprocal restrictions to international trade.

A **trading bloc** is a group of countries that agree to economic integration and freer international trade by reducing or removing trade barriers with one other.

The **tragedy of the commons** refers to the degradation, depletion or destruction of a common pool resource caused by the problems of rivalry and overuse.

A **transfer payment** is a sum of money from the government to another party for which no goods or services are being paid, that is, there is no corresponding output or economic activity for such payments.

Underemployment occurs when people are willing to work full-time but can only find part-time employment, causing a loss in potential national output.

An **undervalued currency** occurs when the value of a currency is below its equilibrium value in the long run.

Unemployment refers to the non-use of a factor of production, such as labour, that is, people who are registered as being available and willing to work at the going real wage rate but who cannot find employment.

Unemployment benefits are welfare transfer payments made by a government to support people who are willing to work but cannot find employment.

The **unemployment rate** measures the percentage of the labour force that is unemployed. It is calculated by the formula: (Total unemployed / Total labour force) × 100.

Unequal distribution of income refers to imbalances of income distributions, that is, very few members of the society account for a disproportionately high concentration of national income.

Unequal distribution of wealth, also call a wealth gap, is defined as the assets, saving and investments that are spread between the citizens of a country.

Unitary price elastic demand occurs when a given price change of a product leads to the same percentage change in the quantity demanded. Hence, the PED value = 1.

Unitary price elastic supply refers to a theoretical situation when the value of PES = 1, that is, the percentage change in quantity supplied is equal to the percentage change in the price of the product.

A **universal basic income** (UBI) refers to the guaranteed minimum level of income that a government guarantees to each individual of the country.

Utility refers to the level of consumer satisfaction from the consumption or use of a good or service.

Utility maximization refers to the assumption that people choose what gives them the largest amount of utility (the greatest level of satisfaction).

Value judgements are the beliefs of individuals and societies about what is right or wrong, or good or bad.

Voluntarily unemployed people are economically inactive individuals who choose not to pursue full-time employment.

A **wage-price spiral** is a circular cause of higher inflation due to trade unions negotiating higher wages owing to higher costs of living, which causes higher costs of firms.

Wants are human desires, that is, things people would like to have or have more of. Wants are infinite (unlimited) as it is human nature to continually want more things.

Wealth is a stock concept referring to the value of accumulated assets; for example, property, land and savings.

Wealth inequality refers to the unequal distribution of assets, savings and investments across the economy.

A **wealth tax** is direct tax based on the market value of assets owned by an individual, household or firm.

The **World Bank** is the international organization that lends money to less economically developed countries (ELDCs) for economic development projects and structural change.

The **World Trade Organization** (WTO) is a global organization that exists to promote trade liberalization, to oversee multilateral trade agreements, and to resolve trade disputes between member states.

X-inefficiency occurs when firms lack the incentive to control costs of production. It represents the difference between the behaviour of rational firms in competitive markets and those in the absence of competition.

Index

A

ability to store 117, 119
abnormal profit 190, 211
absolute advantage 426–9
absolute poverty 333
access to markets 592
accounts
  capital 511
  current 496, 508–11, 515–16, 518–20, 523–4
  financial 511–12, 516–17
  interdependence between 513–14
acquisitions 90
actual growth 22–3, 266
actual output 266
administrative barriers 452–4
ad valorem taxes 130
adverse selection 173
adverse supply-side shocks 251
aggregate demand 243–8, 285
  macroeconomic equilibrium 256–62
  see also economic growth
aggregate supply 249–62, 285
  long run aggregate supply (LRAS) curve 252, 266–7
  macroeconomic equilibrium 256–62
  short run aggregate supply (SRAS) curve 249–51
  see also economic growth
aggregate variables 4
agriculture 222–3
aid 612–14
allocative efficiency 75, 142, 195
allocative inefficiency 198–9, 203, 208
alternative hypotheses 31
altruism 81
anchoring 79–80
ancillary services 210
Annan, K. 43
anti-competitive practices 182, 213, 406
anti-dumping 458–9
anti-monopoly regulation 346, 406–7
anti-social behaviour, and unemployment 286
appreciation 486
art, value of 114
Asian Tigers 42
asymmetric information 171–7, 180
Atkinson, A.B. 341
Atkinson, T. 335
atypical households 294
austerity measures 314–15
Autarky 428
availability bias 80
average revenue 187
average tax rate (ATR) 354

B

balance of payments 459–60, 506–25
balance of trade 509
banking system 580, 620–1
barriers to entry 118, 180–2, 185
barriers to growth 564–87
barriers to trade 436
base year 290
basic economic problem 8
basic economic questions 16
behavioural economics 42–3, 79–86
benign deflation 301
Better Life Index (BLI) 239
biases 79
bilateral trade agreements 470
biodiversity loss 553–4
boom 234
borrowing 10, 33, 306, 365, 370
  government 122, 262, 312–14, 513
  see also interest rate
bounded rationality theory 80
bounded self-control 80–1
bounded selfisÚess 81
branding 49
brand loyalty 49, 88
Brexit 35
British Steel 434
Brundtland Report 529–30
budget deficit 310, 312
bureaucracy 250
business confidence 246
business cycle 4, 234, 367, 386
business development loans 122
business objectives 87
  corporate social responsibility (CSR) 87–8
  growth 89–90
  market share 88
  profit maximization 87–9
  satisficing 88–9
by-products 60

C

capacity 116–17, 119
cap and trade schemes (CATS) 155–6
capital 8–10
capital account 511
capital expenditures 383
capital flight 574–5
capital gain 50
capital transfers 511
carbon taxes 152–3
Carson, R.L. 528
cartel 200–1
central bank 364–5
  intervention 493
ceteris paribus assumption 32, 46, 56, 138, 161
change 6
child benefits 602
choice 6, 12, 22
  basic economic questions 16
  consumer 77
  default 83
  mandated 84
  nudge theory 84–6
  opportunity cost 13–15
  restricted 84
  and scarcity 7–8
choice architecture 80, 83–4
circular economy 44
circular flow of income 24–7, 218–21, 228
claimant count 277
classical economics 38–9
classical microeconomics 38–40
climate change 532, 551
closed economy 25, 218
Coca-Cola 9, 150, 183
collective self-governance 157–8
collusive oligopoly 200–3
command and control (CAC) regulation and legislation 135–6
command economy *see* planned economy
commodities, primary 102–3, 109–10, 119
common access resources (CARs) 148–51
common external tariff 472
common markets 473–4
community surplus 75–6, 142, 439–40
comparative advantage 429–36
competition policies 123, 402–4, 459, 524
competitive supply 59–60
complementary goods 50
concentration ratio 206
concept-based learning 5–7
conflicting macroeconomic objectives 316–24
congested cities 149
constant prices 231–2, 267–8
consumer
  behaviour 137
  choice 77, 213
  deferred consumption 305
  number 52
  preferences 49
  proportion of income 99
  spending 271
  surplus 74
consumer price index (CPI) 290–5
consumption 244–6, 246–7
  level of 124
  per capita consumption (PCC) 549
context 7
contracting out 169–70
coronavirus 42, 191, 248
corporate social responsibility (CSR) 87–8
corporation tax 356–7
corruption 583–5, 623
cost-push inflation 296, 321, 494
crawling peg system 500–1
credit items 506–8
credit ratings 313–14, 518
crime, and unemployment 286
crowding out 393
currency
  appreciation 486, 496
  demand and supply 489–90
  depreciation 486, 494–5
  devaluation 498, 519
  exchange rates 248, 249–50, 485–504
  foreign currency reserves 497, 512

liquidity 503
and living standards 497
overvalued 501
revaluation 498
undervalued 501–2
see also exchange rates
current account 496, 508–11, 515–16, 518–20, 523–4
current expenditures 383
current prices 230, 232
current transfers 510
customs unions 472

D

deadweight loss 159
debit items 506–8
debt 306, 518
government (national) 310–15
debt relief 614–15
debt servicing costs 313
debt to GDP ratio 311
default choices 83
deferred consumption 305
deficit on an account 508
deflation 301–7
deflationary gap 253–4, 282
deflationary spiral 305
deforestation 43, 532, 552
degree of effectiveness 161
degree of necessity 98–9
demand 32, 41, 46
aggregate 243–8
change in 94–6
contraction 53
curve 46–54, 68
decrease 54, 68
diminishing marginal utility 39, 47
excess 69–70
expansion 53
future price expectations 50
income effect 47
income elasticity of demand (YED) 104–10
increase 54, 68
law of 46, 196
management 518
non-price determinants 48–53, 68
price elasticity of demand (PED) 94–103
substitution effect 47
see also goods
demand-deficient unemployment 281–2
demand-pull inflation 296, 494
demand-side policy 365
demerit goods 130, 135, 141, 146–7, 271
depreciation 486
depression *see* economic depression
deregulation 42, 346, 404, 600–1
desertification 551–2
development
barriers to 564–87
measuring 542–60
differentiation 186, 208
digital disruption 254
direct provision 135, 160, 169–70
direct taxes 122, 356–8, 381
discouraged workers 279
discrimination 343–4, 361–2, 582–3
diseconomies of scale 211
disguised unemployment 279
disinflation 303–4
distribution of wealth 4, 123, 217–23
diversification 90, 595–6
dumping 458–9
dynamic pricing 101

E

ecological footprint 241
economic agents 2, 16, 25
economically least developed countries (ELDCs) 222–3, 461, 550, 568–87, 591
economic barriers 567–8
economic depression 40–1
economic development 542–62
economic fairness 123
economic freedom 18
economic goods 11
economic growth 234, 264–73, 299, 320–4, 385, 493, 495–6, 518, 560–1
consequences of 271–2
long-term 266–7
measuring 267–9
negative 264–5
nominal 267
rates of 265
real 267–9
short-term 266
strategies 591–630
see also gross domestic product (GDP)
economic inequality 326–62, 568–9
causes of 341–6
impact of 347–9
reducing 359–62
role of taxation 350–9
economic integration 469–83, 594–5
economic methodology 28–30
ceteris paribus assumption 32
empirical evidence 33
hypotheses 31–2
logic 30
models 31–2
refutation 33–4
theories 31–2
economic models
circular flow of income 24–7
price mechanism 31, 37
production possibility curve (PPC) 19–23
economics, as a social science 2–5, 32
economic sustainability 530–1
economic systems 17
economies of scale 88, 183, 209, 423
education 154–5, 409–10, 519, 538, 547–8, 555, 570–1, 604
effective demand 46
efficiency 6, 19–20, 22, 75, 195–6, 210, 255, 402
elasticity of demand 92–3
see also income elasticity of demand (YED); price elasticity of demand (PED)
elasticity of supply *see* price elasticity of supply (PES)
embargoes 453
empirical evidence 33
employment 271, 275, 524
full 19, 252, 257–8
structural 524
see also unemployment
empowerment 556, 622–3
endemic diseases 577–8
energy, access to 606
energy indicators 549–50
Engel, E. 106
Engel curve 106–7
enterprise 8–9
entrepreneurship 8–9
environmental indicators 550–1
environmental standards 458
see also sustainability
environmental sustainability 531–2
equality 36–7, 325–6
equilibrium 26, 67
see also market equilibrium
equilibrium price 67, 131
equity 6, 36, 124–5, 325–6
European Union 35
excess demand 69–70
excess supply 68
exchange controls 453
exchange rates 248, 249–50, 485–504, 515–17, 523
fixed 497–9, 502–4
floating 486, 502–4, 515
managed 500–1
see also currency
expenditure method 227–8
expenditure reducing policies 519
expenditure switching policies 518–19
export competitiveness 299, 465, 524
export earnings 26
export promotion 593–4
exports 222–3, 420, 489, 494
see also net exports
external balance 367, 387–8
external economies of scale 210
externalities 143
see also negative externalities; positive externalities

F

factor endowment 420
factor immobility 116
factors of production 8–10, 59, 249–50, 254
and equity 125
mobility of 116
fairness 36
financial account 511–12, 516–17
fines 214
finite resources 38
firms 25, 62–3
government support for 122
price takers 179–80, 192
fiscal policy 41–2, 380–99
fishing 43, 148
fixed costs 188–9
fixed exchange rate system 497–9, 502–4
floating exchange rate 486, 502–4, 515
forced pricing strategies 214
foreign aid 612–14
foreign currency reserves 497
foreign direct investment (FDI) 490–1, 512, 517, 581, 609–11
formal sector employment 275
free goods 15
free market economy 18, 37–8, 124–5, 146–7, 217
free rider problem 167–8
free trade 465–7
see also trade protection
free trade areas 471–6
frictional unemployment 285
Friedman, M. 13, 29–30, 32, 361
fruitfulness 32
full employment 19, 252, 257–8
Full Monty, The 283
future price expectations 50, 62, 245

G

game theory 203–4
'Gangnam Style' 14
gender equality 220
gender gap 343–4, 539, 556–7
gender inequality 582–3
Gender Inequality Index (GII) 556–7
generalization 31
geographical immobility 116, 283
Gini, C. 331
Gini coefficient 331–2
global economy 3
Global Entrepreneurship Monitor (GEM), on social enterprises 597
Global Gender Gap Index (GGGI) 343–4
globalization 345–6, 572
global warming 532, 551
goods 11
  balance of trade 509
  complementary 50
  degree of necessity 98–9
  demerit 130, 135, 141, 146–7, 271
  free goods 15
  inferior 49
  luxury 104–6
  merit 123–4, 132, 141, 143–6, 168, 603–4
  necessities 105
  normal goods 48–9, 105–7
  production possibility curve (PPC) 19
  public 124, 135, 165–70
  related 50, 59–60
  substitutes 51–2, 98
  see also demand; price; services; supply
Google 183
Goulet, D. 543
governance 583–5
government 25
  accounts 496, 506–25
  and asymmetric information 174–5
  borrowing 122, 262, 312–14, 513
  budget 386
  budget deficit 310, 312
  command and control (CAC) regulation and legislation 135–6
  competition policy 123
  corruption 583–5, 623
  debt 310–15
  direct provision 135, 160, 169–70
  expenditures 383–4
  failure 138
  fines 214
  fiscal policy 41–2, 380–99
  intervention 17–19, 37–8, 41–2, 121–39, 138, 150–64, 168–70, 213, 601–26
  monetary policy 42, 364–79, 504
  and nudge theory 85–6, 136–7
  ownership 214
  policy ineffectiveness 307
  price controls 125–6
  and price elasticity of demand (PED) 101–2
  priorities 247–8
  provision of information 174–5
  revenue 121–33, 381–3, 460
  spending 26, 220, 228, 244, 247–8, 310, 314–15, 360–1
  stakeholders 162
  subsidies 61, 131–3, 158–9, 250, 448–52
  supply-side policies 346, 401–18
Great Depression 1929-39 40, 42
Great Lockdown 2020 42
gross domestic product (GDP) 229–38, 266–8, 555, 559, 572
  deflator 231–2, 268
  green GDP 272
  per capita 544–5
  real 231–2, 268–9
  see also economic growth
gross national income (GNI) 231–8, 544–5
growth 89–90, 234
  in production possibilities 22–3
guaranteed minimum price 128–9

H

Halpern, D. 85
Happiness Index 239–40
Happy Planet Index (HPI) 241, 558–9
Haq, M. 554
Hardin, G.J. 148
Harford, T. 37
healthcare 410–11, 519–20, 530, 547, 554, 570, 605
health indicators 546–7
heavily indebted poor countries (HIPCs) 615–16
Herfindahl, O.C. 206
Herfindahl–Hirschman Index (HHI) 206
Heritage Foundation, on world's freest economies 18
hidden unemployment 279
Hilti Corporation 422
Hofman, J. 552–3
homogeneous products 181
household income 25, 228–9, 244, 291, 294
human capital 8–10, 342–3, 409–10
Human Development Index (HDI) 554–7
hyperinflation 290, 308
hypotheses 31–2

I

imperfect competition 184, 196
imperfect information 81–2
import expenditure 26
imports 222–3, 420, 490
import substitution 593
impure public goods 166
incentive function 72–3
incentive-related policies 408–9
income 10, 49, 217–18, 555
  consumer 99
  distribution 123, 273, 287, 323–4, 326–7, 387
  household 25, 228–9, 244, 291
  inequality 218
  Minimum Income Standard (MIS) 337
  national 227
  redistribution 123
  universal basic income (UBI) 361
  see also economic inequality; gross national income (GNI); taxation
income effect 47
income elasticity of demand (YED) 104–10
  importance of 109
income method 227
indebtedness 245, 247, 287, 575–6, 614–16
indirect taxes 61, 101–2, 129–31, 150–2, 250, 358–9, 382
industrial policies 412–13
inequality 36–7, 160, 217, 548–9, 582–5
Inequality-adjusted Human Development Index (IHDI) 557
inequality of opportunity 341
infant industries 456
inferior goods 49, 107
inflation 4, 271, 289–309, 524
  causes of 296
  cost-push 296, 321, 494
  and currency 493–4
  deflation 301–7
  demand-pull 296, 494
  disinflation 303–4
  effects of high rate 297–300
  hyperinflation 290, 308
  low 307–9, 317, 320–1, 385
  measuring 290–5
  menu costs 299
  rates around the world 300
  relative rates 493
  shoe leather costs 300
  targeting 366
inflationary gap 253, 258
informal economy 573–4
infrastructure funding 123, 412, 519, 606
injections 26, 229
institutional framework 578–9, 620–1
institutional structures 255
intellectual property rights (IPRs) 581
interdependence 6, 25, 44, 185, 201–2, 283
interest rate 10, 42, 244, 250, 365, 372–3, 517
  zero interest rate policy (ZIRP) 34
internal economies of scale 209–10
international agreements 156–7
International Bank for Reconstruction and Development (IBRD) 618
International Development Association (IDA) 618
International Monetary Fund (IMF) 513, 518, 619
International Poverty Line (IPL) 335–6, 340
international trade 420–35
  absolute advantage 429–30
  comparative advantage 429–36
intervention 6
interventionist policies 409–13, 601–26
investment 220, 244, 265
  foreign direct investment (FDI) 490–1, 512, 517, 581, 609–11
  portfolio investment 491, 512
investment expenditure 26
invisible hand 18, 37–8

J

J-curve effect 520–2
jobs protection 460–1
joint supply 60–1

K

Keynes, J.M. 29, 40–1
Keynes, J.N. 29

Keynesian economics 40–1
aggregate demand 244
aggregate supply 253–5
macroeconomic equilibrium 259–62
Keynesian multiplier 389–90
Ki-moon, B. 163
King, M.L. 81, 361
Kyoto Protocol 156–7

L

labour 8–10
costs 249
flexibility 403
force 265, 276–80
market participation 556–7
minimum age 277
mobility 265, 283
productivity 265
retirement age 277
wage-price spiral 299
see also employment; unemployment
labour market policies 407
labour theory of value 40
labour unions *see* trade unions
laissez-faire economics 29, 37–8, 42
land 8–10
landlocked countries 576–7
land rights 625–6
law of demand 46, 196
law of diminishing marginal returns 57, 118
law of diminishing marginal utility 39, 47
law of supply 56–8
law of unintended consequences 139
leakages 26, 223, 229
legislation 135–6, 153–4, 160–4, 174–5, 213
leisure time 20
life expectancy 241, 546–7
litter fines 158
Live Aid 566
living standards 271, 348–9, 497
logic 30
long run 57
long run aggregate supply (LRAS) curve 252, 266–7
long run average cost (LRAC) 194–5, 209–11
Lorenz, M.O. 330
Lorenz curve 330–1
losses 191
loyalty card schemes 205
luxury goods 104–6
LVMH Group 91

M

Mabuchi 183
macroeconomic equilibrium 256–62
macroeconomics 3–5, 41–2, 226
conflict between objectives 316–24
Malthus, T. 326
managed exchange rate 500–1
mandated choices 84
Mansudae Overseas Projects 490
manufactured products 102–3, 119
marginal cost (MC) 58, 75, 87, 117–19
marginal cost (MR) 189
marginal private benefit (MPB) 142–4, 147
marginal private cost (MPC) 142–4, 146–7, 158–9
marginal rate of transformation (MRT) 20
marginal returns *see* law of diminishing marginal returns
marginal revenue (MR) 87, 189
marginal social benefit (MSB) 142–4
marginal social cost (MSC) 142–4, 146–7, 158–9
marginal tax rate (MTR) 354
marginal utility 39, 47
market, definition 48
market access 572
market-based policies 403–9, 598–9
market concentration 206
market demand curve 48
market development 90
market disequilibrium 68–9
market equilibrium 67–76
market equilibrium wage rate 128
market failure 76, 121, 124
asymmetric information 171–7
externalities 140–63
market power 179–215
public goods 165–70
marketing 210
market leaders 88
market penetration 90
market power 179–215
market share 88
market structure 179
market supply 58–9
Marshall, A. 38, 41
Marshall-Lerner condition 521
Marshmallow Test 81
Marx, K. 29, 38, 40–1
Marxism 40
mass protest 386
maximum price 125–6
McDonald's 227
McNamara, R. 333
megacities 269
menu costs 299
mergers 90, 213
merit goods 123–4, 132, 141, 143–6, 168, 603–4
micro-credit 580
microeconomics 3–5
microfinance 620–1
Mill, J.S. 39
minimum efficient scale (MES) of production 209
Minimum Income Standard (MIS) 337
minimum lending rate (MLR) 370
minimum price 128–9
minimum reserve requirements (MRR) 369–70
minimum wage 362, 408, 603
see also National Minimum Wage (NMW)
mixed economy 18–19
mobile banking 621
mobility of factors of production 116
models 31–2
see also economic models
monetarist revolution 42
monetarists 42
aggregate supply 252, 255
macroeconomic equilibrium 257–62
monetary policy 42, 321, 364–79, 504
monetary union 476–9
money creation 367–8
money demand 371–2
money supply 296, 365, 371–2
monopolistic markets 123, 141, 186, 206–8, 212, 346, 406–7
monopoly 182–3, 196–9
MOOC 598
moral hazard 173
M-Pesa 622
MTR Corporation 383
Multidimensional Poverty Index (MPI) 338–9
multilateral development assistance (MDA) 618
multilateral trade agreements 471
Mundell-Fleming's exchange rate effect 244

N

Nash, J. 204
Nash equilibrium 204
national debt 310
national expenditure 228
National Health Service (NHS) 135
national income accounting 227–41
national income equilibrium 229
nationalization 214, 382
National Minimum Wage (NMW) 10, 128, 138
national poverty line 336
national security 456–7
natural capital 8–10
natural monopoly 199
natural rate of unemployment (NRU) 285–6, 317
natural resources 8–10
necessities 98–9, 105
needs 8, 11
negative economic growth 264–5
negative externalities 123–4, 132, 141, 147–8, 149–52, 160–1
net exports 244, 248–9, 494
net factor income from abroad 231
new classical counter revolution (NCCR) 42
nominal GDP 229
nominal interest rate 372–3
non-collusive oligopoly 184, 201
non-excludability 148, 165
non-governmental organizations (NGOs) 617
non-price competition 185, 205
non-price determinants of demand 48–53
non-price determinants of supply 59–63
non-renewable resources 8
non-rivalrous 165
non-tariff barriers 445
normal goods 48–9, 105–7
normal profit 190
normative economics 29, 34–5
nudge theory 84–6, 136–7
null hypotheses 31

O

occupational immobility 116, 283
official borrowing 513
Official Development Assistance (ODA) 616–17
Oktoberfest 453
Okun, A.M. 308
oligopolistic markets 123, 201–5, 212
oligopoly 184–5, 200–1
open economy 26
open market operations (OMO) 369
opportunistic behaviour 172
opportunity cost 13–15, 20–1, 421, 502
Organization for Economic Co-operation and Development (OECD), Better Life Index (BLI) 239

Ostrom, E. 157
Otto, A. 206
output method 227
over-consumption 168
overfishing 149
over-production 43
over-specialization 591
overvalued currencies 501

P

parallel markets 573–4
Pareto efficiency 22
PayPal 183
Peace Research Institute Oslo (PRIO) 220
peak 234
penetration pricing 205
per capita consumption (PCC) 549
perfect competition 179–81, 192–4
perfect information 78, 180
Phillips, W. 317
Phillips curve 317–20
physical capital *see* capital
Piala Inc. 86
Pigou, A.C. 150–1
Pigouvian tax 150–1
Pigou wealth effect 244
Ping An 540
planned economy 18–19
plastic waste 155
pollution 12, 152–3, 272, 527–8
portfolio investment 491, 512
positive discrimination 35
positive economics 29–30
positive externalities 143–4, 160–1
potential output 235, 266–7
poverty 12, 123, 271, 286, **332**
  absolute 333
  causes of 341–6
  and education 538
  and the gender gap 539
  International Poverty Line (IPL) 335–6
  measuring 334–41
  Multidimensional Poverty Index (MPI) 338–9
  national poverty line 336
  reducing 359–62
  relative 334, 340
  role of taxation 350–9
  and sustainability 535–9
  trap 334, 564–7
Powell, J.A. 359
preferences 49
preferential trade agreements 469–71
price 46
  ceiling 125–6
  competition 205
  constant prices 231–2, 267–8
  consumer price index (CPI) 290–5
  controls 125–6
  current prices 230, 232
  and demand 53, 68–9
  discrimination 101
  disparities 294
  elasticity 92–6, 112–13
  equilibrium 67, 131
  floor 128–9
  future expectations 50, 62, 245
  incentive function 72–3
  inelasticity 92–6, 113–14
  penetration pricing 205
  rationing function 73–4
  rigidity 185, 205
  signalling function 72–3
  stability 289
  and supply 57–8, 68–9
  volatility 592
  wars 202
  see also demand
price elasticity of demand (PED) 94–103, 131, 139, 498, 520
  importance of 101
  and total revenue 99–100
price elasticity of supply (PES) 111–20, 131, 139
  determinants 115–18
price makers 182, 196
price mechanism 31, 37, 71–6, 140
price takers 179–80, 192
primary commodities 102, 109–10, 119
primary sector 538, 571–2
prisoner's dilemma 204
private benefits 141
private costs 141
private goods 166
private sector 16–17
privatization 42, 122, 346, 382–3, 405–6, 599–600
producer surplus 74
product development 90
production 16
  level of 123
  over-production 43
  see also factors of production
production possibility curve (PPC) 19–23, 75, 266–7
productive capacity 402
products
  manufactured 102–3, 119
  see also goods; services
profit 10, 87
  abnormal 190, 211
  formula for calculating 188–9
  margins 57
  maximization 87–9, 187, 193–4, 196–9
  normal 190
progressive taxes 351–2, 397
property rights 581–2, 624–5
proportional taxes 352–3
provision of information 174–5
public goods 124, 135, 165–70
public sector 16–17
public sector borrowing 122
purchasing 210
purchasing power parity (PPP) 232–3, 340–1, 545
pure public goods 166

Q

quantitative easing 370–1
quasi-public goods 166
quotas 445–8

R

rate of substitution *see* marginal rate of transformation (MRT)
rational consumer choice 77
rationing function 73–4
Rawls, J. 585
raw material
  costs 249
  discovery 265
real-world issues 3
recession 234–5
recovery 234
recruitment costs 210
refutation 33–4
Regie Autonome des Transports Parisiens (RAPT) 200
regional trade agreements 470–1
regressive taxes 353
related goods 50, 59–60
relative interest rates 493
relative poverty 334, 340
remittances 491–2
renewable resources 8, 12
rent 10
research and development (R&D) 122, 211, 411–12
reserve assets 512–13
resource allocation 72, 140, 299, 306, 423–4
resource depletion 272
restricted choices 84
retirement age 277
revenue 187
rewards 10
Ricardo, D. 326, 436
Robbins, L. 11
Rowntree, B.S. 334
rule of thumb 79

S

Sachs, J. 565
sales revenue 88
  see also total revenue
Samuelson, P.A. 167
sanitation 607–8
satisficing 88–9
savings 26, 219–20, 298
savings ratio 564
Say, J.B. 29, 40
Say's law 38, 40–1
scarcity 6, 7–8, 11–12, 20, 22, 48, 71, 76
screening 176–7
seasonal unemployment 284
SEAT (Sociedad Española de Automóviles de Turismo) 599
self-esteem, and unemployment 286
Sen, A. 335, 530, 554
services 11
  balance of trade 509
  direct provision 135
  luxury 104–6
  production possibility curve (PPC) 19
shoe leather costs 300
shortages 73
short run 57
short run aggregate supply (SRAS) curve 249–51
short run average cost (SRAC) 193
signalling 175
signalling function 72–3
Simon, H. 88
simplicity 32
single indicators 543–4
slump 234
Smith, A. 18, 29, 37–8, 200, 203–4, 326, 334, 426, 542
smoking 135–6
social benefits 141
social costs 141
social deprivation 287
social enterprises 596–8
socially optimum output 142
social science 2–5
social stability 349
social surplus *see* community surplus 142
social sustainability 530
sovereign wealth funds (SWF) 122
specific taxes 129
speculation 492, 503
spillover effects 143
stakeholders 162
standards of living *see* living standards
start-ups 520

state pensions 602
Stiglitz, J. 238, 241
stock, ability to store 117, 119
stress, and unemployment 286
structural unemployment 282–3, 524
subsidies 61, 131–3, 158–9, 250, 448–52
substitute goods 51–2, 98
substitutes, lack of 183
substitution effect 47
'sugar tax' 124
Sunstein, C. 83
supernormal profit *see* abnormal profit
supply 56
  competitive 59–60
  contraction 64
  curve 58–64, 68
  decrease 64, 68
  excess 68
  expansion 64
  increase 64, 68
  joint 60–1
  law of 56–8
  law of diminishing marginal returns 57
  marginal cost 58, 75
  market supply 58–9
  non-price determinants 59–63, 68
  price elasticity of supply (PES) 111–20
  shortages 73
  surplus 68–9
supply-side policies 346, 401–18, 519–20
surplus 68–9, 74, 142, 439–40
surplus on an account 508
sustainability 6, 43, 149, 163, 241, 272, 321–2
  economic 530–1
  environmental 531–2
  and poverty 535–9
  social 530
sustainable development 527–39

T

tariffs 437–44
taxation 26, 122, 220–1, 244, 246–7, 271, 276, 381, 579
  ad valorem taxes 130
  average tax rate (ATR) 354
  carbon taxes 152–3
  concessions 122
  corporation tax 356–7
  direct taxes 122, 356–8, 381
  and economic inequality 345, 350–9
  flat rates 353
  impact of national debt 314–15
  indirect taxes 61, 101–2, 122, 129–31, 150–2, 250, 358–9, 382
  interventionist policies 601–2
  marginal tax rate (MTR) 354
  personal income taxes 356
  Pigouvian tax 150–1
  progressive taxes 351, 397
  proportional taxes 352–3
  rates in specific countries 386
  regressive taxes 353
  specific taxes 129
  taxable income 352
  tax allowance 351
  tax cuts 409
  wealth taxes 357–8
tecÚology 246, 254
  constant state of 20
  lack of access to 569–70
  and supply 62
telecommunications 606–7
tertiary sector 110
Tesco 215
Tetra Pak 183
Thaler, R. 43, 83–4, 136
Thatcher, M. 382–3, 405
theories 31–2
time
  and price elasticity of demand (PED) 99
  and price elasticity of supply (PES) 116, 119
Todaro, M. 543
total costs 188
total revenue 187
  and price elasticity of demand (PED) 99–100
Townsend, P. 335
tradable permits 155–6
trade agreements 469–71
trade balance *see* balance of trade
trade cycle *see* business cycle
trade liberalization 346, 402, 406, 424, 572
  see also international trade
trade policies 248
trade protection 436–54, 455–67, 519
  see also free trade
trade retaliation 463
trade unions 407–8
trade war 463
trading blocs 471–3
tragedy of the commons 148, 168
training 409–10
transfer payments 227, 360, 384, 602
transport 606
  see also infrastructure funding
triangle model 7
tropical climates 577–8

U

uncertainty 297, 304
underemployment 280
undervalued currencies 501–2
unemployment 4, 234, 275–87, 496
  benefits 408, 602
  causes of 281–5
  claimant count 277
  costs of 286–7
  cyclical 281–2, 306
  discouraged workers 279
  disguised 279
  disparities 280
  frictional 285
  hidden 279
  low 275, 307–9, 317, 366, 385
  measuring 276–80
  natural rate of 285–6, 317
  rate 276–7, 285–6
  seasonal 284
  structural 282–3
  underemployment 280
  voluntary 279–80, 285–6
  see also employment
unfair competition 459
unitary price elastic demand 95
unitary price elastic supply 112, 115
United Nations
  on carbon taxes 153
  on education 571
  Framework Convention on Climate Change (UNFCCC) 1992 156–7
  Happiness Index 239–40
  International Labour Organization (ILO) 277
  Millennium Development Goals (MDGs) 533–5
  on poverty 335
  Sustainable Development Goals (SDGs) 44, 163, 335, 338, 529, 533–7, 582, 631–4
  Universal Declaration of Human Rights 11
United Nations Educational, Scientific and Cultural Organization (UNESCO) 527
universal basic income (UBI) 361
unused capacity 116–17
utility maximization 78
utility theory 38–9

V

value judgements 35–6, 138
variable costs 188
Volkswagen scandal 175
Volkswagon scandal 457
voluntary unemployment 279–80, 285–6

W

wage-price spiral 299
wages 10, 276
  minimum wage 362, 408, 603
  National Minimum Wage (NMW) 10, 128, 138
  see also employment
wants 8, 11
waste disposal management 552–3
wealth 218, 244
  distribution of 4, 123, 217–23, 273, 287, 328–30
  redistribution 305
  taxes 357–8
  see also economic inequality
well-being 6
well-being measures 238–41
WISE ChoiCES concepts 5–7
withdrawals *see* leakages
women's empowerment 622–3
  see also gender gap
working hours 20
World Bank
  and economic development 618
  on poverty 12, 335
World Trade Organization (WTO) 479–83
  multilateral trade agreements 471

Y

YKK 183

Z

zero interest rate policy (ZIRP) 34

Acknowledgements

Text permissions

t = top b = bottom r = right l = left

p.3 © IB Diploma Programme Economics guide, page 7; **p.11** The Universal Declaration of Human Rights, United Nations, Articles 25 and 26; **p.18** Highlights of the 2019 Index of Economic Freedom, © The Heritage Foundation; **p.37** Tim Harford, *The undercover economist: exposing why the rich are rich, the poor are poor, and why you can never buy a decent used car*, © 2006, Oxford University Press, reproduced with permission of the Licensor through PLSclear; **p.149** *These are the world's 20 most congested cities* by Mark Matousek, April 8, 2019, © Insider Inc; **p.158** adapted from 'You can be fined £150 for littering from your car even if a passengers does it' by Luke John Smith, May 29 2019, © The Express; **p.162** U.S. Department of Health & Human Services, 'Influenza (flu) in the workplace', NIOSH, Flu-related Work; Flu Season To Cost Employers $21B, © Challenger, Gray & Christmas; **p.181** 'Overview of Sydney Fish Market', © Tripadvisor LLC; **p.200** 'Getting Around', © 2020 Lonely Planet; **p.220** *Women, Peace, and Security Index, 2019/20 Report*, © Georgetown Institute for Women, Peace and Security; **p.221** List of Countries by Personal Income Tax Rate, © 2020 Trading Economics, Corporate tax rates table, © KPMG International; **p.222** Graph of index of coffee prices 2010–2020, © 2020, Trading Economics; **p.236** *t* Vietnam's GDP, $bn (1990–2020), © 2020 Trading Economics, *b* Honduras GDP per capita, 2020 © Trading Ecnomics, Iceland GDP per capita, 2020 © Trading Economics; **p.237** *Shadow Economies Around the World: What Did We Learn Over the Last 20 Years?* By Leandro Medina and Friedrich Schneider, January 24, 2018, © 2018, International Monetary Fund; **p.240** Helliwell, J., Layard, R., & Sachs, J. (2019). *World Happiness Report 2019*, New York: Sustainable Development Solutions Network; **p.241** *t* About the HPI, © New Economics Foundation, *b* About the HPI, © New Economics Foundation; **p.250** Sales Tax rate, 2020 © Trading Economics; p.269 GDP per capita, select countries (2020), 2020 © Trading Economics; **p.270** *t* Vietnam's GDP $bn, (2014-20), 2020 © Trading Economics, *b* data from Nominal GDP ($bn) and Population (million), 2020 © Trading Economics; **p.271** UBS Global Real Estate Bubble Index 2018, 27 Sep 2018, © UBS; **p.282** Spain's unemployment rate (%) (1195-2020), © 2020 Trading Economics; **p.300** Inflation rates around the world, selected countries, March 2020, © 2020 Trading Economics; **p.301** Deflation rates around the world, selected countries, March 2020 © 2020 Trading Economics; **p.303** Japan's inflation rate (%) 1990–2020, © 2020, Trading Economics; **p.308** Steve Hanke, 'Hanke's Annual Misery Index 2018: The World's Saddest (And Happiest) Countries', Mar 28, 2019, © Forbes Media LLC.; **p.309** Thailand's inflation rate, © 2020 Trading Economics; **p.311** *tl* Government debt (selected countries) March 2020, © 2020 Trading Economics, *br* Country list government debt to GDP (selected countries), March 2020, © 2020 Trading Economics; **p.312** Spain's government debt to GDP, © 2020, Trading Economics; p.313 *t* United States government debt, ($m), 2010–20, © 2020 Trading Economics, *b* US average monthly prime lending rate 2015–20, © 2020 Trading Economics; **p.314** *l* Greece central government debt ((€m) 2007-20), © 2020, Trading Economics, *r* Spain general government debt ((€m) 2007-20), © 2020 Trading Economics; **p.323** 'The Richest in 2020', © Forbes Media LLC.; **p.327** GDP for top and bottom 10 countries in purchasing power parity (2020), © 2020 Trading Economics; **p.329** Credit Suisse Research institute 'Global wealth book' 2018, © Credit Suisse Group AG; **p.331** *t* © 2020 GroundUp.'15% of SA's wealth is in the hands of just 3,500 people, study finds' by James Stent, 10 March 2020. **https://www.groundup.org.za/article/15-sas-wealth-hands-just-3500-people-study-finds/**; *b* Gini coefficients, © World Population Review; **p.333** Poverty headcount ratio at $1.90 a day (2011 PPP) (% of population), © The World Bank Group; **p.335** The World Factbook 2020. *Washington, DC: Central Intelligence Agency, 2020.* **https://www.cia.gov/library/publications/the-world-factbook/geos/by.html**; **p.338** 2018 MPI: Dimensions, indicators, deprivation cutoffs, and weights, United Nations Development Programme (UNDP), Human Development Report Office (HDRO); **p.339** *t* United Nations Development Programme 'Illuminating Inequalities', 11 July 2019, *b* Global Multidimensional Poverty Index 2019, *Illuminating Inequalities*, © United Nations Development Programme and Oxford Poverty and Human Development Initiative; **p.344** *t* Global Gender Gap Report 2020, ISBN 9782940631032, © 2019, The World Economic Forum, *b* 'Women's ownership of land in rural areas can help cut poverty' by Kritika Singh, 03 Aug 2016, © 2016, HT Media Limited; **p.347** 'Life Is Getting Even Harder for Garbage Sifters in the Developing World' by The Conversation, March 2, 2020, © U.S. News & World Report L.P; **p.349** *Global Study on Homicide 2019* © United Nations; **p.356** *t* Countries with the highest rates of personal income tax, © The Heritage Foundation, *b* WIDER Working Paper 2017/55, *Global distribution of revenue loss from tax avoidance*, ISBN 9789292562793, March 2017, © UNU-WIDER 2017; **p.357** *l* OECD, Statutory Corporate Income Tax Rates, **https://stats.oecd.org/Index.aspx?DataSetCode=TABLE_II1**, *r* WIDER Working Paper 2017/55, Global distribution of revenue loss from tax avoidance, ISBN 9789292562793, March 2017, © UNU-WIDER 2017; **p.375** *t* Spain interest rates 2010-20, © 2020, Trading Economics, *b* Argentina interest rates 2010-20, © 2020, Trading Economics; **p.386** Corporate tax rate – by country, © 2020, Trading Economics, Sales tax rate – by country, © 2020, Trading Economics, List of countries by personal income tax rate, © 2020, Trading Economics; **p.387** List of countries by personal income tax rate, © 2020, Trading Economics; **p.411** Nurses by country, © 2020, Trading Economics, Medical doctors by country, © 2020, Trading Economics; **p.460** Countries with the highest tariffs in the world, © World Bank, 2018 data; **p.462** World Bank. 2020. Global Economic Prospects, June 2020. Washington, DC: World Bank. DOI: 10.1596/978-1-4648-1553-9. License: Creative Commons Attribution CC BY 3.0 IGO. This is an adaptation of an original work

by The World Bank. Views and opinions expressed in the adaptation are the sole responsibility of the author or authors of the adaptation and are not endorsed by The World Bank; **p.492** *Personal and Cash Remittances Set Record Highs in 2019 at US$33.5 Billion and US$30.1 Billion,* February 17, 2020, © Bangko Sentral ng Pilipinas (BSP); **p.504** Foreign currency reserves, June 2020, © 2020 Trading Economics; **p.508** IB Economics Guide, page 49; **p.521** France balance of trade ($bn) 2016-20, © 2020 Trading Economics; **p.525** Australia balance of trade (millions of AUD) March 2019-February 2020, © 2020 Trading Economics; **p.529** United Nations, *Report of the World Commission on Environment and Development*, Our Common Future, © United Nations; **p.534** 'Transforming our world: the 2030 Agenda for Sustainable Development', A/RES/70/1, © United Nations. Reprinted with the permission of the United Nations; **p.536** *t* Poverty and Shared Prosperity *Extreme poverty* © 2020 The World Bank Group, *b* Poverty headcount ratio at $1.90 a day (2011 PPP) (% of population), © 2020 The World Bank Group; **p.540** 'The World Bank In Rwanda', © 2020 The World Bank Group; **p.545** Projected GDP per capita Ranking, © StatisticsTimes.com; **p.546** Life expectancy at birth and mortality rate, © 2020 The World Bank Group; **p.547** Healthcare expenditure as percentage of GDP, selected countries, © 2020 The World Bank Group; **p.548** Table 1: Human Development Index and its components, United Nations Development Programme (UNDP), Human Development Report Office (HDRO); Literacy rate, adult total (% of people ages 15 and above, © 2020 The World Bank Group; **p.550** Access to electricity (% of the population), © 2020 The World Bank Group; **p.553** Top five countries for municipal solid waste, 2019, © 2020 The World Bank Group; **p.556** Proportion of seats held by women in national parliaments (%), © 2020 The World Bank Group; **p.557** *tr* Highest female labour force participation ratios, © 2020 The World Bank Group, *b* Inequality-adjusted Human Development Index (IHDI), United Nations Development Programme (UNDP), Human Development Report Office (HDRO); **p.558** Human Development Report 2019 'Beyond income, beyond averages, beyond today: Inequalities in human development in the 21st century', United Nations Development Programme; **p.565** Jeffrey Sachs, *The End of Poverty: How We Can Make it Happen in Our Lifetime,* ISBN 9780241958667, © 2011, Penguin UK; **p.568** adapted from World development indicators, © 2020 The World Bank Group; **p.571** Education index, United Nations Development Programme (UNDP), Human Development Report Office (HDRO); **p.572** The World Factbook, Central Intelligence Agency; **p.573** *t* Trade Map, List of exporters for the selected product in 2019, International Trade Centre, **www.intracen.org/marketanalysis**, *b* Bonnet, Florence, Joann Vanek and Martha Chen. 2019. *Women and Men in the Informal Economy – A Statistical Brief.* Manchester, UK: WIEGO; **p.576** Debt to GDP Ratio by Country 2020, © World Population Review; **p.579** Ease of Doing Business rankings, © 2020 The World Bank Group; **p.580** *The world's 2 billion unbanked*, *in 6 charts* by Camilla Hodgson, Aug 30, 2017, © Insider Inc.; **p.584** The world's most honest and corrupt countries, 2020, © Transparency International 2020, *r* Worldwide governance indicators, © The World Bank Group; **p.604** Education expenditure as a proportion of GDP, © 2020 The World Bank Group; **p.605** Healthcare spending as a percentage of GDP, © 2020 The World Bank Group; **p.611** Egypt: Foreign Investment, © Export Entreprises SA; **p.617** OECD, 'Development aid in 2015 continues to grow despite costs for in-donor refugees', 13 April 2016, **http://www.oecd.org/dac/stats/ODA-2015-detailed-summary.pdf.**

Photo credits

t = top b = bottom r = right l = left

p.vii © Robert Kneschke/stock.adobe.com; **p.1** © Anke Scheibe/Westend61 GmbH/Alamy Stock Photo; **p.3** © 3D generator/stock.adobe.com; **p.4** *t* © Alfa27/stock.adobe.com, *b* © M-SUR/stock.adobe.com; **p.7** © _Jure/stock.adobe.com; **p.8** © RTimages/stock.adobe.com; **p.9** © Monkey Business/stock.adobe.com; **p.12** © Boysitti/stock.adobe.com; **p.13** © Daniel Jędzura/stock.adobe.com; **p.15** © Katiekk2/stock.adobe.com; **p.16** © Andrey Kiselev/stock.adobe.com; **p.17** © PixieMe/stock.adobe.com; **p.22** © Michail Petrov/stock.adobe.com, **p.23** © Uwe/stock.adobe.com; **p.29** © Mushy/stock.adobe.com; **p.30** © Eggeeggjiew/stock.adobe.com; **p.32** © Exclusive-design/stock.adobe.com; **p.33** © MarekPhotoDesign.com/stock.adobe.com; **p.34** © Rocketclips/stock.adobe.com; **p.35** © Lazyllama/stock.adobe.com; **p.36** © Vanessa Volk/stock.adobe.com; **p.38** © Saktanong/stock.adobe.com; **p.39** *t* © DiamondGalaxy/stock.adobe.com, *b* © Mihail/stock.adobe.com; **p.41** © Andrey Popov/stock.adobe.com; **p.43** © Elena/stock.adobe.com; **p.45** © Nimon_t/stock.adobe.com; **p.47** © Picsfive/stock.adobe.com; **p.50** © Paul Hoang; **p.51** *t* © Tyler Olson/stock.adobe.com; *b* © Aunging/stock.adobe.com; **p.52** © Jess Yu/stock.adobe.com; **p.53** © Bennymarty/stock.adobe.com; **p.56** © FrankBoston/stock.adobe.com; **p.58** © Aekkorn/stock.adobe.com; **p.60** © Juefraphoto/stock.adobe.com; **p.62** © Methaphum/stock.adobe.com; **p.63** © Joeyphoto/stock.adobe.com; **p.64** © Weyo/stock.adobe.com; **p.69** © Sas/stock.adobe.com; **p.72** © Norman Chan/stock.adobe.com; **p.73** © William/stock.adobe.com; **p.78** Geralt/22038 images; **p.79** © Paul Hoang; **p.81** © InsideCreativeHouse/stock.adobe.com; **p.82** *t* © Pantovich/stock.adobe.com, *b* © Rawpixel.com/stock.adobe.com; **p.85** *t* © Lightfield Studios/stock.adobe.com, *m* © RomanR/stock.adobe.com; **p.86** © Bodnarphoto/stock.adobe.com; **p.93** © Monticellllo/stock.adobe.com; **p.96** © Ilike/stock.adobe.com; **p.98** © Andrey Popov/stock.adobe.com; **p.102** © Hd3dsh/stock.adobe.com; **p.105** © Pkazmierczak/stock.adobe.com; **p.106** *tl* © Motortion/stock.adobe.com, *tr* © Kudosstudio/stock.adobe.com; **p.109** © Michelaubryphoto/stock.adobe.com; **p.110** © Unique Vision/stock.adobe.com; **p.112** © Grispb/stock.adobe.com; **p.117** © Bjorn Wylezich/stock.adobe.com; **p.118** © Mirko/stock.adobe.com; **p.119** © Paul Hoang; **p.123** © Wollwerth Imagery/stock.adobe.com; **p.124** © Celso Pupo/stock.adobe.com; **p.125** © Pixfly/stock.adobe.com; **p.127** © Chanchai/stock.adobe.com; **p.128** © New Africa/stock.adobe.com; **p.135** *t* © Paul Hoang, *b* © Christopher Dodge/stock.adobe.com; **p.137** © John/stock.adobe.com; **p.141** © U. J. Alexander/stock.adobe.com; **p.143** © VPales/stock.adobe.com; **p.144** © Ronniechua/stock.adobe.com; **p.145** © Antonel/stock.adobe.com; **p.146** © Beats_/stock.adobe.com; **p.149** © Alex/stock.adobe.com; **p.151** © Trgowanlock/stock.adobe.com; **p.152** © Kamilpetran/stock.adobe.com; **p.158** © Vyacheslav Dumchev/stock.adobe.com; **p.160** © Vectorfusionart/stock.adobe.com; **p.161** © WavebreakmediaMicro/stock.adobe.com; **p.162** © Vichie81/stock.adobe.com; **p.166** © P. Meybruck/stock.adobe.com; **p.167** © Tomas/stock.adobe.com; **p.168** © Toa555/stock.adobe.com; **p.169** © Mbruxelle/stock.adobe.com;

p.171 © BillionPhotos.com/stock.adobe.com; **p.172** © Brenda Carson/stock.adobe.com; **p.173** © Oleksandr Delyk/stock.adobe.com; **p.174** © Ron Dale/stock.adobe.com; **p.176** *t* © Sergey Kohl/stock.adobe.com, *b* © Piman Khrutmuang/stock.adobe.com; **p.180** © Uwe/stock.adobe.com; **p.181** © Ymgerman/stock.adobe.com; **p.184** © Akhenatonimages/stock.adobe.com; **p.185** © Sablin/stock.adobe.com; **p.186** © Ansyvan/stock.adobe.com; **p.187** © Thithawat/stock.adobe.com; **p.191** © Jo Panuwat D/stock.adobe.com; **p.200** © Savvapanf Photo/stock.adobe.com; **p.202** © George Serban/stock.adobe.com; **p.205** *t* © VTT Studio/stock.adobe.com, *b* © Georges Blond/stock.adobe.com; **p.206** © Pineapple studio/stock.adobe.com; **p.208** © DifferR/stock.adobe.com; **p.210** © Gguy/stock.adobe.com; **p.215** © Quality Stock Arts/stock.adobe.com; **p.218** © Infra2808/stock.adobe.com; **p.221** © Rocketclips/stock.adobe.com; **p.222** © Andrey Kuzmin/stock.adobe.com; **p.225** © Peshkov/stock.adobe.com; **p.227** © Tupungato/stock.adobe.com; **p.238** © Perytskyy/stock.adobe.com; **p.240** © Song_about_summer/stock.adobe.com; **p.247** © THANANIT/stock.adobe.com; **p.260** © Mongkolchon/stock.adobe.com; **p.262** © Andrey Yalansky/stock.adobe.com; **p.265** © Tichr/stock.adobe.com; **p.272** © Marina_larina/stock.adobe.com; **p.273** © Bennymarty/stock.adobe.com; **p.279** © Prostock-studio/stock.adobe.com; **p.284** © Last19/stock.adobe.com; **p.287** © Photographee.eu/ stock.adobe.com; **p.290** © Swisshippo/stock.adobe.com; **p.298** © 2mmedia/stock.adobe.com; **p.300** © Tupungato/stock.adobe.com; **p.305** © Andrey Yalansky/stock.adobe.com; **p.308** © Vadim0001/stock.adobe.com; **p.314** © Wirojsid/stock.adobe.com; **p.318** © Wirojsid/stock.adobe.com; **p.319** © Andrey Popov/stock.adobe.com; **p.322** © Dj_mono/stock.adobe.com; **p.327** © Gustavo/stock.adobe.com; **p.328** © Tanu/stock.adobe.com; **p.336** © Fablok/stock.adobe.com; **p.337** © Gustavofrazao/stock.adobe.com; **p.340** © Kara/stock.adobe.com; **p.342** © Yurii Zymovin/stock.adobe.com; **p.344** © Blueaperture15/stock.adobe.com; **p.347** © Tinnakorn/stock.adobe.com; **p.348** © Delphotostock/stock.adobe.com; **p.351** © Monkey Business/stock.adobe.com; **p.357** © Apfelweile/stock.adobe.com; **p.358** © Victoria/stock.adobe.com; **p.360** © Living Legend/stock.adobe.com; **p.361** © GL Archive/Alamy Stock Photo; **p.365** © Scanrail/stock.adobe.com; **p.368** © Prostock-studio/stock.adobe.com; **p.377** © Prostock-studio/stock.adobe.com; **p.382** © Vegefox.com/stock.adobe.com; **p.383** © Heorshe/stock.adobe.com; **p.384** © Soonthorn/stock.adobe.com; **p.386** © Arthur/stock.adobe.com; **p.392** © EtiAmmos/stock.adobe.com; **p.394** © Ake1150/stock.adobe.com; **p.396** © Norman Chan/stock.adobe.com; **p.404** © Przemysław Szabłowsk/stock.adobe.com; **p.406** © Zephyr_p/stock.adobe.com; **p.409** © Xtock/stock.adobe.com; **p.410** © Mnirat/stock.adobe.com; **p.411** © Gorodenkoff/stock.adobe.com; **p.412** © Creativenature.nl/stock.adobe.com; **p.416** © Vichie81/stock.adobe.com; **p.419** © Pink candy/stock.adobe.com; **p.421** © Stveak/stock.adobe.com; **p.422** © Bjorn Wylezich/stock.adobe.com; **p.423** © Funtap/stock.adobe.com; **p.427** © Xamtiw/stock.adobe.com; **p.431** © Swisshippo/stock.adobe.com; **p.434** © Phonlamaiphoto/stock.adobe.com; **p.435** © THATREE/stock.adobe.com; **p.437** © Olivier Le Moal/stock.adobe.com; **p.439** © Aleksandra Gigowska/stock.adobe.com; **p.449** © Martinfredy/stock.adobe.com; **p.453** *t* © Sutthinon602/stock.adobe.com, *b* © Rawf8/stock.adobe.com; **p.456** © Elnur/stock.adobe.com; **p.457** © kittisak/stock.adobe.com; **p.458** © Kadmy/stock.adobe.com; **p.461** © S@photo/stock.adobe.com; **p.462** © Ancamilushev/stock.adobe.com; **p.463** © Destina/stock.adobe.com; **p.464** © Richie Chan/stock.adobe.com; **p.465** © EwaStudio/stock.adobe.com; **p.466** © Scanrail/stock.adobe.com; **p.470** © F11photo/stock.adobe.com; **p.472** © Malajscy/stock.adobe.com; **p.473** © Eisenhans/stock.adobe.com; **p.476** © Rawpixel.com/stock.adobe.com; **p.478** © Ilolab/stock.adobe.com; **p.480** © Ricochet64/stock.adobe.com; **p.482** © Coloures-Pic/stock.adobe.com; **p.486** © Gumpapa/stock.adobe.com; **p.490** © JHVEPhoto/stock.adobe.com; **p.492** © Fazon/stock.adobe.com; **p.493** © Nikkimeel/stock.adobe.com; **p.497** © Hossein1351/stock.adobe.com; **p.499** © RomanR/stock.adobe.com; **p.500** © ThamKC/stock.adobe.com; **p.503** © Kebox/stock.adobe.com; **p.507** *t* © Phaisarnwong2517/stock.adobe.com, *b* © Eakkaluk/stock.adobe.com; **p.509** © GoodIdeas/stock.adobe.com; **p.512** © Ar130405/stock.adobe.com; **p.513** © Unshu/stock.adobe.com; **p.517** © Bence/stock.adobe.com; **p.521** © Art_zzz/stock.adobe.com; **p.528** © Wusuowei/stock.adobe.com; **p.531** © Doomu/stock.adobe.com; **p.532** © Napior/stock.adobe.com; **p.533** © United Nations Development Programme; **p.535** **https://www.un.org/sustainabledevelopment/**; **p.536** © Dennisvdwater/stock.adobe.com; **p.537** © Oneinchpunch/stock.adobe.com; **p.538** © Saktanong/stock.adobe.com; **p.539** © Anton Gvozdikov/stock.adobe.com; **p.542** © Andras/stock.adobe.com; **p.544** **https://www.un.org/sustainabledevelopment/**; **p.548** © Monkey Business/stock.adobe.com; **p.549** © Monkey Business/stock.adobe.com; **p.550** © Coprid/stock.adobe.com; **p.551** *t* © From Global-average temperature records by Dr Peter Stott, Head Climate Monitoring and Attribution at the Met Office. Contains public sector information licensed under the Open Government Licence v1.0, *b* © Vom/stock.adobe.com; **p.552** © Elena/stock.adobe.com; **p.553** © Andrea/stock.adobe.com; **p.544** © Tolly65/stock.adobe.com; **p.558** © Monkey Business/stock.adobe.com; **p.560** © Arthon/stock.adobe.com; **p.561** © Schiddrigkeit/stock.adobe.com; **p.565** © Kay Nietfeld/dpa picture alliance/Alamy Stock Photo; **p. 566** © David Davis/stock.adobe.com; **p.569** © Zhu difeng/stock.adobe.com; **p.572** © Robertonencini/stock.adobe.com; **p.575** © Dmitry Chumichev/stock.adobe.com; **p.576** © Chungking/stock.adobe.com; **p.577** © Nechaevkon/stock.adobe.com; **p.580** © Kg58/stock.adobe.com; **p.581** © Poco_bw/stock.adobe.com; p.**582** © Olivier Le Moal/stock.adobe.com; **p.583** © Fizkes/stock.adobe.com; **p.585** © Michele/stock.adobe.com; **p.588** © Zerophoto/stock.adobe.com; **p.592** © Rawpixel.com/stock.adobe.com; **p.596** *t* © Studiographicmh/stock.adobe.com, *b* © William/stock.adobe.com; **p.597** © Artur/stock.adobe.com; **p.598** © Oculo/stock.adobe.com; **p.600** © iQoncept/stock.adobe.com; **p.602** © Vitalii Vodolazskyi/stock.adobe.com; **p.603** © Danmir12/stock.adobe.com; **p.605** © Yury Zap/stock.adobe.com; **p.606** © Madpixblue/stock.adobe.com; **p.607** © Rukawajung/stock.adobe.com; **p.608** © Riccardo Niels Mayer/stock.adobe.com; **p.612** © Adil/stock.adobe.com; **p.615** © Zimmytws/stock.adobe.com; **p.619** © Bumble Dee/stock.adobe.com; **p.621** © Friends Stock/stock.adobe.com; **p.622** © IrynaV/stock.adobe.com; **p.624** © Makaule/stock.adobe.com; **p.625** © Dr322/stock.adobe.com; **p.627** © Imtmphoto/stock.adobe.com; **p.633** © Rudie/stock.adobe.com; **p.634** © Uwe/stock.adobe.com.

FOR THE
IB DIPLOMA
PROGRAMME

Economics

Ensure full coverage of the syllabus with this comprehensive guide, written by highly experienced authors, that explores all topics through inquiry, concepts, content and contexts and their interrelationships with each other.

- Explore economics through the nine key concepts with our newly-developed acronym for a deeper and more integrated understanding of the subject.
- Ensure comprehension with definitions of key terms and diagrams throughout the book.
- Apply new skills and knowledge to everyday life with examples and case studies.
- Foster the attributes of the IB learner profile with inquiry activities and TOK links for every topic.
- Prepare for the new assessment model with exam-style questions.
- Avoid common mistakes and aid understanding with expert tips and hints.

About the authors

Paul Hoang is an experienced Economics teacher and the author of several best-selling titles for IB and IGCSE Economics. He has held the position of Vice Principal and IB Diploma Programme Coordinator at a renowned IB World School in Hong Kong. He is a highly experienced Economics workshop leader.

Sean Wray teaches Economics at the English Schools Foundation's King George V School in Hong Kong. He is the current ESF Curriculum Group Leader for Business and Economics and has written for the IB Review magazine.

Tanusankar Chakraborty is an experienced teacher and workshop leader. He currently teaches IB Economics at the Western Academy of Beijing and has previously held positions including Department Head, Extended Essay Coordinator and Diploma Programme Coordinator.

Dynamic Learning

This book is fully supported by Dynamic Learning – the online subscription service that helps make teaching and learning easier. Dynamic Learning provides unique tools and content for:

- front-of-class teaching
- streamlining planning and sharing lessons
- focused and flexible assessment preparation
- independent, flexible student study

Sign up for a free trial – visit: **www.hoddereducation.com/dynamiclearning**

HODDER EDUCATION
e: education@bookpoint.co.uk
w: hoddereducation.com

ISBN 978-1-510-47914-2
9 781510 479142
FSC