

AQA A-Level Year 2

Economics

Malcolm Surridge & John Wolinski

To Alex, with much love and gratitude for her wonderful support during the writing of this book and its companion volume, from Malcolm

To Yvonne, with thanks and love for her tremendous tolerance and support during the writing of this book and its companion volume, from John

Published by **Cross Academe Limited**, St John's House, 5 South Parade, Oxford OX2 7JL

Orders

Cross Academe Limited, PO Box 105, Rochester, Kent ME2 4BE

Tel: 01634 729825

Fax: 01634 290175

You can also order through the Cross Academe website: www.crossacademe.co.uk

© Cross Academe Ltd 2016

ISBN 978-1-909592-81-0

Copyright notice

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic or mechanical, photocopying, recording or otherwise without the written permission of the publisher or under licence from the Copyright Licensing Agency Limited, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

Photograph acknowledgements

The cover photograph shows a pair of doors at the Bank of England and is reproduced by permission of pictureline/Alamy Stock Photo.

Other photographs have been reproduced by permission of Juice Images/Alamy Stock Photo (p. 12), Andrew Fox/Alamy Stock Photo (p. 18), Jon Parker Lee/Alamy Stock Photo (p. 22), Alex Segre/Alamy Stock Photo (p. 33), JG Photography/Alamy Stock Photo (p. 38), Cultura Creative RF/Alamy Stock Photo (p. 43), Matthew Chattle/Alamy Stock Photo (p. 48), Kevin Britland/Alamy Stock Photo (p. 53 left), Radharc Images/Alamy Stock Photo (p. 53 right), aber CPC/Alamy Stock Photo (p. 61), Ed Simons/Alamy Stock Photo (p. 74), Andrea Heselton/Alamy Stock Photo (p. 83), keith morris/Alamy Stock Photo (p. 90), epa European pressphoto agency b.v./Alamy Stock Photo (p. 104), keith morris/Alamy Stock Photo (p. 109), Ian Patrick/Alamy Stock Photo (p. 117 bottom), David Burton/Alamy Stock Photo (p. 117 top), Purepix/Alamy Stock Photo (p. 120), PURPLE MARBLES/Alamy Stock Photo (p. 131), Ingram Publishing/Alamy Stock Photo (p. 137), Charles Robertson/Alamy Stock Photo (p. 146), Medicshots/Alamy Stock Photo (p. 152), Rosemary Roberts/Alamy Stock Photo (p. 175), Image Source/Alamy Stock Photo (p. 181), Daniel Atkin/Alamy Stock Photo (p. 202), Oramstock/Alamy Stock Photo (p. 215), David Osborn/Alamy Stock Photo (p. 223), John Gaffen/Alamy Stock Photo (p. 228), Paul Stewart/Alamy Stock Photo (p. 236 left), Christopher Briggs/Alamy Stock Photo (p. 236 right), TravelStockCollection – Homer Sykes/Alamy Stock Photo (p. 253), Justin Kase z12z/Alamy Stock Photo (p. 265), GLC Pix/Alamy Stock Photo (p. 276), richard sowersby/Alamy Stock Photo (p. 285), epa European pressphoto agency/Alamy Stock Photo (p. 327), Geoffrey Taunton/Alamy Stock Photo (p. 367), Adrian Sherratt/Alamy Stock Photo (p. 397) and Lou Linwei/Alamy Stock Photo (p. 410).

Design by Ralph Hall

Printed by Bell & Bain Limited, Glasgow

The paper on which this title is printed is sourced from managed, sustainable forests.

Contents

Preface vi

Section 1 Individuals, firms, markets & market failure

Topic 1 Economic methodology & the economic problem

Topic 2 Individual economic decision making

Chapter 1	Consumer behaviour	6
Chapter 2	Imperfect information	11
Chapter 3	Aspects of behavioural economic theory	14
Chapter 4	Behavioural economics & economic policy	20
	<i>Exam-style questions</i>	23

Topic 3 Price determination in a competitive market

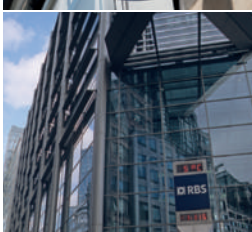
Topic 4 Production, costs & revenue

Chapter 5	The law of diminishing returns & returns to scale	28
Chapter 6	Costs of production	35
Chapter 7	Economies & diseconomies of scale	40
Chapter 8	Marginal, average & total revenue	45
Chapter 9	Profit	50
Chapter 10	Technological change	55
	<i>Exam-style questions</i>	63

Topic 5 Perfect competition, imperfectly competitive markets & monopoly

Chapter 11	Market structures	66
Chapter 12	The objectives of firms	69
Chapter 13	Perfect competition	76
Chapter 14	Monopolistic competition	86
Chapter 15	Oligopoly	92
Chapter 16	Monopoly & monopoly power	107
Chapter 17	Price discrimination	111
Chapter 18	The dynamics of competition & competitive market processes	115
Chapter 19	The concepts of contestable & non-contestable markets	118





Chapter 20	Market structure, static efficiency, dynamic efficiency & resource allocation	122
Chapter 21	Consumer & producer surplus	134
	<i>Exam-style questions</i>	139

Topic 6 The labour market

Chapter 22	The demand for labour & marginal productivity theory	142
Chapter 23	Influences upon the supply of labour to different markets	149
Chapter 24	The determination of relative wage rates & levels of employment in perfectly competitive labour markets	155
Chapter 25	The determination of relative wage rates & levels of employment in imperfectly competitive labour markets	161
Chapter 26	The influence of trade unions in determining wages & levels of employment	166
Chapter 27	The national minimum wage	173
Chapter 28	Discrimination in the labour market	177
	<i>Exam-style questions</i>	183

Topic 7 The distribution of income & wealth: poverty & inequality

Chapter 29	The distribution of income & wealth	186
Chapter 30	The problem of poverty	199
Chapter 31	Government policies to influence the distribution of income & wealth & to alleviate poverty	204
	<i>Exam-style questions</i>	209

Topic 8 The market mechanism, market failure & government intervention in markets

Chapter 32	How markets and prices allocate resources	212
Chapter 33	Aspects of market failure	217
Chapter 34	Competition policy	225
Chapter 35	Public ownership, privatisation, regulation & deregulation of markets	230
Chapter 36	Government intervention in markets	239
	<i>Exam-style questions</i>	243

Section 2 The national & international economy

Topic 9 The measurement of macroeconomic performance

Chapter 37	Uses of national income data	248
	<i>Exam-style questions</i>	259

Topic 10	How the macroeconomy works	
Chapter 38	Aggregate demand & the level of economic activity	262
Chapter 39	The determinants of long-run aggregate supply	270
	<i>Exam-style questions</i>	279
Topic 11	Economic performance	
Chapter 40	Economic growth & the economic cycle	282
Chapter 41	Employment & unemployment	293
Chapter 42	Inflation & deflation	302
Chapter 43	Possible conflicts between macroeconomic policy objectives	311
	<i>Exam-style questions</i>	321
Topic 12	Financial markets & monetary policy	
Chapter 44	The structure of financial markets & financial assets	324
Chapter 45	Commercial banks & investment banks	338
Chapter 46	Central banks & monetary policy	351
Chapter 47	The regulation of the financial system	361
	<i>Exam-style questions</i>	373
Topic 13	Fiscal policy & supply-side policies	
Chapter 48	Fiscal policy	376
Chapter 49	Supply-side policies	392
	<i>Exam-style questions</i>	399
Topic 14	The international economy	
Chapter 50	Globalisation	402
Chapter 51	Trade	417
Chapter 52	The balance of payments	439
Chapter 53	Exchange rate systems	455
Chapter 54	Economic growth & development	468
	<i>Exam-style questions</i>	487
A-Level Paper 3 exam-style questions		489
Index		497



Preface

This book has been written to meet the needs of students studying the second year of AQA A-level economics and is a companion volume to *AQA AS/A-Level Year 1 Economics*. It provides full coverage of subject content in the specification for the second year of AQA A-level economics. The contents of the book have been written to match the relevant subject content, both by topic and by chapter, and they build throughout on the material covered in the companion volume. This volume contains up-to-date data as well as topical and relevant examples to illustrate economic theories and principles. You should use it alongside the AS/Year 1 textbook. There are numerous cross-references between the two volumes because many theories and topics are spread across the elements of the AQA A-level economics specification.

Special features

AQA A-Level Year 2 Economics has a number of special features that have been incorporated to extend your understanding of the requirements of the course:

- **Key concepts** from Year 1 at the start of every chapter, reviewing the key issues covered in the companion Year 1 textbook
- **Overviews** for every chapter, outlining content coverage
- **Real World Economics** case studies, addressing a range of topical issues with accompanying exercises to develop the skill of application
- **Key terms**, highlighted and defined on the relevant pages for ease of reference
- **Number crunching**, for quantitative skills development
- **Review questions** at the end of every chapter, testing knowledge and understanding
- **Exam-style questions** at the end of each topic, with allocated marks, providing effective exam preparation
- Plus **Author tips**, **Key notes** and **Discussion points**

We hope that you enjoy using this book to support your study of AQA economics and that it helps you to prepare thoroughly for your examinations.

Malcolm Surridge & John Wolinski

Section 1

Individuals, firms, markets & market failure





Topic 1

Economic methodology & the economic problem

Key concepts from Year 1

Topic 1 represents an introduction to economics and specifically microeconomics. For this reason, all of the A-level specification for this topic was fully covered in the first five chapters of the Year 1 companion textbook.

- Chapter 1 indicated what economics was about and showed why it is classified as a social science. The methodology of economists was compared and contrasted to that used by natural and other scientists. The difference between positive and normative statements was outlined and their significance explained, particularly in terms of how value judgements influence economic decision making and policy. The chapter concluded by explaining the difference between positive and normative statements, showing how people's views relating to choosing between options are influenced by positive factors and also by moral and political judgements.
- Chapter 2 explained why economic activity is planned to satisfy needs and wants. It looked at the difficulties in measuring the satisfaction of wants. The second part of the chapter focused on three economic decisions – what to produce, how to produce and who is to benefit from the goods and services produced. The ways in which these three decisions are made were demonstrated in the context of a market economy, supported by observations relating to how a mixed economy might impact upon these economic decisions.
- Chapter 3 introduced the concept of economic resources or factors of production. The four factors of production – labour, land, capital and enterprise – were each considered in turn. Specific emphasis was placed on the notion of the environment as a scarce resource. The contrast between renewable and non-renewable resources was examined, particularly with regard to their impact on scarcity.
- In Chapter 4 the basic economic problem was studied through investigation of the significance of four basic economic ideas, namely that human wants are unlimited, wants vary in importance, means (resources) are limited, and that resources have alternative uses. These characteristics lead to the necessity of choice for consumers, but also in terms of choices relate to how scarce resources are allocated to different uses. The chapter concluded with an introduction to the concept of opportunity cost – the idea that all decisions involve a sacrifice of the next best alternative that is foregone.
- Chapter 5 investigated the concept of production possibility and the use of production possibility diagrams in economic analysis. The logic of the production possibility diagram and the shape of the production possibility boundary were explained. The chapter then examined the main uses of production possibility diagrams, demonstrating how they could be used to show resource allocation, opportunity cost and tradeoffs, unemployment of economic resources and economic growth. The chapter concluded with an introduction to economic efficiency in the form of allocative efficiency and productive efficiency, and showed how the latter could be illustrated on a production possibility diagram.



Topic 2

Individual economic decision making

Consumer behaviour

Key concepts from Year 1

This chapter builds on the concept of the rational consumer, introduced in Year 1. It links closely to Chapter 6 of the Year 1 companion textbook – the determinants of demand – by explaining the link between marginal utility and the individual demand curve.

This chapter examines the traditional view of rational economic decision making and economic incentives. Utility theory is then considered, looking at the difference between total utility and marginal utility and explaining the hypothesis of diminishing marginal utility. The chapter concludes with a study of utility maximisation and the importance of analysis of the margin when making choices.

Rational economic decision making

Economists often make assumptions when analysing situations, and investigating consumer behaviour is no exception. A critical assumption made by traditional economists is that consumers are rational. Each consumer will make decisions that will, as a whole, maximise their individual welfare. If this logic is applied to the market, then all consumers in the market act logically.

Economic man (or *homo economicus*) is the name given to the hypothetical individual who:

- behaves rationally. This means that economic man acts in a way that is consistent with his (or her) preferences. Neoclassical economics assumes that individuals have a fixed set of preferences when choosing their consumption of goods and services;
- has complete knowledge. In order to make rational decisions, perfect knowledge is required. It is also assumed that economic man can undertake the complex calculations required to assess and compare the relative value of the various wants that he wishes to satisfy;
- acts solely in self-interest. Decisions made by economic man are intended to satisfy his personal/household wants. No consideration is given to the impact of his decisions on other parties.
- aims to maximise personal utility/satisfaction/happiness. Economic man consumes the goods and services that allow him to enjoy the highest possible level of satisfaction or happiness.

Behavioural economists question the validity of the concept of economic man, arguing that the assumptions outlined above are not realistic in modern society. In particular, we will examine why they challenge the idea that man always acts rationally. In Chapters 3 and 4 we will examine the views of behavioural economists and their impact on economic behaviour and policy.

Key term

Rational economic decision making occurs when individuals compare the benefits and costs of alternative decisions and select the one that maximises their personal net benefit.

Key note

In his book *Principles of Economics* (1998), Harvard University economist Greg Mankiw identified ten economic principles. Four of these relate to ‘how people make decisions’:

- People face trade-offs.
- The cost of something is what you give up to get it.

(These ideas refer to the concepts of ‘choice’ and ‘opportunity cost’, introduced in Chapter 4 of the Year 1 companion textbook.)

- People respond to incentives. (This is explained immediately below.)
- Rational people think at the margin. (This is explained in the section on marginal utility later in this chapter.)

Economic incentives

The traditional view of consumer behaviour is that ‘people respond to incentives’. In this context, the word ‘incentives’ refers to both costs and benefits. If the price of a good rises, then people will buy less of it. Similarly, if the price of a good falls, then people will buy more of it. In this way, price acts as an incentive to buy less (or more) of a good.

A benefit describes the satisfaction that a consumer receives from a good. If a rational consumer believes that, in a new situation, a particular good provides them with more satisfaction, then they will be encouraged to buy more of it.

We saw in the first year how the market mechanism impacted upon consumer behaviour through its influence on the quantity demanded. Changes in the quantity demanded were brought about through price changes (*resultant* changes) or changes in consumers’ perceptions of the benefits of the good (*real* changes). These changes acted as incentives to modify the quantity consumed.

Utility theory: total and marginal utility

It is difficult to quantify happiness, although utility is an attempt by economists to do so. The term ‘util’ is sometimes used as a way of measuring utility or satisfaction. Although utility can be difficult to measure, the market mechanism uses price to decide the allocation of goods. This is based on the assumption that an individual who is prepared to pay a given price for a good must want that good more than an individual who is not prepared to pay that price. However, this approach ignores the importance of income and wealth. Nevertheless, the satisfaction that one individual gains from different goods can be assessed by the price that he or she is willing to pay for each good. This means that the relative utility of different goods to a given individual can be judged.

Key terms

Utility is a measure of the satisfaction or happiness gained from the consumption of a good or service.

Total utility is the aggregate sum of satisfaction or happiness that an individual gains from the consumption of a given amount of a good or service.

Marginal utility is the additional satisfaction or happiness gained from the consumption of one more unit of a good or service.

Economic incentives are costs or benefits that influence economic agents to act in a certain way. These incentives are often, but not necessarily, financial.

For a rational consumer, the more units of a good that a consumer buys, the less utility each additional unit provides for the consumer. For example, if someone is very thirsty, the first bottle of water gives a great deal of utility as it is needed to quench the thirst. A second bottle may also help, but the main thirst has been quenched and so it gives less utility than the first bottle and so on for subsequent bottles.

A rational consumer will only buy a product if the amount of utility it gives is equal to or greater than the price charged. In the instance above, the consumer may be prepared to pay £2 for the first bottle of water, as it provides a great deal of utility. The second bottle is worth less than the first bottle, say 50p, as it helps to quench the remaining thirst. The third bottle may be valued at 10p, as it provides some utility. If the fourth bottle provides no additional utility, then the consumer will value it at zero.

This logic applies to the individual's utility. However, as the UK market consists of about 60 million consumers, each making rational decisions, total utility for a good will follow the same pattern. As a consequence, at a high price fewer consumers will gain enough utility to make purchase worthwhile; at a lower price there will be more customers wanting to buy (demand) the good.

Table 1.1 shows how total and marginal utility may be quantified. The values are based on the choices made by the individual in the paragraph above. Since these choices were expressed in terms of a willingness to pay a given price, their utility (utils) is equated to the number of pence they were prepared to pay.

Table 1.1 Measure of marginal and total utility of bottles of water for an individual

Quantity consumed	Marginal utility (utils)	Total utility (utils)
0	–	0
1	200	200
2	50	250
3	10	260
4	0	260

The hypothesis (law) of diminishing marginal utility

Measuring marginal utility allows the demand curve of an individual to be derived. Table 1.1 indicates that at prices above £2 per bottle, there will be no demand for water. At £2, the individual gains sufficient utility to purchase one bottle. A second bottle of water will only be purchased if the price falls to 50p. Similarly a third bottle will be demanded if the price is 10p. If water is free, then a fourth bottle will be demanded.

The hypothesis of diminishing marginal utility supports the view that demand curves should be downward sloping. As quantity consumed of a good rises, each additional good provides the consumer with a diminished level of utility. It is therefore logical that a rational consumer will be prepared to pay a lower price for the second good than the first good, a lower price for the third good than the second good, and so on. This leads to a demand curve that slopes downwards from left to right (Figure 1.1).

Utility maximisation

A rational consumer will try to achieve as much happiness as possible. This means that they will try to maximise their utility. In Table 1.1 total utility from bottles of water has been maximised (at 260 utils) when consumption is 4 units. However, this does not mean that the consumer will buy 4 units. The decision to consume a good

Key term

The **hypothesis (or law) of diminishing marginal utility** states that, as consumption of a good increases, each additional unit of the good provides less utility than that provided by the previous unit.

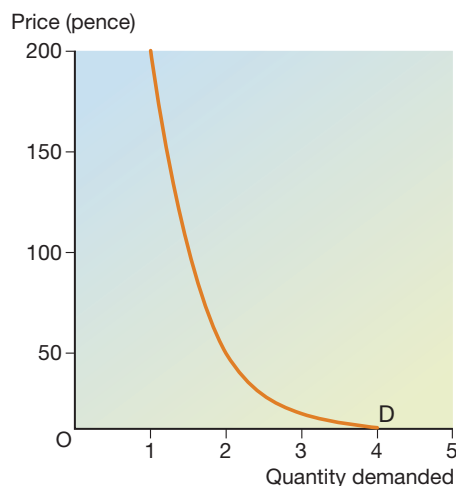


Figure 1.1 *Diminishing marginal utility and the individual's demand curve*

has an opportunity cost measured by the alternative foregone. The consumer has limited resources and has gained very little utility from the third and fourth bottles of water. Thus, unless the price is very low, it is probable that greater utility could be achieved by purchasing other goods.

It is not necessary to prove at A-level, but an individual consumer will maximise their utility by ensuring that the ratio between the additional (marginal) utility of a good and its price is the same for every good. Where MU_a is the marginal utility of good 'a' and P_a is the price of good 'a', a consumer's utility is maximised when:

$$\frac{MU_a}{P_a} = \frac{MU_b}{P_b} = \frac{MU_c}{P_c}, \text{ etc.}$$

This is known as the principle of equi-marginal utility.

The importance of the margin when making choices

The idea of maximisation is a common one in economic theory. For example, in the first year we saw that firms aim to maximise profits and factors of production aim to maximise their rewards. In this chapter we have seen that consumers wish to maximise their utility.

How can an individual consumer be sure that they are maximising their utility? This is where the importance of the margin can be shown. For each decision they make, such as the purchasing of a good, a consumer can measure the additional (marginal) benefit that they gain and compare it to the marginal cost of that decision. If the marginal benefit (MB) exceeds the marginal cost (MC), then their total utility has increased; if the marginal benefit (MB) is below the marginal cost (MC), then their total utility has decreased. Logically, a consumer will buy a good if $MB > MC$ and keep buying additional units until $MB = MC$. At this point total satisfaction is maximised and buying more of the good will cause total satisfaction to fall.

If we take the water purchases as our example and assume that each bottle of water costs 50p, the first bottle improves utility because $MB = 200p$ and $MC = 50p$. For the second bottle, $MB = 50p$ and $MC = 50p$; this is worthwhile because the benefit (just) matches the cost – overall total utility has not changed. For the third bottle, $MB = 10p$ and $MC = 50p$; this bottle should not be purchased since it reduces

overall satisfaction because the additional cost is higher than the additional benefit. Utility is maximised where $MB = MC$. This use of the margin when making choices will be examined in much more detail when studying topic areas such as profit maximisation, labour markets and market failure.

REALWORLD ECONOMICS 1.1

Airlines and the margin

Airlines frequently use the margin when selling 'last-minute' tickets. A typical plane has high fixed costs and low variable costs. As a rule they charge high prices to ensure that both fixed costs and variable costs are covered and profit can be made. One additional passenger adds little to total costs. Thus if a

seat is empty and unlikely to be filled, profit can be increased by charging a very low price provided the price covers the additional cost of flying the passenger.

Discussion points

1 Can you think of other examples of organisations/situations that might lead to this approach to pricing?

2 What is the danger of using this approach to pricing?

3 Ryanair uses a different approach. It charges lower prices initially to ensure that it sells all or most seats in its aircraft. Last-minute seats are sold at a higher price. Why might this be a logical approach for a firm aiming to maximise profit?

Review questions

Total: 20 marks

- 1** When making a decision, a rational consumer (economic man) will:
 - A Accept that their knowledge will be imperfect
 - B Aim to maximise individual utility
 - C Consider the impact of their decision on third parties
 - D Make the same purchasing decisions every time

(1 mark)
- 2** What is meant by the term 'rational decision making'? (3 marks)
- 3** Define the term 'total utility'. (3 marks)
- 4** The table below shows the total utility and marginal utility of baked beans to an individual.

Quantity consumed	Total utility (utils)	Marginal utility (utils)
0	0	–
1	20	20
2	35	(b)
3	(a)	9
4	51	7

- (a) Calculate the total utility of 3 units. (2 marks)
- (b) Calculate the marginal utility of the second unit. (2 marks)
- 5** Explain why a rational decision maker would respond to economic incentives. (4 marks)
- 6** Explain the importance of the margin to a consumer who is trying to maximise their utility. (5 marks)

Imperfect information

Key concepts from Year 1

The concept of imperfect information was introduced as a potential cause of market failure in Chapter 26 of the Year 1 companion textbook. This topic was developed in more detail in Chapter 27, with a particular focus on asymmetric information.

This brief chapter draws on the knowledge gained in Chapter 27 of the Year 1 companion textbook. You are advised to revisit that chapter if necessary. However, whereas the Year 1 book studied imperfect information and asymmetric information as causes of market failure, this chapter looks at the importance of information for decision making, and the impact of imperfect information and asymmetric information on decision making.

The importance of information for decision making

Perfect knowledge is a critical assumption in the model of perfect competition, introduced in Year 1 of the course. For individuals to make rational decisions, they need to be aware of all of the facts needed to make those decisions. In a perfect market, consumers are assumed to have perfect information on the goods being sold as well as on the price offered by all firms providing a good and its alternatives. Similarly, individual workers and owners of factors of production will have perfect knowledge/information relating to the rewards available for their services and the rewards that might be earned if the skills of the factor are improved, say by additional training or by differing combinations of factors of production.

In Year 1 we saw that imperfect information leads to market failure because individuals make decisions that do not have the outcomes expected. In effect, imperfect information can prevent individuals from making rational decisions. It can also cause market failure because government lacks the information needed to measure externalities and take action to overcome a misallocation of resources.

Since economic models, such as the operation of the market mechanism, rely on perfect information, the existence of imperfect information undermines the reliability of conclusions drawn from these models.

Key terms

Imperfect information occurs when a buyer or seller lacks the information needed to make the best choice in a transaction.

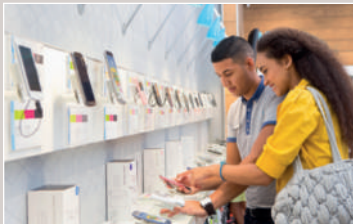
Symmetric information occurs when both the seller and the buyer are well informed about the goods and services and prices in the market.

Asymmetric information occurs when either the seller or the buyer has more information than the other party in a transaction.

REALWORLD ECONOMICS 2.1

Buying a mobile phone

Behavioural economists would question whether consumers can make a rational decision when purchasing a mobile phone, since the marketing skills of the main suppliers have created brand loyalty that can lead to bias in decision making. (This will be covered in Chapter 3.)



Buying a mobile phone – it's hard to make a rational decision

However, the level of information needed to make this decision may make it even harder to make a rational decision. In the past, consumers often had a lack of awareness of the technical details of a good until they received the user manual (often available only after the good had been purchased). This issue has largely been resolved by the use of the internet to make user manuals freely available. However, this can create so much information that it becomes impossible for an individual to make a rational decision (although economic man would be able to cope).

A consumer wishing to choose

between three leading models of mobile phone can access the relevant user manuals, but this would require a lot of reading:

- Apple iPhone – 196 pages
- Samsung Galaxy – 198 pages
- Sony Xperia – 115 pages

A consumer would need to read 509 pages of information just to find out about the features and use of these three phones. Such a task would also require an understanding of all of the technical details.

Discussion point

Is the greater availability of information improving the opportunity for consumers to make rational purchasing decisions?

It should also be noted that the provision of additional information is not necessarily going to lead to more rational decision making. The model of economic man assumes that individuals can process all of the information required to make rational decisions.

The significance of asymmetric information

The model of perfect competition assumes symmetric information. This requires both buyers and sellers to be equally well informed about the goods and services in the market. Consequently, both parties make rational decisions on whether to demand or supply the product at the market price.

Where asymmetric information exists, there is also likely to be market failure because one party to the decision (usually, but not always, the seller) possesses more information on which to base their decision than the other party (usually, but not always, the buyer). For example, the seller of a good is likely to have a better understanding of the technical details of the product than a typical consumer. Where asymmetric information exists, the person with the greatest knowledge will usually have greater bargaining power when it comes to deciding on the value of the good. This will prevent rational decision making by the individual with less information. Asymmetric information may even give monopoly power if it restricts choice and awareness of alternative options.

Asymmetric information may also lead to moral hazard, where a decision maker may take risky decisions because another party will have to bear the greatest burden of any costs arising from that decision. For example, people who spend excessively on credit and then apply for bankruptcy are passing the risks to another person.

Similarly, a business with no security systems may be passing the risk of theft on to the insurance company.

Conclusion

Perfect information is needed for rational decision making. Imperfect information and asymmetric information therefore undermine the concept of a consumer who is able to make rational, objective decisions. As products become more complicated, it could be argued that the provision of too much information is also undermining the ability of consumers to act in a rational way.

This challenges the notion of economic man – a concept that is also being questioned by behavioural economists, who are critical of the idea of a person who thinks objectively and has clear and consistent aims and preferences. These criticisms will be considered in Chapters 3 and 4.

Review questions

Total: 15 marks

- 1 Distinguish between imperfect information and asymmetric information. *(4 marks)*
- 2 Explain why perfect information is critical to the existence of economic man. *(4 marks)*
- 3 Explain why the provision of more information may not help consumers to make rational decisions. *(3 marks)*
- 4 Explain how asymmetric information can lead to moral hazard. *(4 marks)*

Aspects of behavioural economic theory

Key concepts from Year 1

In Year 1 we saw that the neoclassical view of economic analysis assumes that individuals are rational and behave in a way that will maximise their individual utility. Although this model allows for logical economic analysis of situations, it is based on assumptions that can be challenged and can lead to inappropriate economic policy recommendations.

Behavioural economic theory challenges the assumptions behind economic man. In this chapter we examine the key principles underpinning behavioural economic theory and the impact that these principles have on economic decision making by both individuals and organisations. The chapter will focus on specific aspects of behavioural economic theory: bounded rationality and bounded self-control; reasons for biases in decision making; and the importance of altruism and perceptions of fairness.

Principles of behavioural economics

Behavioural economics describes the discipline that tries to mix insights from both psychology and economics in order to look at economic problems through the eyes of normal humans, rather than through the eyes of the rational thinking, perfectly informed economic man. The following principles underpin the views of behavioural economists:

- 1 People's behaviour is influenced by other people.** Individuals' decisions are affected by watching and also copying other people; furthermore, individuals' decisions are often affected by the need for approval by others.
- 2 Self-expectations influence people's behaviour.** Each individual has their own set of values and responsibilities – these will influence their decisions in a given situation.
- 3 People want to do the 'right thing'.** Decisions are often influenced by ethical, rather than financial, factors. For example, about 30% of UK adults regularly donate money to charities.
- 4 People are averse to risk.** Many people are more inclined to take decisions to avoid losses rather than to make gains.
- 5 Decision making is often based on habits.** Certain decisions become routine. In these cases, if circumstances change, individuals find it difficult to change to a more rational decision.
- 6 People's decision making is often influenced by weak computational skills.** Rational decisions may not be taken because many individuals are unable to calculate the

possible value of outcomes. This may also arise from a lack of information.

- 7 **People need to believe that their decisions will be effective.** People may not make decisions if they do not believe they can influence the final outcome.

Bounded rationality and bounded self-control

The idea of bounded rationality was an early challenge to the traditional, neoclassical view of economics. In the 1950s, Herbert Simon took the view that there were three main restrictions on rational decision making:

- the individual's inability to process and evaluate information;
- the limited time available for many decisions; and
- the lack of information available.

Each of these restrictions might prevent or limit the scope for rational decisions. Simon concluded that individuals rely on rough 'rule of thumb' notions when trying to make complicated decisions. Invariably this leads to biases in decisions and a tendency for 'satisficing' – accepting an outcome that does not maximise utility but does provide a reasonable level of satisfaction. Similarly, firms might target a reasonable level of profit rather than aiming to achieve profit maximisation.

Many firms recognise that their customers' decisions are limited by bounded rationality and they therefore present a limited range of information on their products, hoping to focus on the key factors that customers want to know. In this way, customers can process the information more effectively (although there is a possibility that an irrational choice may be made if the firm omits important information).

Bounded self-control means that there are limitations to the level of self-control for most individuals. It often takes the form of decisions that lead to short-term utility at the expense of long-term utility, such as binge drinking.

Another example of bounded self-control can be seen in 'all-inclusive' holiday packages. Customers tend to overconsume both food and alcohol on these packages, as the additional cost of extra food and alcohol is zero. This can have negative consequences for health. Furthermore, it can mean that expenditure on the holiday is much greater than an alternative option, such as self-catering, because the latter type of holiday may lead to a more normal and sensible level of consumption of food and alcohol.

These limits to self-control contradict the traditional hypothesis of diminishing marginal utility. As people consume more of a good, each additional unit gives a lower (diminished) level of utility. However, in cases of bounded self-control, individuals tend to overestimate the value of additional units. For example, excessive food and drink can yield negative utility and so, even if the price is zero, a rational consumer should stop consuming a good that provides negative utility. However,

Key terms

Bounded rationality means that an individual's ability to make a rational decision is restricted by factors such as the individual's inability to process and evaluate

information, limited time in which to make the decision, and imperfect information.

Bounded self-control means that individuals may not make rational decisions, even where they are aware of their irrational actions, because they lack self-control.

people continue to consume the goods because they are free. Gambling is another instance of bounded self-control, where people who suffer losses may be prepared to gamble even more in order to recoup those losses.

Biases in decision making

There are many potential causes of bias in decision making. The four reasons for bias that are described below are those listed in the AQA A-level economics specification.

Rules of thumb

These rules provide mental shortcuts which help to simplify day-to-day economic decision making and can save consumers considerable time. Examples include the purchase of goods on 'special offer' because these must be cheaper than those advertised at the normal price, or buying the latest version of a product because it must be technically superior to earlier versions. People instinctively want to make quick decisions and rules of thumb help to speed up the decision-making process. In many cases, these rules assist rational decision making because they are based on relevant experience, but they can lead to irrational decisions if the mental shortcut means that certain vital information is not incorporated into the decision. In his 2011 book *Thinking, Fast and Slow*, Daniel Kahneman contrasted two approaches to individual decision making:

- **System 1 – Thinking Fast:** This approach features shortcuts in order to reduce the time taken to make a decision.
- **System 2 – Thinking Slow:** This is a more rational method that considers as much evidence as possible in order to reach the best decision.

In general, individuals are more likely to use System 1 for everyday decisions, such as buying groceries, and System 2 for more significant decisions, such as taking a new job or buying a new car.

Anchoring

Anchoring is often used by shops. If a person is told that the 'recommended retail price' is £50, then this provides an 'anchor' – in this case its potential value. If the shop price is £25, then the person is more likely to see this as a good choice because its price is only 50% of its 'value'. The same good just offered at £25, with no other reference, will seem to be less attractive. The use of anchoring in this way is so widespread that the Competition and Markets Authority (CMA) has established certain rules to limit the use of anchoring by shops because it is seen to be a way of encouraging consumers to make non-rational decisions.

Key terms

Biases in decision making describe a situation in which an individual has a mental tendency or inclination towards a particular preference. Consequently, any decisions influenced by

these biases are unlikely to be purely rational, even though the individual might believe that they are weighing up the pros and cons of the decision.

Rules of thumb (often known as heuristics) are general principles to help an individual make decisions, where the principles are based on experience or practice rather than calculation or evidence.

Anchoring is a bias caused by individuals relying too much on an individual piece of information when making a decision.

Experiments have shown that there is no need to use logical numbers. A group of people was asked whether Gandhi had lived to over 114 years of age. The same group was then asked the age when he died. This group gave a much higher average age than another group that had initially been asked whether he lived more than 35 years. This corresponds to Kahneman's view that people adjust away from the initial anchor, but tend to adjust insufficiently. This results in the final estimate being closer to the anchor than it would have been had no anchor been provided. Kahneman found that any anchor, even where it was very unlikely, created a bias in people's views. Thus, providing an anchor can be used as a way of pulling people towards a certain opinion.

Availability

In general, this bias tends to mean that a recent occurrence will have a disproportionate influence on decisions. For example, a recent period of snow will lead to decisions that are based on an overestimation of the likelihood of snow recurring. Each time there is heavy snowfall there is a clamour for more investment in snow ploughs. However, evidence from the Met Office indicates that, over the last 20 years, the majority of UK winters have been milder than the long-term average.

Often 'availability' is influenced by the information that is readily accessible to the individual. For pricing, this might be the most recent price paid or an easily accessible reference price, such as the Argos website. This bias can mean that decisions are taken with only a limited view of the situation.

Social norms

Individuals who do not follow the social norms of behaviour may be shunned by the group or society. As a consequence, the decisions of an individual will be influenced by society and, more significantly, by the groups with which they associate.

Fashion items are an example of where social norms can cause a bias in decision making. Many groups can be identified with certain products or fashion styles and members of a group will therefore be more inclined to buy the products associated with the group. Actions and behaviour can also be influenced by social norms. Psychological research suggests that the desire to conform to social norms and follow a particular trend often overrides more rational decision-making processes about the wisdom of joining the group.

Society as a whole can use laws to enforce certain rules of behaviour. Legislation on non-smoking areas, opening hours for shops on Sundays, the use of seat belts and indecency laws are all examples of where an individual's decisions will be influenced by society's norms.

Availability bias occurs because when people make judgements about the probability of an event happening, they are heavily influenced by situations that they can remember.

Social norms are rules of behaviour that are considered acceptable within a group or society.

Altruism is the disinterested and selfless concern for the wellbeing of others.

The importance of altruism and perceptions of fairness

Altruism directly contradicts economists' traditional view of self-interested, economic man. Altruism means that individuals will make decisions that do not serve their self-interest. For example, voluntary work is a common feature of most economies. Any 'volunteer' is not going to be able to maximise utility in the traditional sense. However, often the individual will derive considerable personal satisfaction from their actions and the favourable impact that those actions have on other people.

Some studies have indicated that some services are best provided through altruistic behaviour. For example, countries such as the UK rely on voluntary giving of blood for medical purposes. Studies in the UK and New Zealand in fact indicate that offering payment would put off some donors. The World Health Organisation (WHO) also recommends voluntary blood donation – partly for altruistic reasons but mainly for safety. It believes that 'The safest blood donors are voluntary, non-remunerated donors from low-risk populations'.

People's sense of **fairness** can also lead to apparently irrational decision making. Role-playing exercises in which a 'dictator' has to decide on how much of a cake to give to a subordinate invariably leads to examples of fairness (although the largest piece is kept by the dictator). Self-interested, economic man would be expected to keep all of the cake, but human instinct suggests that most people have some sense of fairness.

REALWORLD ECONOMICS 3.1

The behaviour of taxi drivers

It is assumed that the demand for taxis is higher when it is raining. Traditional economic theory suggests that it should be easier to find taxi drivers at busy times (such as when it is raining) because this is the time in which they can earn most money. Self-employed taxi drivers can decide their working



Finding a taxi when it's raining – harder than you might think

hours and so a rational taxi driver will work longer hours on busy/rainy days.

A 1997 study by Colin Camerer et al. suggested that this was not the case. They analysed taxi trips in New York and found that as wages went up, the supply of taxi drivers went down. Further investigation led to the conclusion that most taxi drivers set a target income for each day. On busy days the taxi was constantly occupied and so the driver earned money more quickly. From a behavioural perspective, this situation arose from a bias in their decision making. The target income served as an anchor and, once this target had been reached, the driver concluded that the day's work had yielded a successful outcome. In contrast, less busy days meant that the taxi driver worked longer hours.

This outcome is not consistent with economic man, who would maximise his income on the good days. Economic man would also decide that on days with few customers, the benefits would be insufficient to continue working.

Improved information services may well change this situation. Organisations such as Uber allow taxi drivers to adjust pricing according to demand. This improved information should encourage more drivers to offer rides during busy and rainy times – economic man may provide the service that customers need.

Discussion point

Why do neoclassical economists and behavioural economists disagree on how taxi drivers will behave on busy days?

Research by the United Nations indicates that fairness is a major influence on national happiness. Studies show that while an individual's happiness is influenced by level of income, a more significant influence is the individual's income in comparison to other people in the country. Countries with greater equality tend to have the highest levels of happiness.

Review questions

Total: 30 marks

- 1 Which *one* of the following statements is a principle of behavioural economics?
 - A People are not influenced by other people
 - B People are rational
 - C People base decision making on habit
 - D People enjoy taking risks

(1 mark)
- 2 Which *one* of the following statements is an assumption of neoclassical economics?
 - A People aim to maximise utility
 - B People have weak computational skills
 - C People want to do the right thing
 - D People's decisions are influenced by rules of thumb

(1 mark)

The following answers apply to Questions 3, 4 and 5:

- A Anchoring
 - B Availability
 - C Rules of thumb
 - D Social norms
- 3 Bias which occurs because people are heavily influenced by certain situations or experiences is known as:

(1 mark)
 - 4 Bias which occurs because there is overreliance on a single piece of data is known as:

(1 mark)
 - 5 Bias which occurs because the individual wants their decision to be accepted by the group is known as:

(1 mark)
 - 6 Herbert Simon stated that *bounded rationality* describes a situation in which an individual's ability to make rational decisions is restricted by three particular factors. State these three factors.

(3 marks)
 - 7 Define the term 'bounded self-control'.

(3 marks)
 - 8 Explain why bounded rationality and bounded self-control prevent rational decision making.

(6 marks)
 - 9 Explain the difference between 'altruism' and 'fairness'.

(5 marks)
 - 10 Analyse *two* reasons why bias can cause irrational decision making.

(8 marks)

Behavioural economics & economic policy

Key concepts from Year 1

This chapter examines how behavioural economics is challenging the traditional view of rational, economic man.

This chapter builds on Chapter 3, focusing on how behavioural economics shapes economic decision making. There is some focus on its use by firms, but the main focus is on how behavioural economics contributes to government policies, notably in the area of government intervention to overcome market failure. The concepts of choice architecture, framing and nudge are explained. The chapter concludes with an examination of how default choices, restricted choice and mandated choice can be used in the design of government policies.

Choice architecture and framing

The context in which a choice is made can be influenced by an organisation, such as a firm or government, and the ways in which this context is designed (by a choice architect) can have a major influence on the final decision. For example, many firms use telephone calls with the promise of a special offer, on condition that the individual makes an immediate decision. The design of this option is intended to encourage individuals to make an immediate decision based on limited information, as they will not have had enough time to research alternative choices.

In the UK, the DVLA will not give a driving licence until an applicant has answered questions on organ donation. The questionnaire then encourages the applicant to register for organ donation. Since 1995, over 9 million people have signed up for organ donation as a result of this scheme.

Choice architecture can take many forms. Examples include:

- reducing choice overload. This occurs when individuals cannot efficiently process all the information needed to make a choice. In this instance, choice architecture can take the form of limiting the number of alternative options or providing tools to support or advise the decision maker;
- overcoming bias against long-term factors. Individuals tend to underestimate the importance of long-term factors when making decisions. An example is the tendency for individuals to avoid planning their pensions. Government policy initially tackled this bias by using taxation to fund a state pension. More recently, government has used legislation to force businesses to provide a pension scheme for their employees;
- using measures that are easily understood – for example, comparing energy suppliers in terms of monthly bills rather than kilowatt hours of energy;

Key term

Choice architecture is the term used to describe the different ways in which choices can be presented to consumers or individuals.

- providing default decisions. (These will be covered in the final section of this chapter.)

Framing is where choices are presented in a particular way in order to elicit a particular response. For example, when hospital patients are asked if they wish to have surgery, they are more likely to agree to surgery if they are told that there is a 90% chance of survival than if they are told that the mortality rate is 10%.

Nudges

If an individual is free to choose but has been subjected to some form of influence, then this is termed a nudge. For example, individuals can still buy cigarettes, but age restrictions have made them less accessible. Furthermore, the designs of packets are controlled to reduce their attraction to consumers. Research has shown that images are more likely to stop people smoking than words. From 2016 the EU is introducing a regulation that the top 65% of both sides of a cigarette packet must feature images and warnings designed to prevent consumption.

Overall, nudges are used to guide people towards the ‘best’ decision. In order to discourage spillage in men’s urinals, one company makes stickers of a fly that are placed in the centre of the urinals. This reduces the level of spillage as users of the urinal focus their aim on the fly.

Choice architecture and nudges are criticised because they are seen to interfere with the individual’s freedom of choice. However, government policy is to use them if the outcome is desirable. In imperfect markets, rational decisions are not possible anyway since producers are restricting information available to consumers.

Default choices, restricted choice and mandated choice

A default choice is when a decision is made that requires no action on the part of the individual. For example, for most internet browsers, Google is the default search engine.

Individuals’ choices can be influenced by restricting the choices available. For example, the DVLA does not allow a person applying for a driving licence to say ‘No’ to organ transplant. The options are:

- Yes, I would like to register (as a donor)
- I do not wish to answer this question now
- I am already registered on the NHS Organ Donor Register

Applicants can only select one of these three options. The government is using nudges to gain the desired outcome. (As explained earlier, ‘Yes’ is more likely to be chosen and, furthermore, ‘No’ is not presented as an option, as such.)

Key terms

Framing occurs when the choices are presented in a way that is intended to encourage a given response.

Nudges are ways of influencing individuals’ choices in a particular direction, but without removing their freedom of choice.

Default choice is the decision that is made if an individual takes no action.

Restricted choice is when individuals can only select from a limited range of options.

Mandated choice occurs when people are required by law to make a certain decision.

When people are required by law to make a certain decision, this is a mandated choice. For example, in the UK car owners are required by law to take out insurance that covers injury and damage to other road users (typically insured through third party insurance).

Usually, mandated choice is used when the government deems it to be in the public's interest. For third party insurance, it provides innocent victims of car accidents with compensation. In times of war, conscription is a mandatory requirement for certain individuals to offer their services to the armed forces.

REALWORLD ECONOMICS 4.1

UK government nudges

Since 2010 government has adopted many policies based on theories of behavioural economics.



Choice architecture in a school canteen

Nudges have been particularly popular.

Richard Thaler and Cass Sunstein, who introduced the nudge concept, argued that choice architecture can be used to present 'better' choices more favourably, so that individuals are more likely to make 'good' decisions.

In school canteens, the food offered can have a major bearing on pupils' diets. However, more subtle uses of choice architecture can also influence pupils' decisions.

The placement of food can be used to encourage or discourage consumption, with food at eye level more likely to be chosen for consumption. The layout of the menu can also affect decisions – foods displayed at the beginning or end of a list of options are more likely to be chosen than those in the middle.

Discussion point

Should school canteens deliberately try to influence pupils' food choices?

Review questions

Total: 15 marks

Answers A to D below apply to Questions 1 to 4:

- A Influencing choices but without removing freedom of choice
- B Presenting options in a certain way in order to encourage a given response
- C The different ways in which options can be presented
- D Where individuals can only select from a limited range of options

- 1 Choice architecture means: (1 mark)
- 2 Framing means: (1 mark)
- 3 Nudges describe: (1 mark)
- 4 Restricted choice describes: (1 mark)
- 5 Explain the difference between default choice and mandated choice. (5 marks)
- 6 Explain two ways in which choice architecture might improve an individual's decision making. (6 marks)

Topic 2 Exam-style questions

A-LEVEL PAPER 1

SECTION A Context – Individual economic decision making

Extract A Data on an individual's utility from food

Units of food consumed	0	1	2	3	4	5	6	7	8	9	10
Total utility (£)	0	30	50	62	69	74	77	79	80	79	77

Extract B The Competition and Markets Authority (CMA)

The Competition and Markets Authority (CMA) is a government organisation that recognises the importance of behavioural economics. Its role is to ensure that markets work well. Traditionally, the CMA and its predecessors – the Competition Commission and Office of Fair Trading (OFT) – focused on supply-side blockages that have led to inefficient operation of markets. These blockages are based on traditional, neoclassical economic thinking and include factors such as market concentration, supplier behaviour and barriers to entry. Typically, the CMA would intervene if a firm abused its market power. However, in recent years behavioural economists have begun to influence the thinking of the CMA so that it now intervenes in markets in which there are demand-side blockages. Common demand-side blockages include market failure caused by consumers finding it difficult to access information to make decisions; their inability to process and interpret the data available; and difficulties in acting on information in a non-biased way. The table below contrasts these two different approaches.

Supply-side blockages (classical view)	Demand-side blockages (behavioural view)
Market concentration	Accessing information
Supplier behaviour, such as cartels	Assessing information
Barriers to entry	Acting on information

Extract C The gym market and the CMA

The CMA has been involved in improving the operation of markets within the gym market using both neoclassical and behavioural economic approaches. In 2013, the CMA's predecessor – the OFT – investigated demand-side blockages that made it difficult for consumers to make rational decisions on gym membership. This investigation was triggered by information that showed that there were 4.5 million adults with gym membership, but only 27% used the gym regularly. In total it was estimated that £37 million per annum was wasted on gym membership in the UK.

January is the peak month for opening gym membership, with enrolments 50% higher than most other months. This arises from a mix of excessive indulgence at Christmas and New Year's resolutions that will invariably be doomed to fail. However, the waste is compounded by the fact that most contracts for gym memberships last for a minimum of 12 months. (Before the investigation, LA Fitness offered 24-month contracts.) Furthermore, many contracts allowed for significant price increases, which were outside the control of the consumer once the contract had been signed.

The OFT ruled that gym membership contracts should be deemed to be unfair if they have lasted more than 12 months, shown big price increases or there has been a major change in the service provided, such as reductions in equipment and opening hours. It also ruled that automatic renewal of contracts by default, without the customers' agreement, is unfair.

In 2015, there have been more traditional interventions by the CMA. In February 2015, two low-cost fitness gyms – Pure Gym and the Gym Group – agreed to merge. However, the CMA investigation recommended that the merger should be refused, even though the two gyms had a combined market share of only 2%. The CMA ruled that they had a large market share of the emerging low-cost gym market, which was vital in preventing other gyms from exploiting their customers.

In August 2015, the CMA agreed to a takeover by Pure Gym of LA Fitness (an upmarket gym). Although this meant a much larger market share for Pure Gym, the CMA did not prevent the takeover because Pure Gym and LA Fitness cater for different customers, mostly in different towns. The takeover would not therefore have a detrimental effect on the market's competitiveness.

Questions

Total: 40 marks

- 1 Use the information in Extract A to calculate the number of units of food consumed by a rational individual on an all-inclusive holiday, where additional units of food are free. (2 marks)
- 2 Explain why the CMA is now recognising that default choices can explain some cases of market failure in the gym market. (4 marks)
- 3 Using the data in Extract A, explain, using a diagram, why the hypothesis of diminishing marginal utility can be used to explain the downward-sloping demand curve. (9 marks)
- 4 Assess the view that it is vital for the CMA to understand and use both traditional, neoclassical economic analysis and behavioural economics in order to make markets operate fairly. (25 marks)

SECTION B Essays

Total: 40 marks

Behavioural economists recognise that imperfect information makes it difficult for economic agents to make rational decisions. However, although it is a reason why consumers do not act in the rational way assumed by neoclassical economists, other factors, such as bias in decision making, can have a significant impact too. Behavioural economists recognise that there are many reasons why individuals do not make rational decisions.

- 1 Explain possible causes of bias that might prevent an economic agent, such as an individual, from making rational decisions. (15 marks)
- 2 Assess the view that imperfect information is the main reason why consumers do not make rational decisions when purchasing goods for consumption. (25 marks)



Topic 3

Price determination in a competitive market

Key concepts from Year 1

Topic 3 introduces the concept of a competitive market in which resources are allocated through the interaction of demand and supply. Since demand and supply are essential foundations for analysing microeconomic topics, all of the A-level specification for this topic – ‘Price determination in a competitive market’ – was fully covered in Chapters 6 to 11 of the Year 1 companion textbook.

- Chapter 6 introduced the concept of demand and investigated the factors that influenced the demand for goods and services, focusing on price, income, wealth, the price of substitutes and complements, and individual preferences.
- Chapter 7 introduced the concepts of price, income and cross elasticity of demand and showed how to calculate each of these elasticities. The relationship between income elasticity of demand and normal and inferior goods was studied, before moving on to an explanation of the link between cross elasticity of demand and substitute and complementary goods. Price elasticity of demand and its influence on firms’ total revenue and consumers’ total expenditure were then considered. The chapter concluded with an explanation of the factors that influenced each of these elasticities of demand and an interpretation of the meaning of the numerical values of these elasticities of demand.
- Chapter 8 introduced the concept of supply and investigated the factors that determine the supply of goods and services, such as price, costs of production and changes in technology. The supply curve was introduced and its shape explained, through recognising that suppliers are motivated by profit and so higher prices provide an incentive to expand production. The chapter concluded by looking at the factors that cause shifts in the supply curve.
- Price elasticity of supply was explained in Chapter 9, including how to calculate its value. Diagrams were used to show the visual appearance of supply curves with different price elasticities. The chapter concluded with an explanation of the factors that influence price elasticity of supply and an interpretation of the meaning of the numerical values of different price elasticities of supply.
- Chapter 10 brought demand and supply together, looking at how equilibrium prices and quantity/output were determined in competitive markets. The effect of both changes in demand and supply were examined and consideration was given to situations in which both supply and demand changed simultaneously. The effects of different price elasticities of supply and demand were also considered, when examining shifts in supply or demand. The chapter concluded by studying supply and demand in some real-world markets.
- Chapter 11 examined situations in which changes to equilibrium price and quantity in a particular market affected equilibrium price and quantity in different but interrelated markets. The interrelationships studied were joint demand, demand for substitute goods, composite demand, derived demand and joint supply.



Topic 4

Production, costs & revenue

Key concepts from Year 1

In Year 1, the topic entitled 'Production, costs & revenue' was introduced in Chapters 12 to 16 of the Year 1 companion textbook. Chapters 12 and 13 provided an introduction to certain key concepts:

- Chapter 12 of the Year 1 book focused on production and productivity. It studied the production process – the conversion of inputs into outputs of goods and services. The concept of productivity was introduced, with a specific focus on labour productivity – its meaning, factors that influence it and the benefits of high labour productivity.
- Chapter 13 dealt with the concepts of specialisation, division of labour and exchange. The chapter examined the extent to which these concepts are beneficial to individuals, firms and countries. It concluded by looking at why specialisation requires exchange to take place and the necessity of money as a means of exchange.

Since these concepts provide the foundation for the next few chapters, you are advised to check that you are familiar with the ideas in Chapters 12 and 13 of the Year 1 companion textbook.

The law of diminishing returns & returns to scale

Key concepts from Year 1

As indicated in the introduction to this topic area, you should be familiar with production, productivity, specialisation and division of labour.

All of the material in this chapter is new. However, it does provide the basic understanding needed to explain the relationship between inputs and outputs and the consequent impact of this relationship on the shapes of cost curves, such as average costs, which were introduced in Year 1.

This chapter explains the difference between the short run and the long run. It then considers the differences between marginal, average and total returns. The law of diminishing returns is examined. The chapter concludes by studying the concept of returns to scale and explaining the difference between increasing, constant and decreasing returns to scale.

The difference between the short run and the long run

The relationship between inputs and outputs varies over time. For this reason it is vital to define time periods before examining the nature of output/returns.

The quantity used of some inputs (factors of production) cannot easily be changed. For example, if a firm wants to increase output, it may take a long time to acquire and develop land so that it is suitable for production of the firm's particular goods or services. Other factors, such as capital, can be difficult to put into productive use quickly.

However, some inputs (factors of production) can be changed quickly, if the firm wants to increase output. Raw materials or stocks can often be acquired quickly. Labour may also be very flexible, as workers may work additional hours to increase production, if necessary. It may also be possible to recruit some labour (usually unskilled labour) at short notice, in order to expand output rapidly.

Economists use inputs (the factors of production) to define time periods: the **short run** is the time period in which the amount of *variable factors of production* can vary but in which *fixed factors of production* are fixed. Based on this logic, the **long run** is the time period in which the amount of **both variable factors of production** and *fixed factors of production* can vary. These time periods vary between industries. For example, an oil company may take years to discover, extract and refine more oil. In contrast, a firm providing gardening services may be able to increase all of its factors of production in a very short time.

Key terms

The **short run** is the time period in which it is only possible to change the level of input of variable factors of production.

The **long run** is the time period in which it is possible to change the level of input of all of the factors of production.

The difference between marginal, average and total returns

Total returns are usually measured by the physical level of output produced, such as 57 cars or 320 packets of tea. For this reason, total returns are often referred to as total product or total physical product.

Average returns are found by dividing the total returns by the quantity of input. Input can be measured by a combination of factors of production, but it is common practice to measure the input of a specific factor of production, such as labour.

The calculation of average returns, based on input of labour, is:

$$\frac{\text{Total returns (or total physical product)}}{\text{Units of labour}^*}$$

Note: * In this example, ‘units of labour’ is the quantity of workers. However, the input of labour can also be measured by other methods, notably ‘man hours’ – the number of hours worked.

If total returns are 57 cars and the number of workers is 19, then average returns = $57/19 = 3$ cars per worker.

Marginal returns are the additional output from one extra worker (or unit of input). Based on the above example, if 20 workers produce 61 cars, then the marginal output of the twentieth worker is 4 units ($61 - 57$).

The marginal output of the n^{th} worker is calculated as follows:

Total returns from ‘ n ’ workers *minus* total returns from ‘ $n-1$ ’ workers.

Units of labour	Total returns	Average returns	Marginal returns*
0	0	–	–
1	4	4	4
2	11	5.5	7
3	19.5	6.5	8.5
4	29	7.25	9.5
5	39	7.8	10
6	48	8	9
7	56	8	8
8	61	7.63	5
9	63	7	2
10	63.5	6.35	0.5
11	63.5	5.77	0

Note: * Marginal returns represent the change in input from one quantity of input to the next. Therefore, on a graph, the marginal return of the first unit of labour should be plotted halfway between 0 and 1 unit of labour. Similarly, the marginal return of the fifth unit of labour should be plotted midway between 4 and 5 units of labour, and so on. However, this mathematical precision is generally ignored for the sake of convenience.

Table 5.1 Total returns, average returns and marginal returns

Key terms

Total returns are the quantity of output produced by a given quantity of inputs over a period of time.

Average returns are the quantity of output produced per unit of input.

Marginal returns are the additional quantity of output produced by one extra unit of input.

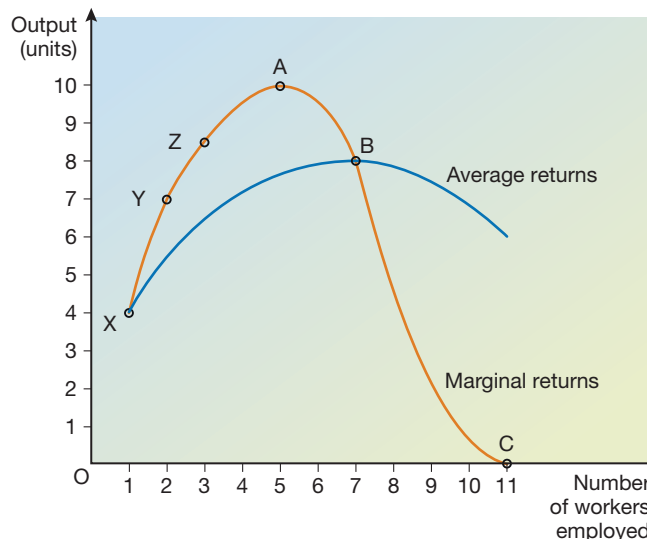
Table 5.1 shows an example of the relationship between total returns, average returns and marginal returns. It is based on the assumption that the firm is operating in the short run. In the short run we assume that the quantity of capital is fixed but that labour is variable because new or temporary workers can be recruited and existing workers can vary their hours of input.

The law of diminishing returns

If increasing quantities of a variable factor of production (input) are combined with a fixed level of input of other factors of production, then eventually the **marginal return** will diminish.

In Table 5.1, labour is the variable factor of production and capital is the fixed factor of production. The law of diminishing returns applies to both marginal returns and average returns. For very low levels of input of factors of production, both the marginal returns and average returns increase as additional variable factors of production are added. However, there comes a point where additional variable factors of production lead to a lower (diminishing) marginal return. Subsequently, there is also a lower (diminishing) average return that commences after the marginal return begins to decline.

Figure 5.1 The law of diminishing returns



In Table 5.1, marginal returns peak at 10 units from the fifth worker. Average returns peak when 6 and 7 units of labour are used.

Figure 5.1 shows marginal returns and average returns graphically. The 'y' axis measures returns (output) and the 'x' axis measures the variable factor of production (labour). It is assumed that other factors such as land and capital are fixed.

The marginal returns line shows the addition to total output (production) from each worker. The first worker only produces 4 units (Point X). The second worker adds 7 units of output (Point Y) and the third worker has a marginal return of 8.5 units (Point Z). At low levels of input, these increases arise from a more efficient use of factors of production. Usually it is impossible for one or two workers to carry out every job efficiently and time is wasted because workers are likely to be constantly moving between different tasks. However, after a certain level of units of labour,

there is an efficient balance between the quantity of fixed factors, such as capital and land, and the variable factor (labour). This enables the firm to be very efficient. In this example, this occurs when about 5 to 7 workers are employed, with marginal returns peaking at 10 units from the fifth worker and average returns peaking at 8 units per worker from 6 to 7 units of labour.

If more and more labour is added, the situation changes because there is then insufficient capital and land for the workers to operate efficiently. Therefore, extra workers do not increase production so much. This is known as diminishing marginal returns. In Figure 5.1, this occurs at Point A. The fifth worker causes output to increase by 10 units, but subsequent workers have lower marginal returns. Point B marks the point at which average returns (total returns/number of workers) starts to diminish, having peaked at 8 units per worker. As more workers are added, both marginal returns and average returns continue to diminish. Point C shows that the eleventh worker adds zero units of output. Even if the workers are not paid, it would not be worthwhile employing more of them if the firm is aiming to maximise profit.

The law of diminishing returns only applies in the short run. This law arises because it is inefficient to keep adding more workers if there is no additional space or capital equipment to facilitate their work. In the long run, a firm faced with diminishing returns from its variable factors of production would be able to increase its quantity of 'fixed' factors of production so that there is a more efficient balance of inputs.

Returns to scale

Returns to scale only apply in the long run. In the short run, only the quantity of variable factors of production, such as raw materials, can be changed. These short-run changes are likely to lead to an inefficient balance between the fixed and the variable factors of production. Since it is impossible to vary all factors of production, a firm wishing to increase its output by 10% in the short run might need to increase its labour force by, say 25%, because it cannot vary the capital that it uses. In the long run it is possible to increase all factors of production. If we assume that there is an optimum balance between labour and capital in the production process, then the long-run strategy of a firm would be to increase each factor of production by the same percentage. Thus, to achieve a 10% increase in output in the long run the firm is likely to increase the quantity of all factors of production by, say, 10%.

Author tip

It is not usual practice to calculate returns to scale, but it is the percentage (or proportionate) change in output divided by the percentage (or proportionate) change in input.

The difference between increasing, constant and decreasing returns to scale

Returns to scale can show increasing returns to scale, constant returns to scale or decreasing returns to scale.

Key terms

Returns to scale is the proportionate (or percentage) change in output of a firm or industry resulting from a proportionate (or percentage) increase in all inputs.

Increasing returns to scale occur if the percentage (or proportionate) increase in output is greater than the percentage (or proportionate) increase in input.

Constant returns to scale occur if the percentage (or proportionate) increase in output is equal to the percentage (or proportionate) increase in input.

Decreasing returns to scale occur if the percentage (or proportionate) change in output is less than the percentage (or proportionate) increase in input.

Illustrating returns to scale

If the quantity of output rises by 20% in response to a 10% increase in all inputs (factors of production), then the production process is achieving increasing returns to scale. As scale increases, it is quite common for greater efficiency to be achieved because different internal economies of scale may occur.

Decreasing returns to scale occur if the production process becomes less efficient as production is increased. This can arise because the firm becomes too large to be coordinated and managed effectively or communication breakdowns become more commonplace. These diseconomies of scale can lead to lower (decreasing) returns to scale.

It is possible that the increase in inputs has no impact on efficiency – the percentage growth in output exactly matches the percentage growth in input. This situation represents constant returns to scale.

These different categories of return to scale can be shown using the formula:

$$\frac{\text{Percentage (or proportionate) change in output}}{\text{Percentage (or proportionate) change in input}}$$

- If this calculation is > 1.0, we have increasing returns to scale.
- If this calculation is = 1.0, we have constant returns to scale.
- If this calculation is < 1.0, we have decreasing returns to scale.

Table 5.2 Returns to scale

Units of input	Total output (total returns)	Ratio of outputs to inputs	Returns to scale
0	0	–	
10	4	0.4	N/A
20	12	0.6	Increasing
30	30	1.0	Increasing
40	52	1.3	Increasing
50	75	1.5	Increasing
60	96	1.6	Increasing
70	112	1.6	Constant
80	128	1.6	Constant
90	135	1.5	Decreasing
100	140	1.4	Decreasing
110	143	1.3	Decreasing

Table 5.2 provides an example of changing returns to scale. Typically, at low levels of input and output the firm is unlikely to benefit from mass production, division of labour, specialisation and internal economies of scale. Consequently, production is inefficient and the ratio of output to input is relatively low. As the scale of the firm increases (from 10 units of input to 60 units of input), the firm experiences increasing returns to scale as the larger size of the firm enables it to benefit from the advantages of division of labour, economies of scale, etc. Between 60 and 80 units of input, constant returns to scale are experienced, suggesting that diseconomies of scale are counterbalancing any economies of scale. Above 80 units of input, decreasing returns to scale occur, suggesting that diseconomies of scale outweigh any economies.

Conclusion

The law of diminishing returns and the concept of returns to scale both have implications for costs of production:

- Periods of diminishing returns (in the short run) and decreasing returns (in the long run) mean that the firm is becoming less efficient as its output rises. These situations are likely to lead to rising average costs of production.
- Periods of increasing returns (in the short run and in the long run) mean that the firm is becoming more efficient as its output rises. These situations are likely to lead to falling average costs of production.

The impact of the law of diminishing returns and the concept of returns to scale on costs of production will be dealt with in more detail in the next chapter.

REALWORLD ECONOMICS 5.1

Diminishing returns and returns to scale in retailing and distribution

In retailing, firms often face a situation in which some factors of production are fixed and so changes in potential demand can only be met by changing the quantity of variable factors of production. For retailers, their main fixed factors are their shops. Retailing is very seasonal, with many shops recording much higher sales in the pre-Christmas period.

This seasonality can cause high levels of inefficiency. Many stores are designed to have a scale that can cope with the



The local convenience store, Tesco-style

pre-Christmas peak. For the remainder of the year this fixed input is generally underused. However, materials (stock) and labour are variable factors. In the peak period, retailers purchase more stocks of goods to be sold and most will increase their staffing levels. For example, for Christmas 2015, Sainsbury's and Marks and Spencer both employed an additional 15,000 staff. This process occurs annually and these firms know, from experience, the optimum level of additional staffing required during this period. However, it is difficult to maintain efficiency because these large increases put pressure on fixed factors, such as the shop premises, which become very overcrowded.

For distribution companies the potential for diminishing returns is even greater. The Royal Mail recruited 19,000 temporary workers for the Christmas post in 2015 – approximately 20% of its normal workforce.

Although returns to scale are a measure of the efficiency of converting inputs to outputs, supermarkets have found that customer behaviour can have a big impact on their efficiency. Traditionally families had one large shopping trip per week, usually at the weekend. Shoppers now tend to visit supermarkets more often but with smaller purchases and a greater focus on fresh food. As a consequence, large superstores, once seen as the stores that offered the highest returns to scale, have become less efficient. The supermarket chains are therefore focusing on establishing smaller, local convenience stores. These provide high 'output' because shoppers are prepared to pay higher prices for convenience and freshness.

Discussion point

Why is it important for retailers to understand the concepts of diminishing marginal returns and returns to scale?

Review questions

Total: 40 marks

- 1 Output increases by 20% from a 25% increase in all inputs. This would be an example of:
- A Constant returns to scale
 - B Decreasing returns to scale
 - C Diminishing returns
 - D Increasing returns to scale

(1 mark)

Questions 2 and 3 are based on the data shown in the following table, which is based on labour being the only variable input.

Units of labour	Total returns	Marginal returns	Average returns
1	10	10	10
2	24	(b)	12
3	42	18	(c)
4	54	12	13.5
5	(a)	6	12
6	58	-2	9.7

- 2 (a) Calculate total returns from 5 units of labour. (1 mark)
 (b) Calculate marginal returns from 2 units of labour. (1 mark)
 (c) Calculate average returns from 3 units of labour. (1 mark)
- 3 After how many units of labour do diminishing returns occur?
- A 3
 - B 4
 - C 5
 - D 6
- (1 mark)
- 4 What is meant by the term 'short run'? (3 marks)
- 5 What is meant by the term 'long run'? (3 marks)
- 6 Define the term 'constant returns to scale'. (3 marks)
- 7 Average output is 8 units when 3 units of labour are employed; average output is 9 units when 4 units of labour are employed. What is the level of marginal returns from the fourth unit of labour? (3 marks)
- 8 Explain *one* possible diseconomy of scale that might cause decreasing returns to scale. (4 marks)
- 9 Explain why marginal returns fall in the short run as additional variable factors are added. (5 marks)
- 10 What is the difference between 'diminishing returns' and 'decreasing returns to scale'? (6 marks)
- 11 Explain *two* possible economies of scale that might cause increasing returns to scale. (8 marks)

Costs of production

Key concepts from Year 1

Costs of production were introduced in Chapter 14 of the Year 1 companion textbook. The key concepts covered were:

- **fixed costs** – costs that do *not* vary directly with output, in the short run;
- **variable costs** – costs that do vary directly with output, in the short run.
- The concepts of total costs (TC) and average costs (AC), or average total costs (ATC), were also introduced and defined. These definitions are reproduced below.

Before progressing, you should ensure that you are aware of these terms and how TC and AC (ATC) are calculated. You should also recognise the difference between short-run and long-run costs.

This chapter extends the ideas covered in Chapter 14 in the first year. The concept of marginal cost is introduced. The concepts of diminishing marginal returns and returns to scale, from Chapter 5 of this book, are used to explain the shape of the marginal, average and total cost curves and how factor prices and productivity affect firms' costs of production and their choice of factor inputs.

The difference between marginal, average and total costs

The relationship between different costs is shown in Table 6.1. Please note that this is identical to Table 14.2 in the Year 1 companion textbook but with the addition of a column showing marginal costs.

Table 6.1 Total costs, average costs and marginal costs

Production (units)	Fixed costs (FC) (£)	Total variable costs (TVC) (£)	Total costs TC (£)	Average fixed costs (AFC) (£)	Average variable costs (AVC) (£)	Average total costs (ATC) (£)	Marginal costs (MC)
0	30	0	30	na	na	na	na
1	30	20	50	30.00	20.00	50.00	20
2	30	30	60	15.00	15.00	30.00	10
3	30	36	66	10.00	12.00	22.00	6
4	30	38	68	7.50	9.50	17.00	2
5	30	40	70	6.00	8.00	14.00	2
6	30	45	75	5.00	7.50	12.50	5
7	30	54	84	4.29	7.71	12.00	9
8	30	68	98	3.75	8.50	12.25	14
9	30	87	117	3.33	9.67	13.00	19
10	30	110	140	3.00	11.00	14.00	23
11	30	146	176	2.73	13.27	16.00	36
12	30	198	228	2.50	16.50	19.00	52

The reasons for the shapes of the marginal, average and total cost curves

Figure 6.1 reproduces the data in Table 6.1 in graphical forms so that the shapes of the marginal, average and total cost curves can be seen.

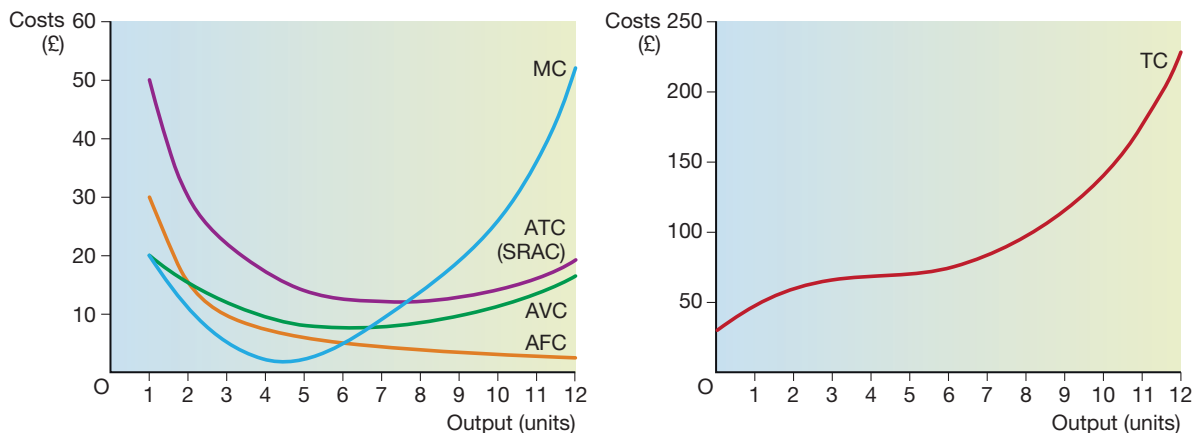


Figure 6.1 The marginal, average and total cost curves

(a) Marginal and average costs

(b) Total costs

In the short run, fixed costs remain the same. However, the fixed factors of production will impose a limit on the amount that a firm can produce.

At low levels of output, variable factors of production are unlikely to work efficiently and so returns are low. As a consequence, the cost of making each item tends to be very high when few units are produced.

As more workers are employed, division of labour and specialisation can be used in order to improve labour productivity. This means that efficiency improves and so the costs of each unit produced should fall.

If output continues to rise, diminishing returns occur. Thus efficiency will fall when output gets close to the maximum capacity of the firm. This is because there is very little flexibility as all of the fixed factors of production are being used fully. With low levels of return, the cost of making additional items rises considerably.

Figure 6.1(a) shows how marginal cost starts quite high (at £20 for the first unit) and then falls to only £2 for the fourth and fifth units. However, it then rises gradually (at first) but very dramatically as maximum capacity is approached. Average cost falls until the seventh unit, partly because fixed costs are spread over more units of output (so AFC falls) and partly because AVC is also falling until the sixth unit.

Figure 6.1(b) shows the shape of the total cost line. Since fixed costs are incurred regardless of output, total costs are £30 at zero units of output. The gradient of the line then falls because a more efficient use of inputs leads to better returns, which

Key terms

Total costs (TC) are the sum of fixed costs and variable costs.

Average costs (AC) or average total costs (ATC) are the total costs divided by the number of units produced.

Marginal costs (MC) are the addition to total costs arising from making one more item.

means that additional units of output are adding little to total costs (marginal costs are falling). Once marginal costs begin to rise, the total cost line becomes steeper.

The problem of overused fixed factors of production can be solved in the long run by introducing additional fixed factors of production, such as more capital equipment and new buildings. The firm will also benefit from internal economies of scale. (These will be discussed in Chapter 7.) Capital equipment will no longer be overused and so production becomes more efficient because firms can plan the optimum mix of the different factors of production. In the short run, firms may face situations in which they have insufficient fixed assets, such as a factory that is too small.

In the long run, firms can limit sudden rises in costs because they can plan their fixed capital more carefully. As a firm approaches its maximum capacity, it can introduce additional machinery or build a new factory to ensure that it can supply all of the products that it wishes.

How factor prices and productivity affect firms' costs of production and their choice of factor inputs

In Chapter 5, Table 5.2 was used to illustrate returns to scale. The first two columns of Table 5.2 are reproduced in Table 6.2 so that the impact of factor prices on costs of production can be shown. The final column of Table 5.2 is also shown (as Column 5).

Column 1 Units of input	Column 2 Total output (total returns)	Column 3 Total costs (£)	Column 4 Average total costs (£) (Column 3/Column 2)	Column 5 Returns to scale
10	4	10	2.50	N/A
20	12	20	1.67	Increasing
30	30	30	1.00	Increasing
40	52	40	0.77	Increasing
50	75	50	0.67	Increasing
60	96	60	0.63	Increasing
70	112	70	0.63	Constant
80	128	80	0.63	Constant
90	135	90	0.67	Decreasing
100	140	100	0.71	Decreasing
110	143	110	0.77	Decreasing

Table 6.2 Returns to scale

If we assume that each unit of input costs £1, then Column 3 shows the total cost. For example, if there are 10 units of input, then total costs will be $10 \times £1 = £10$. Similarly, the total costs of 20 units of input will be $£10 \times 2 = £20$.

Column 4 shows the average cost (average total cost), which is calculated by dividing total costs by total output.

Table 6.2 shows the impact of returns to scale on average costs. Column 5 shows that returns to scale (measured by output/input in Table 5.2) were increasing until 60 units of output were used. ATC continually fell as units of input rose from 10 through to 60 units. From 60 to 80 units of output, there were constant returns to scale – this coincided with the range of inputs at which ATC was constant (and at its lowest level of £0.63). Once decreasing returns occurred (above 80 units of input), ATC rose.

Conclusion: There is an inverse relationship between returns to scale and average costs.

Factor prices and costs of production

What happens if factor prices, such as wages for labour, change? Column 3 shows that 10 units of input (all four factors of production) incur a total cost of £10. If each input doubles in price from £1 to £2 per unit, the cost of 10 inputs becomes $10 \times £2 = £20$. Thus the average cost becomes $£20/4 = £5$ per unit of output. Similarly, when 20 units of input are used the AC becomes $£40/12 = £3.33$ per unit of labour. In effect, a doubling of the price of factor inputs leads to a doubling of average total costs.

Productivity and costs of production

What happens if productivity changes? Productivity is output/input (Column 2/Column 1). Column 2 shows that output is 4 units when 10 units of input are used. If productivity doubles, then each input produces twice as many units of output. For example, total output from 10 units of input will double from 4 to 8 units. Referring back to Table 6.2, the ATC from 10 units of input will now be $£10/8 = £1.25$. Similarly, total output when 20 units of input are used becomes 24 units of

REALWORLD ECONOMICS 6.1

The use of marginal cost by airlines

Typically, the airline industry has tended to operate at output that is below capacity. This situation occurs because of:

- **seasonal demand** – airlines must plan their capacity to cope with peak demand that will only occur for a limited period;
- **brand image** – major airlines believe that they need to provide

- a wide range of routes;
 - **airport allocations** – once an airline decides not to operate a particular route, the airports will allocate takeoff times to a different airline;
 - **high fixed costs** – fixed costs account for 80–90% of total costs for airlines.
- Marginal cost is the cost of one

additional item. Although its normal pricing must ensure that it covers fixed costs, empty seats on a flight that is about to depart incur minimal cost for the airline. Food provided is likely to be the marginal cost of an empty seat, if the flight is about to depart. This situation has encouraged airlines to set a price that covers variable costs only.

Usually, high prices are charged to people that place a high value on flights – primarily business customers. If customers who book early cover fixed costs, then it makes sense to offer last-minute tickets at a price that covers the marginal costs of that ticket. This is known as marginal cost pricing.

Discussion point

Is this a sensible strategy for airlines? Assess the possible ‘pros and cons’ of using marginal cost pricing on a regular basis.



Spare capacity in the airline industry has resulted in marginal cost pricing

output. Thus the $ATC = £20/24 = £0.83$ per unit of labour. In effect, a doubling of productivity leads to a halving of the average total costs of production.

Factor prices, productivity and the choice of factor inputs

To date we have assumed that, in the long run, there is an ideal balance between the factors of production (labour, land, capital and enterprise).

The balance will be changed if factor prices change. Other things being equal, an increase in a factor price, such as wages for labour or rent for land, will reduce profit for the entrepreneur. However, different factors of production can often be substituted for each other. As a rule, higher factor prices for a particular factor of production will lead to a switch away from that input and towards other inputs (factors of production).

Productivity also influences the choice of factor inputs. If labour productivity increases by 50%, say through better training, then this will encourage entrepreneurs to use more labour, possibly replacing other factor inputs.

Review questions

Total: 30 marks

Questions 1 to 4 are based on the table below.

Output (units)	Total costs (£)	Marginal costs (£)	Average costs (£)
1	50	20	50
2	60	10	30
3	66	6	?
4	76	?	19
5	95	19	19
6	?	31	21

- Calculate total costs of 6 units of output. *(1 mark)*
- Calculate the marginal cost of the fourth unit of output. *(1 mark)*
- Calculate average total costs of 3 units of output. *(1 mark)*
- Calculate the fixed costs. *(1 mark)*
- What is meant by the term 'marginal cost'? *(3 marks)*
- Explain the link between factor prices and costs of production. *(4 marks)*
- Explain the link between productivity and costs of production. *(5 marks)*
- Explain how diminishing marginal returns affect the shape of the marginal cost curve. *(6 marks)*
- Explain *two* factors that might lead to a firm deciding to replace labour with capital equipment. *(8 marks)*

Economies & diseconomies of scale

Key concepts from Year 1

Most of this topic was covered in Chapter 15 of the Year 1 companion textbook. That chapter introduced the concept of economies of scale and examined the difference between internal economies of scale and external economies of scale. Diseconomies of scale and their reasons were explained. The chapter concluded by studying the relationship between economies of scale and diseconomies of scale and their impact on the shape of the long-run average cost curve. The concept of the minimum efficient scale (MES) was also explained, in the context of the shape of the long-run average cost (LRAC) curve. Ensure that you are familiar with this material before commencing this chapter.

This chapter extends the content of Chapter 15 in the Year 1 book. It explains the relationship between returns to scale and economies or diseconomies of scale and explains the significance of the minimum efficient scale for the structure of an industry and barriers to entry. The chapter concludes with an introduction to, and examination of, the L-shaped long-run average cost curve.

The relationship between returns to scale and economies or diseconomies of scale

The definitions below summarise the meanings of returns to scale and economies or diseconomies of scale.

Returns to scale describes the proportionate (or percentage) change in output of a firm or industry resulting from a proportionate (or percentage) increase in all inputs.

Returns to scale can be:

- **increasing returns to scale** – if the percentage (or proportionate) increase in output is greater than the percentage (or proportionate) increase in input;
- **constant returns to scale** – if the percentage (or proportionate) increase in output is equal to the percentage (or proportionate) increase in input;
- **decreasing returns to scale** – if the percentage (or proportionate) change in output is less than the percentage (or proportionate) increase in input.

Internal economies of scale are the advantages that an organisation gains due to an increase in its size. These advantages cause an increase in productive efficiency and thus a decrease in the average (total) cost of production.

Internal diseconomies of scale are the disadvantages that an organisation experiences due to an increase in size. These cause a decrease in productive efficiency and therefore an increase in the average (total) cost of production.

Author tip

The new material in Year 2 is focused purely on internal economies of scale. You should also ensure that you revise external economies and diseconomies of scale (the benefits and problems of being a firm within a large localised industry) for your A-level examinations.

Returns to scale are clearly very closely related to economies of scale. However, whereas economies of scale are measured in terms of unit costs, returns to scale are measured in terms of output divided by input. Of course, if output per unit of input increases, it is almost certain to lead to lower unit costs (and thus economies of scale). However, if, for example, the improvement in output has been achieved by paying much higher wage rates to workers, it is possible for returns to scale to improve but unit costs to rise too – this would mean diseconomies of scale.

Similarly, bulk buying is an economy of scale which cuts unit or average costs. If a firm makes no changes to its production process but its raw materials cost 10% less than before, its average costs of production will fall. This will mean that economies of scale are achieved (lower average costs) but returns to scale are the same (constant) because the ratio of outputs to inputs is still the same.

The key difference between the two concepts is the way in which they are measured, with economies/diseconomies measured as changes in average costs but returns to scale measured as the ratio of output to input.

The relationship between economies of scale, diseconomies of scale and the shape of the long-run average cost curve

This relationship was explained in Chapter 15 of the Year 1 book. Figure 7.1 shows the shape of the long-run average cost curve. Typically, it is shown as a U-shaped curve.

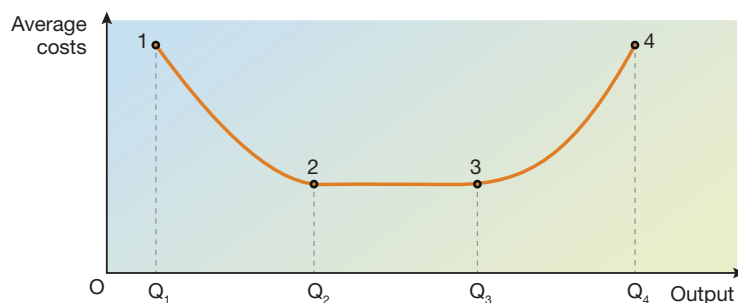


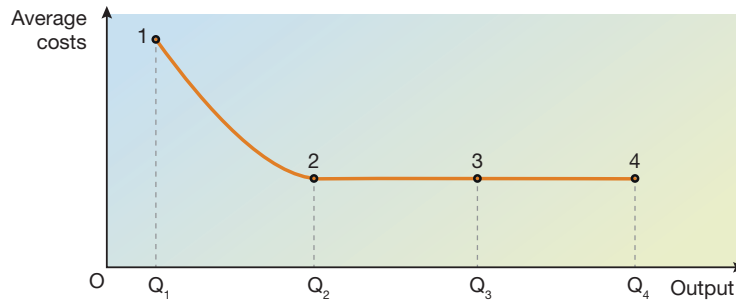
Figure 7.1 The U-shaped long-run average cost curve

As output increases from Q_1 to Q_2 , there is greater scope for internal economies of scale. This leads to a steady decrease in average costs (from Point 1 to Point 2 on the average cost curve). As the curve approaches Point 2, there are some diseconomies of scale evident. In themselves, these diseconomies would cause average costs to rise, but at this level of output the diseconomies of scale are outweighed by the economies of scale and so average cost continues to fall, albeit more slowly. Between output levels Q_2 and Q_3 the economies of scale are exactly matched by the diseconomies of scale and so average costs remain the same. Between Points 2 and 3 we have constant returns to scale. Between output levels Q_3 and Q_4 the diseconomies become more significant, leading to an increase in average costs (shown by the increase in AC between Points 3 and 4).

The L-shaped long-run average cost curve

Traditionally, economists have believed that the long-run average cost curve is U-shaped, as shown by Figure 7.1. However, this view has been challenged in recent times, with a number of empirical studies suggesting that the long-run average cost curve is L-shaped.

Figure 7.2 The L-shaped long-run average cost curve



As with the U-shaped curve, exponents of the L-shaped long-run average cost curve agree that at low levels of output, increasing economies of scale are achieved as output increases. There is also agreement that beyond a certain point (Point 2 in this case) the impact of economies and diseconomies of scale tend to cancel each other out. However, the L-shaped curve is based on the assumption that, at very high levels of output, diseconomies of scale do not dominate. Beyond output Q_3 , average costs remain constant. (Many studies suggest that average costs continue to fall beyond Points 2 and 3, but only slightly.)

The L-shaped long-run average cost curve occurs because of the impact of technical and production economies and managerial costs. Studies suggest that these economies are still achievable at very high levels of output. Managerial diseconomies of scale lead to lower efficiency, but this is countered by the increased efficiency of the production process. Overall, this leads to constant (or slightly declining) average costs.

Over 60 years ago, Joel Dean pioneered studies of production costs in the USA and concluded that in a range of industries, such as leather goods, stockings, furniture manufacturing and shoe retailing, the long-run average cost curve was closer to an L-shape than a U-shape. Subsequent studies in other countries have confirmed this conclusion, notably Gupta's study of Indian manufacturing. In his study, 18 of the 29 industries indicated LRAC curves that were L-shaped. However, this study did reveal that results varied between industries.

The concept of the minimum efficient scale of production

The minimum efficient scale shows the lowest level of output at which a firm is able to achieve average costs that are competitive with large-scale rivals.

The LRAC shows that, as output increases, long-run average cost declines steadily at first, mainly as a result of greater economies of scale. The LRAC then reaches a point at which constant returns to scale occur. After a period of constant returns to scale, average costs either:

- increase (for a U-shaped LRAC curve);
- stay constant or slightly decrease (for an L-shaped LRAC curve).

In both Figures 7.1 and 7.2, the minimum efficient scale is shown by Point 2, with Q_2 being the level of output at which the minimum efficient scale occurs.

The minimum efficient scale is significant for two main reasons:

- **The structure of an industry:** The extent to which economies and diseconomies of scale impact upon costs varies considerably between industries. For example, in car manufacturing it is estimated that an output of 5 million cars is required

Key term

Minimum efficient scale

is the lowest level of output at which long-run average cost is minimised.

before expenditure on research and development can be justified. If fewer than 5 million cars are produced, research and development costs will be too high and impact negatively on the car manufacturer's average costs. This means that the minimum efficient scale for car manufacturers is likely to be very high.

- **Barriers to entry:** The scope for economies of scale is viewed as one of the key barriers to entry into an industry. Invariably, small firms are unable to compete with established firms that can produce low-cost products on a large scale. The minimum efficient scale provides an indication of this barrier. In Figure 7.1, it can be seen that if a new firm has an output of Q_1 , then it will be unable to match the average costs of production of established firms whose output lies between Q_2 and Q_3 . The difference between the new firm's average costs and the average costs at Point 2 (which is the minimum efficient scale or MES) indicates the extent to which the new entrant will be uncompetitive. The MES is a guide to new entrants – this is the level of output that they need to be able to achieve in order to be competitive with existing firms.

REALWORLD ECONOMICS 7.1

The L-shaped LRAC

Some studies in the UK and Europe support the view of the L-shaped long-run average cost curve.

Studies of 18 different UK airports in the 1970s and 1990s, by Doganis et al., indicated an L-shaped long-run average cost curve. For passenger numbers below 3 million, the long-run average costs fell significantly as output increased. However, when passenger numbers were between 3 million and 5 million, the long-run average costs tended to even out.

The car industry also displays features consistent with the L-shaped LRAC. The large numbers



The minimum efficient scale for commercial vehicles is much lower than for cars

Activity	Optimum level of output per annum (cars)
Production/technical economies	
■ Casting engine blocks	1 million
■ Pressing car panels	1 to 2 million
■ Final assembly of car	250,000
Non-production activities	
■ Sales	2 million
■ Finance	2.5 million
■ Research and development (R&D)	5 million

Table 7.1 Optimum level of output per year for certain activities in the car industry

of mergers and takeovers of car manufacturers have occurred because of the scope for economies of scale and because of the high level of the minimum efficient scale for many activities involved in car manufacturing.

Table 7.1 provides a summary of the optimum level of output per year, in terms of benefits from internal economies of scale, for certain activities within the car industry. The optimum level occurs where average costs are minimised.

Table 7.2 shows an index of average costs per car for varying levels of output.

These data have influenced the structure of the car industry. Although assembly lines are efficient with output of 250,000 cars per annum, larger plants enable car manufacturers to benefit from full economies for pressing panels and casting engine blocks. Thus firms need sales in excess of 1 million to be competitive in terms of production. However, it is the non-

Output per annum (cars)	Index of average costs
100,000	100
250,000	83
500,000	74
1,000,000	70
2,000,000	66

Table 7.2 Index of average costs per car for varying levels of output

production economies of scale that encourage even larger firms. R&D costs are higher than the minimum efficient scale unless 5 million cars are produced. This has encouraged integration of car firms, so that R&D results can be fully exploited. In many cases, car firms share R&D

expenses (and benefits) because they lack the scale to operate competitively in terms of R&D.

Above 2,000,000 cars, average costs show little or no further reduction.

Sources: 'The economics of British airports' by Rigas Doganis and 'Report on the economic prospects for the automotive industry in the UK and Europe' by Garel Rhys

Exercises

Total: 15 marks

- 1 Measured as passenger numbers, the minimum efficient scale for UK airports in the 1970s was:
 A 2 million
 B 2.5 million
 C 3 million
 D 5 million

(1 mark)

- 2 Calculate the percentage fall in average costs for a car firm that increases its output from 250,000 cars per annum to 1,000,000 cars per annum. **(2 marks)**

- 3 Explain how the minimum efficient scale in the UK car industry has encouraged mergers between car manufacturers. **(6 marks)**

- 4 Refer to Table 7.1. Explain why the optimum level of output for research and development is much higher than the optimum level of output for the final assembly of cars. **(6 marks)**

Review questions

Total: 30 marks

- 1 Which pair of cost curves shows a decrease and then an increase as output rises?
 A ATC and MC
 B AFC and TC
 C ATC and AFC
 D MC and TC **(1 mark)**
- 2 What is meant by the term 'constant returns to scale'? **(3 marks)**
- 3 Define the term 'minimum efficient scale'. **(3 marks)**
- 4 Explain the difference between increasing returns to scale and economies of scale. **(4 marks)**
- 5 Explain the circumstances in which a business might have constant returns to scale. **(5 marks)**
- 6 Explain how the minimum efficient scale for an industry might act as a barrier to entry for small firms. **(6 marks)**
- 7 Analyse two reasons why empirical studies might conclude that a firm has an L-shaped LRAC curve. **(8 marks)**

Marginal, average & total revenue

Key concepts from Year 1

This chapter and Chapter 9 build on Chapter 16 of the Year 1 companion textbook. That chapter introduced the concepts of total revenue and average revenue and showed how they are calculated. The reason why the average revenue curve is the firm's demand curve was also explained. Diagrams were used to illustrate the relationships between total revenue and average revenue. Check that you are familiar with these topics before you commence this chapter.

This is a short chapter that aims to ensure that you have a full understanding of revenue. It focuses on marginal revenue, which was not covered in Year 1. It explains the difference between marginal revenue, average revenue and total revenue and the relationships between:

- average revenue and marginal revenue;
- marginal revenue and total revenue.

Please note that you should also be aware of why the average revenue curve is the firm's demand curve – this is covered in Chapter 16 of the Year 1 companion textbook.

The difference between marginal, average and total revenue

A firm's revenue is a measure of the money received from the sale of its goods and services. However, when studying total revenue, average revenue and marginal revenue, economists usually focus on the revenue from a particular good. These types of revenue are calculated as follows:

$$\text{Total revenue (TR)} = \text{Price per unit (p)} \times \text{Quantity of units sold (q)}$$

For example, if a firm sells 200 goods at a price of £7 per unit, then its total revenue is $£7 \times 200 = £1400$.

$$\text{Average revenue (AR)} = \text{Total revenue (TR)} / \text{quantity of goods sold (q)}$$

For example, if a firm receives total revenue of £6000 and sells 300 goods, then its average revenue is $£6000/300 = £20$.

$$\text{MR of unit 'x'} = \text{Total revenue of 'x' units} - \text{Total revenue of 'x - 1' units}$$

Key terms

Total revenue (or sales revenue) is the total money received from the sale of a firm's goods and services.

Total revenue can also refer to the total money received from the sale of a particular good or service.

Average revenue is the average receipt of money for each good or service that is sold.

Marginal revenue is the addition to total revenue as a result of the sale of one more unit of output.

For example, if a firm receives total revenue of £5500 from selling 25 goods and a total revenue of £5400 from selling 24 goods, then the marginal revenue of the twenty-fifth good is $£5500 - £5400 = £100$.

The relationship between average revenue and marginal revenue, and the relationship between marginal revenue and total revenue

The nature of total revenue, average revenue and marginal revenue depends on the market structure in which the firm operates:

- In a perfectly competitive market, equilibrium price is determined by the interaction of demand and supply. All products are homogeneous and firms are price takers, so all firms charge the same price. Furthermore, no single firm can influence the price of the good and so, no matter how many goods it sells, an individual firm will sell all of its goods at the same price.
- In an imperfect market, such as oligopoly or monopoly, the firm's product is differentiated from any competitors. Thus we have a downward-sloping demand curve.

Table 8.1 shows the revenue of a firm in a perfectly competitive market in which the equilibrium market price is £6. (Please note that the data are identical to Table 16.1 in the Year 1 companion textbook, with the exception of the marginal revenue column.)

Table 8.2 shows the revenue of a firm in an imperfectly competitive market in which the firm's demand curve is downward sloping. (Please note that the data are identical to Table 16.2 in the Year 1 companion textbook, with the exception of the marginal revenue column.)

The relationship between average revenue and marginal revenue, and the relationship between marginal revenue and total revenue, in a perfectly competitive market

Table 8.1 Total revenue, average revenue and marginal revenue for a particular good

Quantity	Price (£)	Total revenue (£)	Average revenue (£)	Marginal revenue (£)
1	6	6	6	6
2	6	12	6	6
3	6	18	6	6
4	6	24	6	6
5	6	30	6	6
6	6	36	6	6
7	6	42	6	6
8	6	48	6	6*

Note: * Since the marginal revenue of the eighth good shows the additional revenue from moving from 7 goods to 8 goods, on a graph it is usual to plot the point halfway between 7 and 8 (at 7.5). Similarly, the marginal revenue of the first good is plotted at 0.5, the second good at 1.5, and so on.

In a perfectly competitive market the average revenue is the price because every item produced by a firm is sold at the same price (the market equilibrium price). Thus the AR line is a horizontal line from the point on the 'y' axis that marks the equilibrium price – £6 in this example (Figure 8.1). The MR line is identical to the AR line. It is horizontal because each additional good is sold for the same price. In this case, every good sold brings in an extra £6 in revenue.

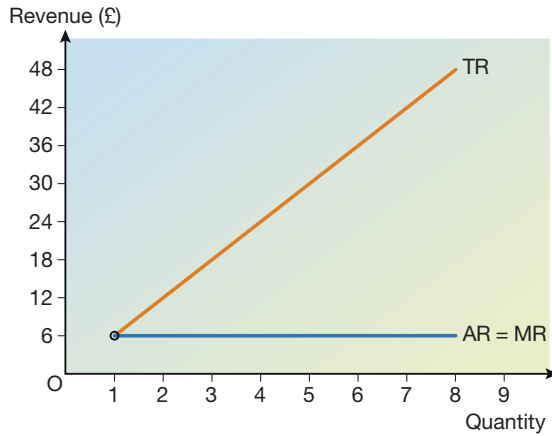


Figure 8.1 TR, AR and MR in a perfectly competitive market (when price is fixed)

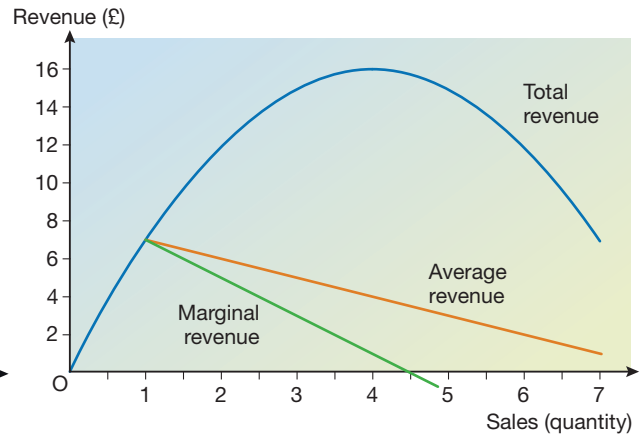


Figure 8.2 TR, AR and MR in an imperfectly competitive (concentrated) market

It can be seen that the total revenue line is a straight line, sloping upwards from left to right. The gradient of this line is determined by the marginal revenue (MR). As the price is £6, each extra good brings in £6 (the MR). The gradient of the TR line thus shows a £6 increase in total revenue (TR) for every single unit increase in the quantity.

The relationship between average revenue and marginal revenue, and the relationship between marginal revenue and total revenue, in an imperfectly competitive (concentrated) market

As can be seen from Figure 8.2, TR increases as quantity rises from zero to 4 units, but peaks at that point. Up to this quantity the marginal revenue is positive and so the revenue gained from selling an additional unit outweighs the loss of revenue from having to charge a lower price for each unit sold. However, as quantity moves from 4 units to 7 units, the marginal revenue becomes negative and so total revenue falls because the revenue from the extra unit sold does not outweigh the revenue lost from having to charge a lower price. For example, moving from selling 5 units to 6 units means having to cut the price from £3 to £2. The sixth unit brings in an extra £2, but the 5 units have each been sold for £1 less (£2 instead of £3), meaning that $£1 \times 5 = £5$ has been sacrificed. Overall, the £2 gained from the sixth unit has been outweighed by the £5 lost from the first 5 units. This gives an overall fall in total

Quantity	Price (£)	Total revenue (£)	Average revenue (£)	Marginal revenue (£)
0	8	0	8	
1	7	7	7	7*
2	6	12	6	5
3	5	15	5	3
4	4	16	4	1
5	3	15	3	-1
6	2	12	2	-3
7	1	7	1	-5

Note: * Since the marginal revenue of the first good shows the additional revenue from moving from zero goods to 1 good, on a graph the point should be plotted halfway between 0 and 1 (at 0.5), and so on. However, this mathematical precision is generally ignored for the sake of convenience.

Table 8.2 Total revenue, average revenue and marginal revenue in an imperfectly competitive (concentrated) market

revenue of £3 (£5 – £2). This can be confirmed by a look at Table 8.2, which shows that TR has fallen from £15 (from selling 5 units) to £12 (from selling 6 units).

Conclusion: *When MR is positive, TR rises; when MR is negative, TR falls.*

In a concentrated market, average revenue falls as output rises. However, marginal revenue falls more rapidly. In order to sell more units, a firm must lower its price (because the demand curve slopes downwards). Thus a lowering of price by £1 lowers the average revenue by £1. However, because all of the previous units are now being sold at a lower price, the MR falls by more than £1 (and can be negative).

Table 8.2 shows the exact relationship between AR and MR in a concentrated market. For each additional unit of output, price (and thus AR) falls by £1 and MR falls by £2. This relationship holds true for all demand curves. Figure 8.2 confirms this relationship – the decline (gradient) of the MR line is twice that of the AR line.

Conclusion: *As output rises, MR declines at twice the rate of AR.*

REALWORLD ECONOMICS 8.1

Increasing total revenue in the energy market

In 2008, Ofgem conducted a survey of energy prices. This survey revealed that energy companies were charging different prices to different customers for a broadly homogeneous product (gas or electricity).

This practice, known as ‘price discrimination’, will be covered in detail in Chapter 17. This article just focuses on how price discrimination can be used to increase total revenue.

Ofgem identified various ways in which energy companies were able to increase total revenue. Arguably the most significant method was the use of long-term contracts, usually for a minimum of one year, which

meant that customers were unable to switch their suppliers for at least 12 months, and sometimes much longer. However, unlike many long-term contracts, these contracts did not involve fixed prices and so a customer might find themselves paying a higher price than the original contract, but without the option to change supplier until the contract expired.

Ofgem’s survey focused more on the methods that led to different customers being charged different prices. These methods are outlined below.

(1) Standard credit v direct debit tariffs

Energy companies prefer automatic, monthly direct debit payments from customers because they are more certain and easier to administer. On average, direct debit payment saves the energy companies £25 per customer per annum. However, Ofgem found that customer tariffs did not reflect this saving. Gas customers who did not pay by direct debit paid an extra £94,

while electricity customers paid an additional £38. Dual fuel customers who did not pay by direct debit paid £133 extra. Ofgem was concerned that this price discrimination led to higher prices being paid by the most vulnerable members of society, such as those without bank accounts.

(2) In-area v out-of-area price differentials

Originally, energy suppliers were based in geographical regions, such as Southern Electric and Northern Gas. After privatisation, each supplier was allowed to compete in other regions. Ofgem found that, on average, ‘in-area’ customers paid £24 more for their energy. These customers tended to be those who had kept with the original supplier, usually because of a lack of information on competitors’ prices. Many of these customers were elderly and also lacked internet access and so found it difficult to get information and/or switch suppliers.

(3) Off-gas grid

People who lived in remote areas



The London offices of Ofgem – the government regulator for gas and electricity markets in Britain

without a mains gas supply also paid higher prices for electricity – on average £40 more than other customers. Again, they tended to be more vulnerable consumers, usually on low-incomes in rural areas.

Many examples of maximising total revenue feature methods of charging higher prices to more affluent customers who are prepared to pay higher prices for a particular brand or good. In the energy market, the opposite seems

to have applied. Furthermore, consumers of large quantities of energy tend to pay lower rates, largely because most tariffs involve a fixed charge plus a payment for the amount of energy used. For customers on low incomes, especially pensioners living alone, energy prices per unit consumed tend to be much higher than for other customers.

Ofgem concluded that the energy suppliers were not acting in

the interests of certain consumers and so it decided to take action to help those consumers.

Source: Ofgem

Exercises Total: 20 marks

- 1 Explain two methods that Ofgem might use to overcome the 'exploitation' of these consumers. (8 marks)
- 2 Should Ofgem intervene in this market? Justify your view. (12 marks)

Review questions

Total: 20 marks

- 1 Define the term 'marginal revenue'. (3 marks)
- 2 Does Table 8.3 show a perfectly competitive market or a concentrated market? Explain your reasoning. (5 marks)

Quantity	Total revenue (£)	Average revenue (£)	Marginal revenue (£)
1	23	23	23
2	46	23	23
3	69	23	23

Table 8.3 Total revenue, average revenue and marginal revenue for a particular good

The remaining questions are based on Table 8.4.

Units of output	Total revenue (£)	Average revenue (£)	Marginal revenue (£)
1	33	33	33
2	60	30	(c)
3	81	(b)	21
4	96	24	15
5	105	21	9
6	108	18	3
7	(a)	15	-3

Table 8.4 Total revenue, average revenue and marginal revenue for Good 'X'

- 3 (a) Calculate total revenue from 7 units of output. (1 mark)
 (b) Calculate average revenue from 3 units of output. (1 mark)
 (c) Calculate marginal revenue of the second unit of output. (1 mark)
- 4 Explain the relationship between marginal revenue and total revenue in Table 8.4. (4 marks)
- 5 Explain the relationship between marginal revenue and average revenue in Table 8.4. (5 marks)

Profit

Key concepts from Year 1

The concept of profit was introduced in Chapter 16 of the Year 1 companion textbook.

In this chapter we consider how profit is calculated. The difference between normal and abnormal (supernormal) profit is explained. The chapter concludes with an examination of the role of profit in a market economy.

Profit is the difference between total revenue and total costs

$$\text{Profit} = \text{TR} - \text{TC}$$

A firm sells 500 units for £6 and has an average cost of production of £4 per unit. It will earn total revenue of $500 \times £6 = £3000$ and its total costs will be $500 \times £4 = £2000$. Profit = $£3000 - £2000 = £1000$. This leads to a second method for calculating profit:

$$\text{Profit} = \text{Quantity (q)} \times [\text{Selling price (p)} - \text{Average cost (AC)}]$$

$$\text{Profit} = q \times (p - AC)$$

In the above example, profit is $500 \times (£6 - £4) = 500 \times £2 = £1000$.

The difference between normal and abnormal (supernormal) profit

When measuring profit, economists use a different approach to other disciplines, such as accounting. Economists are primarily concerned with scarcity, and so they recognise the ‘opportunity cost’ of the resources used to make a profit. If the entrepreneur could have earned a profit of £12,000 from the next best alternative use of the resources at their disposal, then the entrepreneur will only use those resources if the profit (TR – TC) equals or exceeds £12,000. This £12,000 is the entrepreneur’s ‘normal profit’. If profit falls below the normal profit, the entrepreneur will no longer wish to continue operating the firm.

Since ‘normal profit’ is necessary for the firm to keep operating, economists classify it as a cost – in effect, it is the cost of the entrepreneur. Thus total costs to an economist will include ‘normal profit’ because this is necessary to attract the fourth factor of production – enterprise.

Key terms

Profit is the difference between the total (sales) revenue of a firm and its total costs: Profit equals total revenue minus total costs.

Profit maximisation is making the highest possible level of profit.

Normal profit is the minimum level of profit needed to keep a firm operating in its present market.

Abnormal profit (or supernormal profit) is any profit over and above the level of normal profit.

Any profit made over and above ‘normal profit’ is described as ‘**abnormal profit**’ or ‘**supernormal profit**’. Thus any profit shown in Table 9.1 is ‘abnormal profit’ because the normal profit is included in the total costs as the cost of enterprise.

Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7
Output (units)	Price	Total revenue	Total costs	Profit	Average revenue	Average costs
q	p (£)	TR (£)	TC (£)	π (£)	AR (£)	AC (£)
0	25	0	30	-30	na	na
1	24	24	50	-26	24	50
2	23	46	60	-14	23	30
3	22	66	66	0	22	22
4	21	84	68	16	21	17
5	20	100	70	30	20	14
6	19	114	75	39	19	12.5
7	18	126	84	42	18	12
8	17	136	98	38	17	12.25
9	16	144	117	27	16	13
10	15	150	140	10	15	14
11	14	154	176	-22	14	16
12	13	156	228	-72	13	19

Table 9.1 Total revenue, total costs, average revenue, average costs and profit

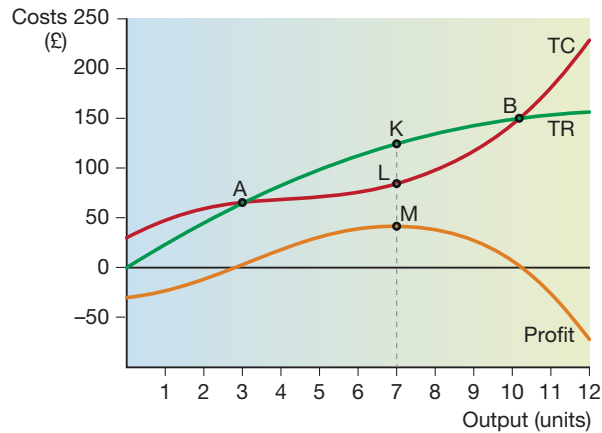
Table 9.1 shows how total revenue, total costs, average revenue and average costs all change as output increases. The revenue data are based on the assumption that the firm is operating in a concentrated market, such as an oligopoly. It faces a downward-sloping demand curve and so as the selling price falls, output rises. The average total cost (ATC) is based on the reasoning outlined in Chapter 6. ATC is high at low levels of output because fixed costs are the same, regardless of output, and so AFC is very high. As output rises, AVC begins to fall, leading to a fall in ATC, but as the firm approaches capacity output, AVC (and thus ATC) rises significantly.

Column 5 shows the profit at each level of output. It is calculated by subtracting total costs (Column 4) from total revenue (Column 3). It is also calculated by $q \times (AR - AC)$ – that is Column 1 multiplied by (Column 6 – Column 7). At zero units of output, there is a loss because fixed costs have to be paid, even if there is no output. For 1 or 2 units of output, some revenue is earned but a loss is still made. As output increases, more specialisation and division of labour leads to greater efficiency and so average costs fall. After the seventh unit, average costs begin to rise, slowly at first but then more noticeably as output gets closer to the firm’s maximum capacity.

It can be seen in Column 5 that profit is negative (a loss is made) if output is between 0 and 2 units because $TR < TC$ (and $AR < AC$). At 3 units of output, profit is zero. From 4 units to 10 units, a profit is made because $TR > TC$ ($AR > AC$). For 11 and 12 units, a loss is made. A profit-maximising firm will choose to produce 7 units of output and earn £42 profit (abnormal profit).

Figure 9.1 shows how profit changes as output increases. A loss is made when output is low because $TC > TR$. At Point A the two lines intersect and so $TR = TC$ and profit is zero. $TR > TC$ between Points A and B, at which point $TR = TC$ again. When output increases beyond Point B, $TC > TR$ again, and so a loss is made.

Figure 9.1 Total revenue, total costs and profit



The highest level of profit is made when the vertical gap between TR and TC is at its greatest. Measured in whole units, this occurs at 7 units of output. Point K shows TR = £126 and Point L shows that TC = £84. Profit at 7 units is £126 – £84 = £42 (shown by Point M).

Figure 9.2 Average revenue, average costs and average profit

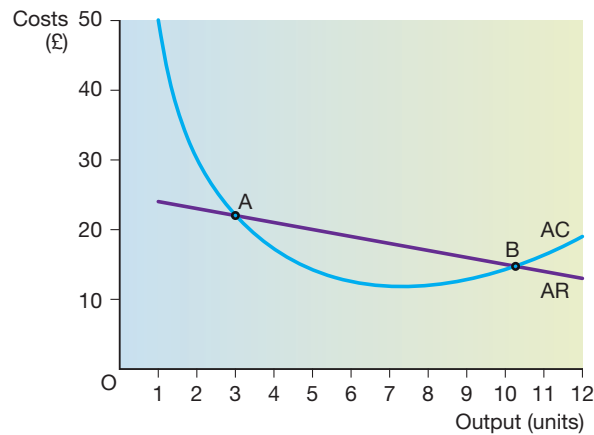


Figure 9.2 shows how AR and AC (ATC) change as output increases. A loss is made when output is low because AC > AR. At Point A the two lines intersect and so AR = AC and profit is zero. AR > AC between Points A and B, at which point AR = AC again. When output increases beyond Point B, AC > AR again, and so a loss is made.

The role of profit in a market economy

Profit plays a number of roles in a market economy, mostly focused on ensuring that the market economy allocates resources efficiently. To an extent, some of these roles overlap. The main roles of profit in a market economy are as follows:

- incentivising the owner and giving signals to indicate which goods and services should be provided. Firms are motivated by profit and will therefore constantly strive to find ways to maximise profit. If demand for a good increases, the price and potential profit is likely to rise too. Firms seeking profit will allocate more resources to the production of these profitable goods;
- reacting to changes in supply. If resources become scarce, their cost is likely to rise, leading to a decrease in market supply. This is likely to lower profit margins

and so firms will take this as a signal either to increase the price or to reduce output of these goods, in order to maintain their profit. If people value these scarce resources, they will be prepared to pay the increased price;

- acting as a signal to new market entrants. Firms that do not operate within a particular market will use changes in profit as a signal for whether it is worth entering that market.

REALWORLD ECONOMICS 9.1

Profits and the supermarkets

German-owned supermarket, Lidl, has announced a record total UK annual revenue of over £4 billion in its most recent accounts – an increase of 21% in one year alone. This growth in total revenue has been accompanied by growing profits. Analysts believe that Lidl and Aldi have identified changes in consumer tastes and behaviour, exploiting these changes to increase both their total revenue and profit levels in the UK.

Figure 9.3 shows how Lidl and Aldi have increased their market share of UK groceries.

The cornerstone of their success has been the limited range of items that they stock. Prior to the recession, UK consumers valued the breadth of choice provided by supermarkets. However, this led to a wide expansion of product ranges. In 2010, an Aldi store would stock 800 items, whereas a large Tesco store would have 50,000 different products

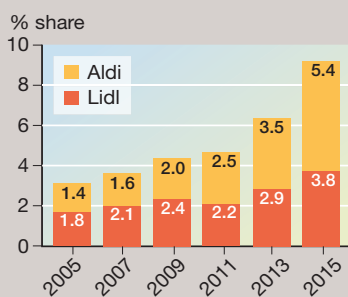


Figure 9.3

from which customers could choose. This situation meant that, for an individual product, Aldi could get a better bulk-buying discount than Tesco, despite the fact that overall Tesco was 25 times bigger.

The recession led to a change in consumer priorities. Low prices, rather than product variety, became the key feature for many consumers. They recognised that while lower-priced goods might often be inferior, they could still offer better value for money.

UK shopping habits have also changed. The traditional, large-scale weekly shop has been replaced by more frequent shopping trips. These trips focus more on fewer but fresher items. The size of many of the larger UK supermarkets was designed to cope with an influx of customers, making large purchases at the weekend. This dramatic change in shopping behaviour means that most supermarkets are working well below their capacity and therefore their fixed costs are much higher than necessary. In contrast, Lidl and Aldi have smaller shops that are used much more intensively (and cost effectively).

At a time when UK supermarkets have suffered declining profit, Lidl and Aldi have more profit to retain, which helps to fund their new store openings.



Lidl and Aldi's combined market share of UK groceries could reach 15% by 2020, according to credit rating agency Moody's

The German discounters are attracting middle-class customers by introducing a deluxe range of own-brand products. More than 31% of their customers are now (2015) from social classes A and B – an amazing increase from only 12% in 2013. However, this has led to a 50% increase in the range of different products sold. At a time when the UK supermarkets are cutting product range, Aldi and Lidl are expanding their ranges.

Sources: Article by Daniel Miller for thisismoney.co.uk, 15.7.15 and other sources

Questions Total: 20 marks

- 1 From the article, identify and explain *two* different roles of profit. (8 marks)
- 2 Has the growth in profit at Lidl and Aldi arisen from supply factors more than demand factors? Justify your view. (12 marks)

- providing a source of finance for investment. Firms that read market signals correctly will make abnormal profits in the short run, before they are competed away. These profits can be retained in the business and used to invest in research and development and new capital equipment;
- rewarding enterprise. Owners and shareholders receive dividends as their share of the profits. Abnormal profits will thus encourage owners and shareholders to direct funds towards profitable companies and will encourage owners and shareholders to provide finance to firms to help the market economy to grow;
- influencing owners of all factors of production. Profitable firms are usually in a position to pay higher rewards to other factors of production, such as high wages to its workers. In this sense, profit will encourage all factors of production to move towards industries that are earning profit;
- rewarding and reducing risk. Profit plays two roles in relation to risk. First, it rewards entrepreneurs who are prepared to take risks because such entrepreneurs can identify opportunities that other firms do not recognise, particularly when it comes to identifying new markets. Secondly, profit can be retained in the business so that there is finance available during economic downturns.

Overall, profit helps economies to operate efficiently because only firms that use resources effectively will continue to supply. Firms that do not identify consumer wishes, or are unable to produce cost effectively, will not survive.

Review questions

Total: 25 marks

Output	Price	Total revenue (£)	Total costs (£)	Profit
1	30		38	
2	24		48	
3	20		54	
4	17		59	
5	15		67	

Table 9.2 Data on Product A

Complete the TR and profit columns in Table 9.2 in order to answer Questions 1 and 2.

- 1 At which level of output does Product A maximise profit? (3 marks)
- 2 At which level of output does Product A make only 'normal profit'? (3 marks)
- 3 A firm's selling price is £6 and it sells 20 units of output. Its profit is £50. Calculate its average total costs (ATC). (2 marks)
- 4 What is the difference between 'normal profit' and 'abnormal profit'? (4 marks)
- 5 In 2015 a firm's accounts show that its total revenue exceeds its total costs by £3000, but the firm decides to cease trading. Briefly indicate what this situation says about the firm's normal profit and its supernormal profit. (3 marks)
- 6 Analyse why profit is so important to the effective operation of the market economy. (10 marks)

Technological change

Key concepts from Year 1

This is a new topic. However, the significance of technological change will be considered in the context of production, productivity and costs of production. These topics were examined in Chapters 12 and 14 of the Year 1 companion textbook. Costs of production have been further examined in Chapter 6 of this book.

This chapter examines the importance of technological change. Initially, it describes the difference between invention and innovation. The impact of technological change on methods of production, productivity, efficiency and firms' costs of production is then considered. The chapter goes on to examine how technological change can lead to the development of new products and new markets, but also lead to the destruction of existing markets. It concludes with a study of how technological change can influence the structure of markets.

The difference between invention and innovation

Invention allows firms to introduce new goods and services into the market, such as the car or the computer. Innovation usually involves the taking of an idea and modifying it, such as laptop computers and mobile phones.

Innovation is often divided into two types:

- **product innovation**, which is the creation and launch of a good or service that is significantly different to an earlier good or service. Product innovation is used to create differentiated products in concentrated markets, enabling firms to benefit from both higher demand and higher prices;
- **process innovation**, which is the creation of a new way of making, providing or delivering a particular good or service. Usually this helps to reduce average total costs (ATC) so that a firm can improve its price competitiveness or earn a higher profit margin per item.

Invention and innovation enable businesses to compete effectively in an increasingly competitive global environment. They are both closely connected to technological change.

Key terms

Invention is the creation of a new idea or product.

Innovation is the successful exploitation of new ideas.

Technological change describes the process of adapting new applications of practical or mechanical sciences to industry and commerce.

Technological change can affect methods of production, productivity, efficiency and firms' costs of production

New inventions and new processes enable economies to make technological progress. In turn, technological progress provides the foundation for many new inventions and innovation.

Technological change creates dynamic markets. In order to stay competitive, new firms must adopt new processes and products. Technological change can have a major impact on markets by introducing new products. New or updated products enable a firm to stay ahead of competitors. In rapidly changing, technologically advanced industries, innovation is vital in presenting new products to the market. In addition, innovation enables firms to balance their product portfolio through the creation of new growth products to replace those that are becoming outdated.

The other main ways in which technological change can influence markets are by affecting methods of production, productivity, efficiency and firms' costs.

Technological change and methods of production

Overall, technological change tends to move production away from labour-intensive methods of production to more capital-intensive production. Although this may involve high expense in the short run, in the long run it should lead to greater productivity because capital can increase the speed of production. Furthermore, new methods of production can improve quality. Thus demand tends not only to increase but also to become more price inelastic. A further benefit of new methods of production can often be seen in greater flexibility. Production processes often allow mass customisation, which results in mass-produced goods that have unique features. For example, modern car plants now use methods of production that allow the firm to produce different variations of the basic model, so that a particular car can be tailor-made according to an individual customer's preferences.

Changing methods of production can also have a significant impact on macroeconomic policies. A shift away from labour-intensive towards capital-intensive methods has led to technological unemployment, since machines have replaced people's jobs. However, the greater efficiency of capital equipment has tended to lead to higher economic growth and raised living standards, and therefore lower levels of cyclical or general unemployment. Technological change has also influenced the distribution of income and wealth. Typically, it has increased the reward received by owners of capital (capitalists) at the expense of unskilled workers. In more recent times, the growing sophistication of capital equipment has led to unemployment amongst more skilled manual and clerical or managerial workers. This has widened inequalities of income and wealth in the UK.

Technological change and productivity

New technology increases the speed of production. This can have wide-ranging effects – not only does it increase the productivity of individual factors of production but it also means that an economy can meet more of the needs of the population. In effect, it is reducing the extent to which scarcity exists.

Technological change can improve the productivity of all factors of production. In industries such as agriculture, new technology has led to a much greater fertility

of land. As a result, a given area of land can produce far more produce. Labour productivity and capital productivity can also be increased by improvements in technology.

Technological change and efficiency

Up until this point we have considered two major measures of economic efficiency: productive efficiency and allocative efficiency.

- **Productive efficiency:** Productive efficiency occurs when average costs are at their minimum point. Technological change, particularly where it involves improving production processes, tends to lead to lower average costs. It therefore helps firms and the economy as a whole to achieve productive efficiency.
- **Allocative efficiency:** The impact of technological change on allocative efficiency is less certain. Improved technological communication systems can overcome or reduce the level of market failure that arises from information failure. However, owners of capital are often the major beneficiaries of improvements brought about by technological change, and so it can lead to a less perfect allocation of resources because of greater levels of inequality.

Technological change and firms' costs of production

High-cost technology

Typically, firms adapting to technological change requires large sums of money to be spent on high-cost technology in the form of items such as research and development and capital equipment. These high fixed costs often lead to much higher average costs for low levels of output because average fixed costs are very high at low levels of output. However, for very high levels of output the average fixed costs are much lower. New technology improves the efficiency of variable factors, such as labour, and so lower variable costs per unit are much lower for all levels of output. Overall, this means that average costs are lower, but the minimum efficient scale is higher, than before the technological change took place.

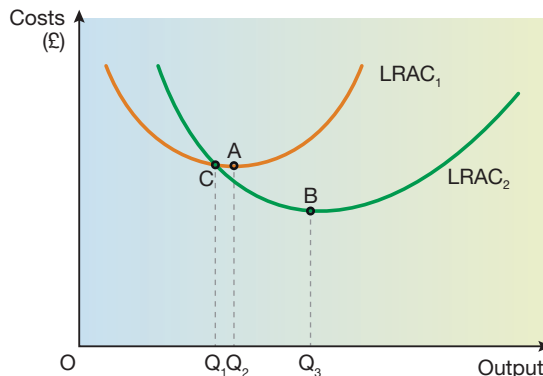


Figure 10.1 *The impact of high-cost technology on a firm's costs of production*

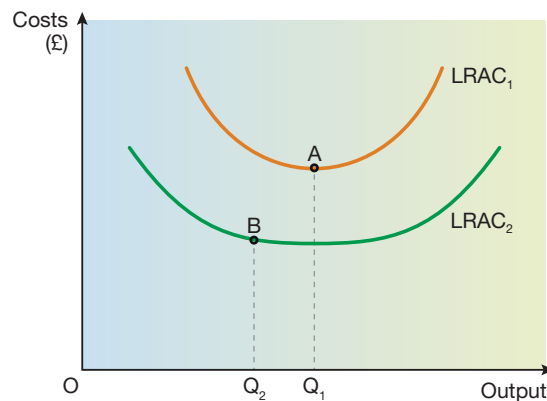
The impact of high-cost technology is shown in Figure 10.1. $LRAC_1$ illustrates the long-run average cost curve prior to the introduction of high-cost new technology. Point A shows the minimum efficient scale of production on this cost curve. The introduction of high-cost technology, leads to a new long-run average cost curve, shown by $LRAC_2$. Point B shows the minimum efficient scale of production on this cost curve. Due to the high expense required, the minimum efficient scale of

production changes from Q_2 to Q_3 as a result of the introduction of new technology. Consequently this change in technology is likely to act as a greater barrier to entry, as firms now need a higher level of output in order to be competitive. For low levels of output, it can be seen that the new technology can lead to increased average costs. Point C shows the point at which $LRAC_2$ and $LRAC_1$ have identical average costs. Thus, for output level Q_1 , there is no difference in the average costs of the two lines. For output levels below Q_1 , $LRAC_1$ shows lower average costs. For output levels above Q_1 , $LRAC_2$ shows lower average costs. Thus, a firm wishing to adopt the new technology must be certain that its output will exceed Q_1 .

Low-cost technology

In more recent times, information and communication systems have become much more accessible and cheaper to acquire. Consequently, in some industries, particularly in the tertiary (services) sector, firms can increase the scale of output without incurring high costs for the new technology that is being used.

Figure 10.2 *The impact of low-cost technology on a firm's costs of production*



The impact of low-cost technology is illustrated in Figure 10.2. $LRAC_1$ shows the long-run average cost curve prior to the introduction of high-cost new technology. Point A shows the minimum efficient scale of production on this cost curve. The introduction of low-cost technology leads to a new long-run average cost curve, shown by $LRAC_2$. Point B shows the minimum efficient scale of production on this cost curve. Due to the low capital expense required, there is a significant fall in average costs at all levels of output. Moreover, since the technology is accessible to firms of all sizes, the minimum efficient scale of production falls from Q_1 to Q_2 as a result of the introduction of new technology. This change in technology means that small firms are more likely to be able to enter the market because fixed costs do not represent such a significant barrier to entry.

Technological change can lead to the development of new products and new markets, and may destroy existing markets

Technological change and the development of new products

We have studied productive efficiency and allocative efficiency. These describe efficiency at a particular point in time and are therefore often classified as static efficiencies. Dynamic efficiency measures whether scarce resources are used efficiently over time, in order to satisfy consumer wants. Technological change is a crucial factor

in achieving dynamic efficiency. It enables firms to invent and innovate and thus meet new and changing consumer wants. However, it is also a consequence of dynamic efficiency. As the pace of change increases, firms need to replace existing products more frequently in order to stay competitive.

Technological change can lead to new products in order to:

- **replace outdated products or processes.** As the pace of change increases, the length of time that specific products remain in the market (their product life cycle) tends to fall significantly. In order to retain competitive advantage in a market, a business must anticipate products and processes that are likely to become outdated and must undertake innovation;
- **increase product ranges.** New technology can allow a business to update its product range continually and stay ahead of competitors. In rapidly changing industries, new product development is vital in the creation of new growth products to take over from those that have reached their decline stage;
- **increase value added.** New technology helps firms to develop goods or services that have unique qualities, allowing them to achieve a higher degree of product differentiation. Products that are clearly differentiated from those of competitors usually sell in higher quantities and their uniqueness tends to lead to more inelastic demand, allowing firms to increase prices and so earn higher profit margins. In addition, if new technology uncovers a new invention, this can be patented, giving the firm a guaranteed monopoly of manufacture for about 20 years;
- **enter new markets.** Technological change can help firms to enter new markets. It enables firms to identify opportunities in new markets if the technology opens up opportunities to release radically different products. In some cases, technological change can lead to the *creation* of new markets.

Technological change and the development of new markets

On occasions, technological change can be so radical that it leads to the development of a new market. Examples of such change include the typewriter, train, car, aeroplane, telephone, radio, television, mainframe computer, laptop computer, tablet, electronic music, photographic film and digital photography.

In most cases, new products establish a place in an existing market. In the car industry, there has been significant technological development over the last 100 years, but fundamentally technological advances just provide new cars that displace older, less technologically advanced models. Thus, inventions such as the telephone tended to act as a complement to traditional communications, such as the postal service. Similarly, television did not replace radio, although it has been a factor in the decline of radio.

Technological change and the possible destruction of existing markets

In the previous section we saw how technological change can lead to the establishment and development of new markets. Sometimes the new markets created by technological change and/or innovation challenge existing markets.

The idea of new markets replacing old markets has been established for many years. In 1942, Joseph Schumpeter introduced the term ‘creative destruction’ to describe this process. He argued that innovation and new technology create new products

Key term

Creative destruction is an incessant process by which innovation and new technology constantly lead to the introduction of new production units that replace outdated ones.

and new markets. However, in doing so, they undermine established products and markets. Often, these new products and markets completely replace the old – the process known as creative destruction. The concept of creative destruction will be analysed in detail in Chapter 18.

Technological change can influence the structure of markets

Technological change can influence the structure of markets in a number of ways.

The number of firms in a market

As explained earlier in this chapter, high-cost technology can severely restrict the number of firms capable of competing in a market. In industries such as car manufacturing, the minimum efficient scale is estimated to be about 2 million cars per annum. As a consequence, industries in which firms rely on expensive technology tend to be oligopoly markets with relatively few competitors. In some cases, such as gas and water supply, it is only economically possible to have one supplier for each household. However, in many markets, such as retailing, new technology is reducing the minimum efficient scale and allowing small firms to compete on more equal terms with larger competitors. In these markets, technological advances are increasing the number of firms in the market and making the markets more competitive.

The degree of homogeneity and differentiation of goods and services in a market

Technological change can often allow businesses to create a unique, differentiated product. In some industries, such as mobile phones, most firms endeavour to maintain this differentiation through methods such as patents and copyright, which prevent other firms from copying the original ideas and designs. For patents, this differentiation can be maintained for a period of 20 years. However, although a particular good or process can be protected for this length of time, rapid changes in technology are likely to make both the good and process outdated within this period. In effect, the firm is unlikely to reap the benefits of a patent for the full 20 years. Copyright has a much longer duration. For books and artistic work, it lasts for 70 years after the death of the author. Once this protection expires, there is no protection and so, if an idea is still valuable, any firm can produce it. This is common practice in the pharmaceutical industry. Once a medicine loses its patent, many firms produce cheap, ‘generic’ alternatives that include the same ingredients. This changes the market from one in which there is a single differentiated medicine to a market with lots of identical or slightly differentiated medicines.

Since markets have become more global, and technology has become more accessible to firms, technological change often leads to more homogeneity of goods and services. This can lead to less differentiation in markets.

Information and knowledge in the market

Information and communications technology has greatly improved information and knowledge in most markets. This has enabled consumers to make more rational decisions because they are fully aware of the goods and services and their prices. Greater information also makes it easier for new firms to enter a market and compete with established firms. However, technological advances have led to a much greater

breadth of knowledge and information globally. As a consequence, it becomes far more difficult for individuals to understand the nature of goods and services being offered. Furthermore, the asymmetric nature of this information tends to make it harder for consumers and gives greater bargaining power to suppliers, who can focus their particular knowledge base on their own specific products.

Barriers to entry

Traditionally, technology has acted as a barrier to entry because only the very largest firms can afford to use the technical equipment necessary to produce at low costs. For industries that are technologically advanced, the minimum efficient scale tends to be much higher, again acting as a barrier to entry for new firms.

However, recent changes in technology have altered this situation. Many online firms now have very little requirement for capital expenditure. Capital costs and sunk costs are very low and so they can enter and exit an industry quite easily, with limited financial risk and little waste.

REALWORLD ECONOMICS 10.1

Creative destruction in higher education

Higher education (HE) is experiencing record growth in many countries. In the USA and Europe, 8.5 million students will graduate this year. HE is also booming in emerging economies, particularly in Asia.

However, as a business its format has changed little in the last 2000 years – groups of young students still meet at appointed times to listen to scholars passing on their wisdom.

In its present state, higher education thrives because university



In higher education, the traditional lecture is under threat from more flexible, online access

degrees add considerable earning power to university students, in comparison to those students who do not progress on to higher education. However, in countries such as the UK, the cost of tuition is making the decision less clear-cut. In many industries, growing income levels have led to labour-intensive organisations becoming uncompetitive. At present, most degrees are still delivered using labour-intensive methods.

Changes in technology are having major impacts on this market:

- the content of degrees is becoming dated more rapidly, placing pressure on graduates to retrain and update their education during their working life. The number of mature students studying for degrees is increasing as a result;
- online learning is enabling students to gain degrees without attending lectures on a full-time basis.

Many people believe that online learning is more suited to the needs of society. It offers more flexibility and can be fitted around other commitments, such as a full-time job.

With education becoming a more global market, there is also pressure to find cheaper ways of delivering degrees. Online learning reduces the cost of delivering degrees and is likely to be used as a competitive weapon.

Will online learning replace traditional degrees? Traditional degrees are unlikely to be able to compete on cost, but, as with many markets, the perception of quality will be critical to the survival of many universities.

*Sources: Article in *The Economist*, entitled 'Creative destruction', 28.6.14 and other sources*

Exercise

Assess the view that higher education, in its present form, is likely to experience creative destruction.

(15 marks)

Competition and cooperation

The global economy enables firms to benefit more fully from invention and innovation because new ideas can be exploited on a much larger scale. In general, these technological advances have encouraged greater competition between firms, as many new technologies are used to improve the efficiency of firms of all sizes. Where the technological advances involve high costs, it has often encouraged firms to cooperate in joint ventures. Global expansion requires technical know-how, but also an understanding of different markets. Firms with technologically advanced ideas will often cooperate with other firms, sometimes even a competitor, because their joint expertise makes them more efficient. As a consequence, technological advance frequently creates a spirit of cooperation as well as competition.

Review questions

Total: 35 marks

- 1 Figure 10.3 shows how technological change has led to a shift in a firm's long-run average costs from $LRAC_1$ to $LRAC_2$.

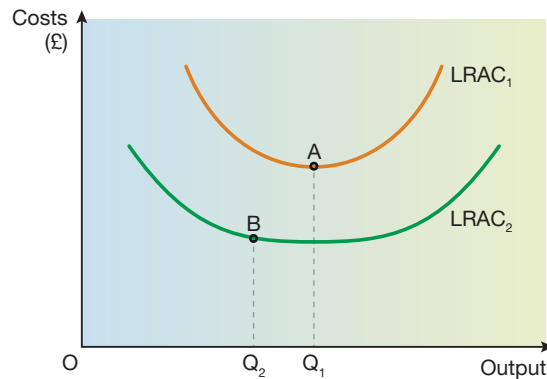


Figure 10.3

This change shows:

- A Increased efficiency and an increase in the minimum efficient scale
 - B Increased efficiency and a decrease in the minimum efficient scale
 - C Decreased efficiency and an increase in the minimum efficient scale
 - D Decreased efficiency and a decrease in the minimum efficient scale (1 mark)
- 2 Define the term 'creative destruction'. (3 marks)
- 3 Distinguish between the terms 'invention' and 'innovation'. (4 marks)
- 4 Explain how technological change can: (4 marks)
- (a) reduce barriers to entry; (4 marks)
 - (b) increase barriers to entry. (4 marks)
- 5 Explain *one* way in which technological change can increase productivity of land. (5 marks)
- 6 Explain why creative destruction tends to result from new entrants to the market rather than established firms. (6 marks)
- 7 Analyse *two* ways in which technological change can influence the structure of markets. (8 marks)

TOPIC 4 Exam-style questions

A-LEVEL PAPER 1

SECTION A Context – Production, costs and revenue

Extract A Manufacturing costs of the Apple iPhone

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015
Average costs (\$) – 2015 prices	250	184	186	193	199	215	218	228	236
Output (millions)	1.4	11.6	20.7	40.0	72.3	125.0	150.3	169.2	231.2

Note: Costs since 2011 relate to models with higher specification.

Extract B Factors influencing manufacturing costs of the iPhone

The iPhone's rapid growth enabled it to benefit from internal economies of scale, particularly in areas such as the bulk buying of components. However, continued expansion led to some coordination difficulties and so diseconomies of scale have emerged, leading to higher average costs for some components. There have also been issues with technical difficulties because some manufacturing processes have not kept pace with other advances in technology. A further issue has been the nature of its suppliers. Samsung provides 26% of the components of an iPhone. As a competitor, it wants Apple to have high costs, but iPhone components provide profit for Samsung.

Apple makes its own software and designs its own chips. It uses its bargaining power and sheer scale to persuade suppliers to offer its components at low cost, as the component suppliers can achieve high labour productivity and also sell so many items when supplying Apple.

In recent years, some components have become more difficult to source because suppliers have found it hard to produce the quantities that Apple requires. The costs of the screen display and applications processors have risen significantly as output has increased (the latter component being one that Samsung supplies).

Sources: www.techrepublic.com, investor.apple.com, recode.net and <http://wallstreetpit.com>

Extract C Apple and creative destruction

Apple's ability to use new technology in order to open up new markets is exemplified by the sales of its three most notable portable devices: the iPod, the iPhone and the iPad. Table A shows the number of these items sold over the last 10 years.

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
iPod	39.4	51.6	54.8	54.1	50.3	42.6	34.2	26.4	14.4	0
iPhone	0	1.4	11.6	20.7	40.0	72.3	125.0	150.3	169.2	231.2
iPad	0	0	0	0	7.5	32.4	58.3	71.0	68.0	54.9

Apple's creativity could be argued to have had a destructive effect on many major businesses. Before the millennium, Sony was the electronics company that others envied. The Sony Walkman was a revolutionary product that displaced portable radios, but when Apple introduced the iPod in 2001, Sony had no response. Its market share of global portable music is now only 6%.

The iPhone was launched in 2007. Although mobile communication was available from products such as the BlackBerry and certain mobile phones, the superior design of the iPhone led to its rapid growth. Firms that did not react to this shift in the market, which encouraged individuals to communicate while on the move, lost sales volume. The most notable example was Nokia. In 1998 it became the market leader for global handset sales and in 2005 it sold its 1 billionth phone. Its success was built on excellent hardware. However, it could not compete with the iPhone's superior software. From a share of over 49% of the smartphone market in 2007, it fell to 3% in 2013. Microsoft bought Nokia for £4.6 billion in 2014; in 2007 it had been worth £150 billion.

Mobile phones, such as the iPhone, have also been recognised as a destructive force behind the upheaval of the photographic industry. Canon and Fuji adopted digital photography quite quickly and they have their own market niche. However, the one-time market leader – Kodak – has been destroyed by digital photography. In 2012 it filed for bankruptcy.

The iPad created a new market, but one that was closely related to the laptop market. Companies such as Dell have struggled to maintain sales as the iPad's sales have increased.

Table A also shows another destructive force. The iPhone has diversified into a range of sizes and specifications. For many buyers, it is their mobile music device that is key and it is no coincidence that the most successful year for iPhones (2015) was the year in which Apple decided to cease production of the iPod (although it has indicated that it will be reintroducing it). Since 2013 iPad sales have also fallen, as people switch from iPads to iPhones.

Questions

Total: 40 marks

- 1 Using the information in Extract A, estimate the likely minimum efficient scale for the iPhone. (2 marks)
- 2 Explain how the data in Extract A might indicate the existence of barriers to entry into this market. (4 marks)
- 3 The data in Extract A suggest an upward-sloping average cost curve for iPhones. Analyse *two* possible reasons why the iPhone might have an upward-sloping average cost curve. (9 marks)
- 4 To what extent have Apple's recent product developments led to creative destruction? Justify your view. (25 marks)

SECTION B Essays

Total: 40 marks

Economists recognise that profit is of vital importance in a market economy. Firms can increase profit by increasing revenue and/or by reducing costs. However, in a contestable market, price rises can lead to lower levels of sales. Furthermore, the U-shaped average cost curve suggests that higher output will lead to higher average costs.

- 1 Explain the importance of the role of profit in a market economy. (15 marks)
- 2 Assess the view that increasing output is not a practical way of increasing profit because it inevitably leads to a fall in average revenue (AR) but an increase in average total costs (ATC). (25 marks)



Topic 5

Perfect competition, imperfectly competitive markets & monopoly

Market structures

Key concepts from Year 1

The concept of market structures was primarily covered in Chapter 17 of the Year 1 companion textbook. The range of market structures was outlined and the factors used to distinguish between different market structures – the number of firms, the degree of product differentiation and ease of entry – were each explained.

Key term

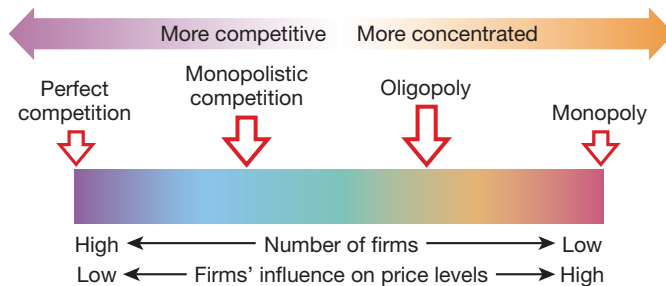
A market structure describes the characteristics of the market that affect the ways in which firms compete and also the welfare of consumers within that market.

This chapter reintroduces the different market structures that are the core of Topic 5. It provides a little more information on the spectrum of market structures, by adding to the concepts covered in Chapter 17 of the Year 1 companion textbook. It is essential that you familiarise yourself with that chapter before reading this one.

The spectrum of competition, ranging from perfect competition at one end to pure monopoly at the other

Figure 11.1 is a reproduction of Figure 17.1 from the Year 1 companion textbook. It shows the range (spectrum) of market structures relevant to the AQA A-level economics specification.

Figure 11.1 *The range of market structures*



It should be noted that the two extremes of the market spectrum – perfect competition and pure monopoly – are largely theoretical concepts.

It should also be noted that duopoly is a market structure that is not shown in Figure 11.1 because it is not included in the A-level specification. It lies between oligopoly and monopoly in the spectrum and describes a market in which there are two competitors. Duopoly markets operate in a very similar (or identical) way to oligopoly and so an understanding of oligopoly provides a foundation for understanding duopoly.

Factors used to distinguish between different market structures

Table 11.1 is a reproduction of Table 17.1 from the Year 1 companion textbook. It provides a summary of the key distinguishing features of the four main market structures.

Factor	Perfect competition	Monopolistic competition	Oligopoly	Monopoly
Number of firms	Many	Many	Few	One
Product differentiation	None	Some	Relatively high	Unique product
Ease of entry/ barriers to entry	No barriers	Few barriers	Many barriers	Very many barriers
Information/ knowledge	Perfect knowledge	Widespread knowledge, but not perfect	Restricted knowledge	Restricted knowledge
Influence of firms on price	Zero	Low	Quite high	Very high

Table 11.1 Factors used to distinguish market structures

The distinguishing features of each market structure impact on other characteristics of the market structures. Some of the key differences and similarities are noted below.

- **Costs:** With large numbers of small firms, perfectly competitive markets and monopolistic competition tend to feature firms producing at low levels of output. Consequently, average costs in these market structures are likely to be much higher than monopoly, where a firm may produce all of the goods in a market within a country. Oligopolists often produce on a global scale and so average total costs tend to be lower because concentrated markets allow more opportunities to achieve economies of scales. However, it should be noted that diseconomies of scale may be more likely too. Moreover, a small concentrated market may not offer opportunities to cut average costs, especially for oligopolists.
- **Demand curve:** In perfect competition, firms must accept the market price. In monopolistic competition, there are many competitors, each producing differentiated products. However, because of the ease of entry into this market and the high volume of alternatives, demand for any individual brand tends to be price elastic. In oligopoly, there are fewer alternatives and so demand is more price inelastic, although the elasticity of demand is also very dependent on the nature of the good – many oligopolists produce non-essential goods, which can have elastic demand. As a market structure, monopoly is the most likely situation in which demand is price inelastic, since by definition there is no alternative product in the market. However, sometimes goods in other markets may compete with a monopolist's good, thus making demand more elastic.
- **Price taker:** In perfect competition, firms must accept the market price and so they are price takers. In other markets, products are differentiated and so firms set the price of their product. However, the scope to increase price is limited by the elasticity of demand and so monopolistically competitive firms face more restrictions (from the marketplace) when setting price, in comparison to oligopolists and monopolists.
- **Profit:** In the short term, all firms can benefit from abnormal profits resulting from favourable changes in their market. However, in the long run new entrants into a market compete this profit away, until only normal profits are earned. Since there is freedom of entry into perfect competitive and monopolistically competitive markets, there is no scope for long-run abnormal profit in these market structures.

Table 11.2 provides a summary of other factors that can distinguish the different forms of market structure. These are based on some of the factors that have been covered in Topic 4.

Table 11.2 Other factors used to distinguish market structures

Factor	Perfect competition	Monopolistic competition	Oligopoly	Monopoly
Average total costs	High	High	Low	Low
Demand curve facing individual firm	Perfectly elastic	Very elastic	Less elastic	Less elastic
Price taker?	Yes	No	No	No
Short-term abnormal profits?	Yes	Yes	Yes	Yes
Long-term abnormal profits?	No	No	Yes (usually)	Yes

Review questions

Total: 25 marks

- 1 In which market structure are there the most barriers to entry?
 - A Monopolistic competition
 - B Monopoly
 - C Oligopoly
 - D Perfect competition

(1 mark)

- 2 Which market structure is the least concentrated?
 - A Monopolistic competition
 - B Monopoly
 - C Oligopoly
 - D Perfect competition

(1 mark)

- 3 Which market structure has differentiated products and a few suppliers?
 - A Monopolistic competition
 - B Monopoly
 - C Oligopoly
 - D Perfect competition

(1 mark)

- 4 Define the term 'market structure'.

(3 marks)

- 5 Explain the difference between long-run profit in a monopolistically competitive market and long-run profit in an oligopoly market.

(6 marks)

- 6 Explain why a perfectly competitive market is more likely to have higher average total costs than imperfectly competitive markets.

(5 marks)

- 7 Explain the difference between the demand curve for an individual firm's good in a perfectly competitive market and the demand curve for a good supplied by a monopolist.

(8 marks)

The objectives of firms

Key concepts from Year 1

The objectives of firms were introduced in Chapter 18 of the Year 1 companion textbook. That chapter contrasted the different objectives that firms might pursue, focusing primarily on profit objectives such as short-run profit maximisation and long-run profit maximisation. It also examined the range of possible objectives that a firm might have, such as survival, growth, quality, maximising sales revenue and increasing market share. It is important that you are able to discuss these objectives and apply them to firms whose aims appear to extend beyond profit maximisation.

This chapter focuses in detail on profit maximisation. It examines the assumptions behind profit maximisation and explains the profit maximising rule ($MC = MR$). The reasons for, and the consequences of, a divorce of ownership from control are examined in the context of firms' objectives. The chapter concludes with a study of the satisficing principle.

The traditional theory of the firm and the assumption that firms aim to maximise profits

Classical economic theory is based on the assumption that all factors of production aim to maximise their satisfaction or reward. Thus, entrepreneurs will want to maximise the profits of their firms in order to maximise their personal rewards.

There are two types of profit maximisation: short-run profit maximisation (which tends to fit most closely with economic theory) and long-run profit maximisation.

- **Short-run profit maximisation:** This objective assumes that firms will try to make maximum profit from every action that they take. In theory, if every action maximises profit in the short run, then maximum profit will be achieved over any other time period. For example, a firm aiming to maximise short-run profit will not provide refunds to consumers who have purchased a product that is the wrong size, because this is an error by the consumer and therefore there is no legal requirement for the firm to provide a refund. Behavioural economists challenge this conclusion. The behaviour of the firm would upset the consumer. As a consequence, they would not perceive this firm so favourably as competitors when deciding on future transactions. As a result, a decision that maximises profit in the short run might endanger long-run profit maximisation.
- **Long-run profit maximisation:** This objective assumes that firms are prepared to make lower profits, and even possibly losses, in the short run, in order to achieve the best possible level of long-run profit. In the example above, many firms would refund the customer in order to build up goodwill and brand loyalty.

The analysis in this chapter will be based on short-run profit maximisation.

In Chapter 9 we examined profit. Profit is calculated by subtracting total cost from total revenue (profit = TR – TC). Table 9.1 and Figure 9.1 respectively showed how profit could be calculated and shown graphically. Figure 9.2 showed how the average profit could be derived from average revenue and average total costs.

Table 12.1 shows the total revenue and total costs of a firm in a perfectly competitive market in which the equilibrium price is £10. The firm is a price taker. The total costs show the normal pattern – relatively high for low levels of output because of the impact of fixed costs. As output rises, increasing returns to scale lead to a fall in the gradient of the TC line, but this gradient then increases as decreasing returns to scale occur.

Table 12.1 Total revenue, total costs and profit for a firm in perfect competition

Output	Total revenue (£)	Total costs (£)	Profit (£)
0	0	18	-18
1	10	25	-15
2	20	28	-8
3	30	30	0
4	40	32	8
5	50	35	15
6	60	39	21
7	70	44	26
8	80	51	29
9	90	61	29
10	100	75	25
11	110	97	13
12	120	127	-7

In Table 12.1 the highest profit is £29 – this is earned from an output of 8 and also 9 units. This £29 is abnormal profit. Since normal profit is included in the total costs, a profit-maximising firm will choose to make 9 units rather than 8, as then normal profit is earned from more units.

In Chapter 1, we looked at the importance of the margin when making decisions. Marginal analysis can pinpoint where an optimum situation (such as profit maximisation) occurs.

The profit-maximising rule (MC = MR)

The concepts of marginal revenue and marginal cost were introduced in Chapters 8 and 6 respectively. Based on the figures in Table 12.2, we can calculate the MR and MC for our perfectly competitive firm.

At zero units of output, the firm will be losing money because it has to pay its fixed costs (£18 in Table 12.1). However, each additional unit made then affects the profit that is shown in the final column. For example, the first good brings in £10 additional revenue and adds £7 to costs. Because MR exceeds MC by £3, this first unit of output contributes £3 towards paying off the fixed costs or making a profit. The second unit has MR = £10 and MC = £3 and so it contributes/adds £7 to the firm’s profit. Each additional unit of output leads to an increase in profit until we reach the ninth unit

Key terms

Marginal revenue (MR) is the addition to total revenue as a result of the sale of one more unit of output.

Marginal cost (MC) is the addition to total costs arising from making one more item.

Output	Marginal revenue (MR) (£)	Marginal costs (MC) (£)	Change in total profit (£)
0	–	–	–
1	10	7	+3
2	10	3	+7
3	10	2	+8
4	10	2	+8
5	10	3	+7
6	10	4	+6
7	10	5	+5
8	10	7	+3
9	10	10	0
10	10	14	–4
11	10	22	–12
12	10	30	–20

Table 12.2 Total revenue, total costs and profit for a firm in perfect competition

of output. The ninth unit brings in £10 additional revenue and adds £10 to costs. At 9 units of output, $MR = MC$. This is the profit-maximising condition and so this firm's profit maximising output is 9 units. If the firm produced a tenth unit of output, the profit would fall because the tenth unit of output brings in an additional £10 in revenue but adds £14 to costs. Therefore profit falls by £4. This situation is confirmed by Table 12.1: total profit falls from £29 to £25 when the tenth unit of output is produced. Figure 12.1 presents these data graphically.

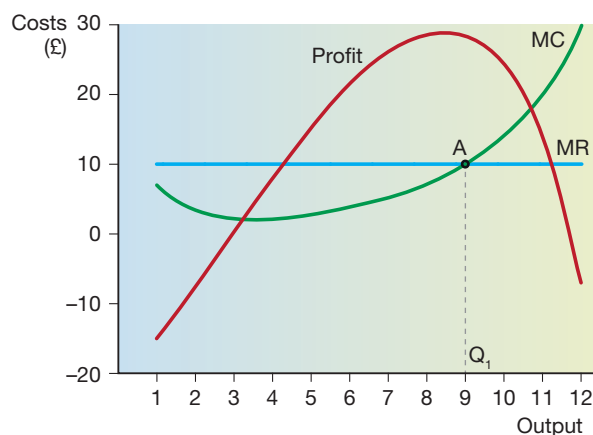


Figure 12.1 Profit-maximisation in perfect competition, using MR and MC

In Figure 12.1, Point A shows where MC cuts MR from below. This occurs at output Q_1 (9 units). A profit-maximising firm will thus produce 9 units of output and sell them at the market price of £10.

What happens in an imperfectly competitive market, with a downward-sloping demand curve?

Table 12.3 shows total revenue, total costs, marginal revenue and marginal costs for a firm in an imperfectly competitive market.

In this market the firm is not a price taker. However, it must accept that it needs

to cut price in order to sell more output because the demand curve (the AR line) is downward sloping from left to right. This means that MR keeps falling as output rises (and becomes negative from the eleventh unit of output in this example).

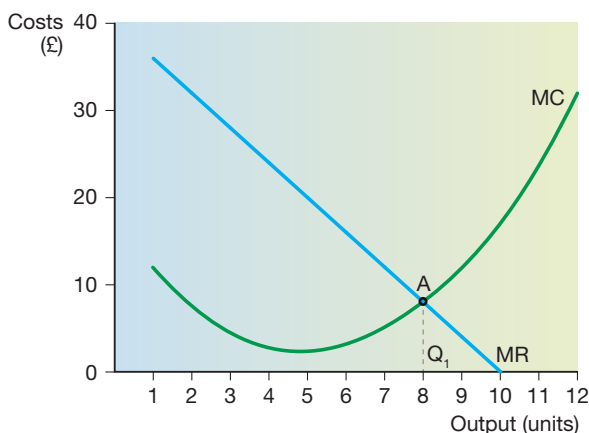
Table 12.3 Revenue, costs and profit in an imperfectly competitive market

Output	Price (AR) (£)	Total revenue (£)	Total costs (£)	Profit (£)	Marginal* revenue (MR) (£)	Marginal costs (MC) (£)	Change in total profit (£)
0	38	0	30	-30	-	-	-
1	36	36	42	-6	36	12	+24
2	34	68	49	19	32	7	+25
3	32	96	53	43	28	4	+24
4	30	120	55	65	24	2	+22
5	28	140	57	83	20	2	+18
6	26	156	60	96	16	3	+13
7	24	168	65	103	12	5	+7
8	22	176	73	103	8	8	0
9	20	180	85	95	4	12	-8
10	18	180	102	78	0	17	-17
11	16	176	125	51	-4	23	-27
12	14	168	157	11	-8	32	-40

Note: *It can be seen that the MR line slopes downwards twice as steeply as the AR (price) line.

Using marginal analysis, the firm will increase its profit by £24 from the first unit (which has MR of £36 and MC of £12). As output rises, MR > MC for each unit of output up to the seventh unit. However, the eighth unit is where MR = MC and so profit is maximised at this output. In Figure 12.2, Point A shows where MR = MC and Q₁ shows the profit-maximising output (8 units).

Figure 12.2 Profit maximisation in an imperfectly competitive market, using MR and MC



The reasons for, and the consequences of, a divorce of ownership from control

Traditionally, the entrepreneur undertook two entrepreneurial functions: ownership and control. The owner(s) of a firm would also be the individual(s) who managed the activities of the firm. In the case of sole traders and partnerships, the owners and managers tend to be the same person. Consequently, the two entrepreneurial

functions remain with the entrepreneur(s) – the sole trader or partners – and so there is often little or no divorce of ownership and control in these smaller firms. However, in public limited companies (PLCs) the divorce is prevalent. As limited companies form the major part of the UK economy, this means that the divorce of ownership and control is widespread.

Reasons for the divorce of ownership from control

- As firms grow in size, their management becomes more complex. It is therefore vital to employ specialist managers rather than encourage owners (who may have no managerial qualifications or capabilities) to organise the firm.
- The growth of firms, particularly public limited companies, means that much more finance is needed from shareholders. In order to obtain this finance, companies need to be able to attract finance from people who are interested in taking a share of the profit but who have no interest in getting involved in the firm.
- Institutional investors are organisations that hold money on behalf of others. For example, insurance companies will hold customers' premiums until there is a requirement to pay people who have suffered insured losses. There are many large organisations that are prepared to buy shares as a way of storing funds but also as a way of increasing their value. Institutional investors do not like to become involved in the management of firms in which they own shares. However, they are very keen to ensure that the organisation is controlled by skilled managers.
- As a company grows, managers and shareholders become increasingly separated because shareholders find it difficult to access the information needed to assess the quality of managers' decision making. This gives greater autonomy to managers.
- The divorce also provides checks and balances to ensure that the business operates smoothly. If owners make the decisions, there is no second point of view. However, shareholders can use general meetings in order to monitor the quality of managers' control of the organisation.

Consequences of a divorce of ownership from control

- The divorce creates a principal–agent problem, whereby one person (the agent) is able to make decisions on behalf of another person (the principal). The agents are the managers and the principals are the shareholders. The problem occurs when the agent makes decisions that are not in the interests of the principal.
- By giving control to people who see the management of the firm as a long-term career, decision making is more likely to take a long-term view. In contrast, owners and shareholders are often accused of focusing purely on short-term profits.

Key terms

Ownership is the provision of finance to a firm in return for a share of profit. It involves undertaking financial risks.

Control is the actual management of a firm and the organisation of decision making within it.

Divorce of ownership and control is the separation of the two entrepreneurial functions, so that control is exerted by individuals who are not necessarily the main owners of the firm or organisation.

The **satisficing principle** is when firms aim for minimum acceptable levels of achievement in terms of profit, revenue and market share.

- Managers are more likely to take a stakeholder view of decisions because they have to work alongside these other stakeholders. Consequently, the objectives of a firm are more likely to be adapted to the needs of other people with an interest in the firm, such as workers, customers, suppliers and the local community.
- Managers may prioritise the needs of one particular stakeholder – the managers themselves. This can lead to objectives that are not in the best interests of the firm.

REALWORLD ECONOMICS 12.1

Banks and the divorce of ownership and control

The banking crisis of 2008 reawakened debate about the divorce of ownership and control in UK limited companies.

Traditionally, banks competed with building societies. Building societies were non-profit-making organisations owned by their customers, with one vote per shareholder (each customer was a shareholder), and so they had a reputation for looking after the interests of customers first. However, in the 1980s and 1990s most of the large building societies converted into banks, which merged or were taken over by the larger banks. This led to an increased focus on profit as the banking industry's main objective.



An effigy depicting 'fat cat bankers', November 2009

Relaxation of controls of banks in the 1980s also led to a change in their objectives. From being non-risk-taking organisations that tended to follow a policy of 'satisficing', banks began to target higher and higher profits. As the nature and complexity of banking activities and their scale of operations changed, the divorce of ownership and control became more pronounced. Owners, many of whom were institutional investors (who own over 60% of bank shares), were happy if profits increased. In order to increase their size, and thus benefit from economies of scale, many banks appeared to focus on sales maximisation in the belief that this would in turn lead to profit maximisation. As a consequence, managers in the investment divisions of the major banks were rewarded for their ability to make profit while other employees were rewarded for succeeding in increasing sales revenue through activities such as payment protection insurance (PPI).

This led to a situation in which many managers earned bonuses from activities that were both high risk and, in some cases, of questionable ethical behaviour. Banks argued that these bonuses were a reward for the managers' success in boosting banks' profits.

Problems arose from information

and communication difficulties, which led to banks accepting responsibility for loans that earned high interest rates but which had very high risks of default. When the recession led to default on many of these debts, the banks were unable to survive and many relied on government bailouts. The government thus became a major shareholder of some banks, notably RBS and Lloyds, though it is currently (2016) trying to sell those shares as quickly as it can but constrained by the potential financial losses of selling them while the banks' share prices are still low. The government is having difficulties in persuading banks to reduce their culture of paying high bonuses to managers. In the economy as a whole, bonuses represent about 5% of incomes; in the banking system, they represented over 33% prior to the recession and, in 2014, bonuses were still as high as 24%. Banks now argue that these bonuses are needed to retain key staff.

Exercises Total: 15 marks

- 1 Explain two reasons for the divorce of ownership and control in the banking industry. (6 marks)
- 2 Analyse why this divorce of ownership and control still exists in banking, despite government intervention in the banking system. (9 marks)

- The checks and balances, mentioned earlier, can slow down decision making because owners may want to ratify strategic managerial decisions. This may lead to lost opportunities.

Overall: The divorce of ownership and control does make it more likely that a firm will have a variety of different objectives and not focus solely on profit maximisation. One of these objectives, profit satisficing, is considered in the next section.

The satisficing principle

The behavioural economist, Herbert Simon, introduced the satisficing principle. He argued: ‘We must expect the firm’s goals to be not maximising profits but attaining a certain level or rate of profit, holding a certain share of the market or certain level of sales’. Simon argued that it was not possible for a firm to know whether its profits were maximised. Simon used the word ‘satisficing’ to indicate that firms would need a profit target, but that maximisation of profit was not a realistic goal.

The satisficing principle has been criticised because of its ambiguity. There is no definition of what a satisfactory level of profit would be. There is also no indication of the balance between the different goals: making profit, achieving a certain market share and achieving a certain level of sales.

Profit maximisation is based on the assumption that firms’ aims are governed solely by their owners. However, the divorce between ownership and control in most PLCs means that they will also need to meet the needs of other people (stakeholders) who have an interest in the firm, such as customers, the local community, government, workers and suppliers.

These pressures from other stakeholders in the firm are likely to lead to a degree of compromise between profit and other objectives. This will in turn lead to a situation where the main financial objective of the firm is to make a satisfactory profit, rather than maximising profit. In practice, shareholders are unlikely to have sufficient understanding of the day-to-day activities of the firm to recognise whether the profit achieved is the maximum possible profit.

Review questions

Total: 30 marks

- 1 Explain why profit maximisation occurs when marginal costs equal marginal revenue. *(5 marks)*
- 2 What is meant by the term ‘divorce of ownership and control’? *(3 marks)*
- 3 Explain *two* possible reasons for the growth of the ‘divorce of ownership and control’. *(8 marks)*
- 4 Explain *two* possible consequences of the growth of the ‘divorce of ownership and control’. *(8 marks)*
- 5 Explain why the desire to meet stakeholders’ needs might lead to firms following the ‘satisficing principle’. *(6 marks)*

Perfect competition

Key concepts from Year 1

Competitive markets were introduced and explained in Chapter 19 of the Year 1 companion textbook. The model of perfect competition was discussed, with an explanation of the assumptions made and how market equilibrium is determined by the interaction of supply and demand in a perfectly competitive market.

This chapter focuses on equilibrium from the perspective of the individual, profit-maximising firm. It demonstrates the formal diagrammatic analysis of the perfectly competitive model, showing both short-run and long-run equilibria. The chapter then studies the implications of the following key assumptions of the model: large numbers of producers, identical products, freedom of entry and exit, and perfect knowledge. These implications are studied in terms of their impact on the behaviour of the individual firms and the industry as a whole. The model is used to demonstrate why firms operating in perfectly competitive markets are price takers. The chapter concludes by examining the proposition that, given certain assumptions, perfect competition will result in an efficient allocation of resources.

The formal diagrammatic analysis of the perfectly competitive model in the short and long run

Key terms

A **firm** is an individual business unit providing particular goods and services, such as Next or Ford.

An **industry** comprises a group of competing firms that provide the same goods and services, such as clothing retailers or car manufacturers.

Perfect competition exists where the following conditions exist:

- There are many buyers and sellers, none large enough to influence market price.
- All products are identical (homogeneous).
- There are no barriers to firms wishing to enter or exit the market.
- Buyers and sellers have perfect knowledge.
- Firms aim to maximise profit; consumers aim to maximise satisfaction.

(These assumptions were explained in Chapter 19 of the Year 1 companion textbook.)

The way in which a perfectly competitive market operates in the short run can be explained through Figures 13.1 and 13.2.

In a perfectly competitive market, no individual firm is large enough to affect market price. Equilibrium market price is set by market supply and market demand, as shown in Figure 13.1. The equilibrium point is where $D = S$, and so equilibrium *market* price is P_E . Individual firms in a perfectly competitive market are **price takers** – every unit of output in the market is sold at price P_E . Figure 13.2 shows short-run equilibrium for the individual *firm* within this market. Note how the output for the firm is measured in units, whereas the market output is expressed in millions of units.

In Figure 13.2, all of the firm's output is sold for P_E . Thus the demand for the individual firm's output is shown by a horizontal line at P_E . Every unit that the individual firm

produces is sold for P_E . This means that the average revenue (AR) for each unit sold by an individual firm is P_E , and so the horizontal line from P_E also represents the average revenue (AR line). Furthermore, every additional unit is sold by the firm for P_E , so the firm's MR is also P_E for each unit of output produced.

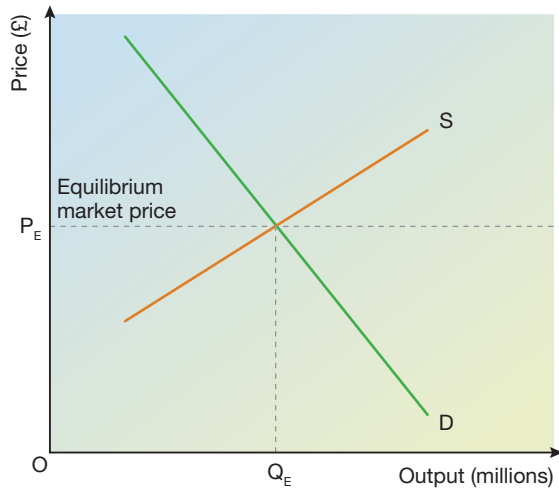


Figure 13.1 Perfect competition – market equilibrium in the short run

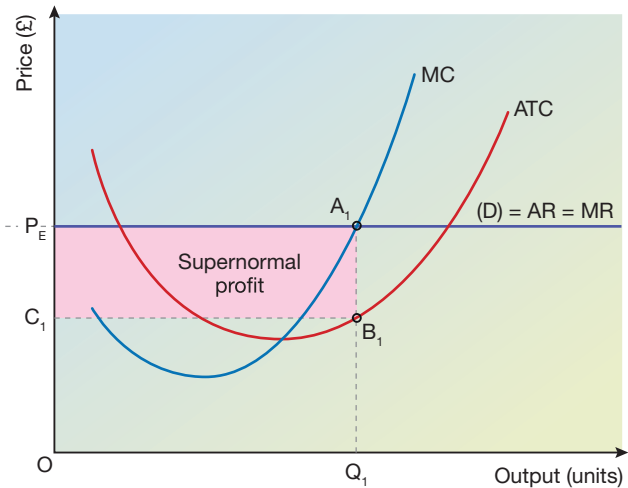


Figure 13.2 Perfect competition – the single firm's equilibrium in the short run

As a price taker, the firm must accept the market price (P_E) but can choose its level of output. In order to maximise profit, it chooses output where marginal cost equals marginal revenue, with marginal cost cutting the marginal revenue line from below. In Figure 13.2 this is at Point A_1 and so equilibrium output for this firm is OQ_1 . If the firm had produced one more unit of output, MC would have risen above P_E and so exceeded MR, causing total profit to fall slightly. Similarly, if the firm had produced less than OQ_1 , then it would not have maximised profit because it would have failed to produce some units for which MR exceeded MC.

$$\text{Profit} = \text{Total revenue (TR)} - \text{Total costs (TC)}$$

$$\text{TR} = p \times q$$

Therefore $\text{TR} = OP_E \times OQ_1$. This is shown by area $OP_EA_1Q_1$.

$$\text{TC} = \text{ATC} \times q$$

Therefore $\text{TC} = OC_1 \times OQ_1$. This is shown by area $OC_1B_1Q_1$.

Total profit is shown by area $C_1P_EA_1B_1$.

(Note that as normal profit is included in costs, this profit is 'supernormal profit'.)

Figures 13.1 and 13.2 show the *short-run* equilibrium for the market and individual firm respectively.

In the long run, the existence of supernormal profits and absence of barriers to entry into a perfectly competitive market attracts new firms into the market. Although a single firm cannot affect market supply, the entry of many new firms will increase market supply and thus affect the market equilibrium price. Figure 13.3 shows the long-run market equilibrium in a perfectly competitive market. Figure 13.4 shows the long-run equilibrium for a single firm.

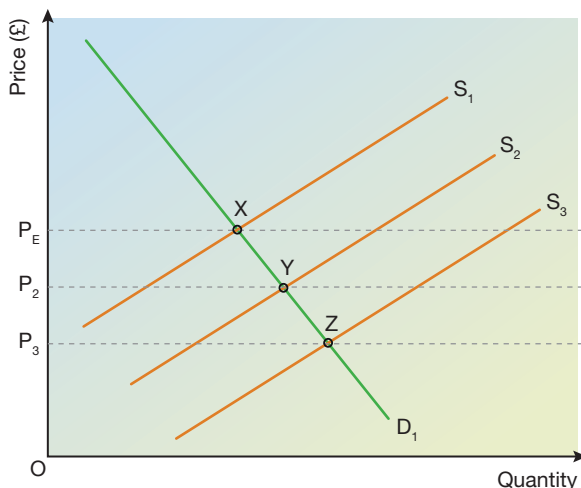


Figure 13.3 Perfect competition – market equilibrium in the long run

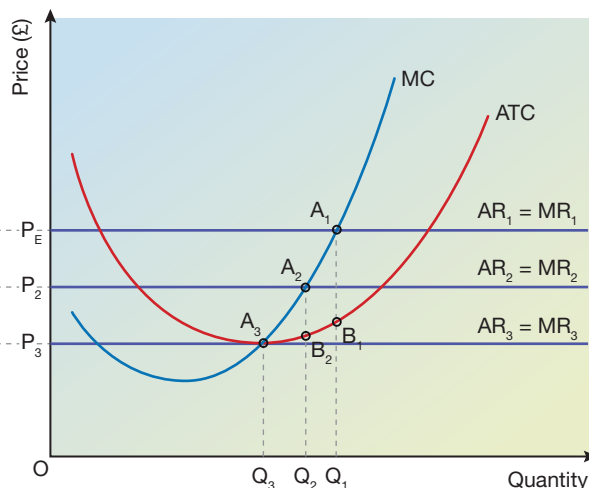


Figure 13.4 Perfect competition – the single firm's equilibrium in the long-run

In Figure 13.3, D_1 and S_1 are the original demand and supply lines (as shown in Figure 13.1). The original equilibrium for the *market* is shown by Point X, giving a market equilibrium price of P_E . Since this price enabled individual firms to earn supernormal profits (as shown in Figure 13.2), new firms will enter the market in order to earn supernormal profits for themselves. Although an individual firm is too small to affect the market supply, the entry of a large number of new firms will increase the market supply. In Figure 13.3 this is shown by the increase in supply from S_1 to S_2 . As a result, the equilibrium point changes from X to Y and so the equilibrium price falls from P_E to P_2 . As price takers, individual firms must accept this new, lower price. Figure 13.4 shows the impact of this fall in price on the individual firm. The firm in the market is now receiving a lower average revenue (AR), shown by AR_2 , and also the same, lower marginal revenue (MR), shown by MR_2 . At price P_2 , the firm's equilibrium point is A_2 (where MC cuts MR_2 from below) and so the firm produces OQ_2 units. At this output, ATC is B_2 and so profit per unit is $A_2 - B_2$. Supernormal profit has fallen but still exists and so new firms continue to enter the market, causing a further increase in market supply. In Figure 13.3 this is shown by line S_3 . When supply shifts as far as S_3 , the new equilibrium point in the market is Z, giving an equilibrium market price of P_3 . In Figure 13.4 it can be seen that the individual firm's AR line is now A_3 . At price P_3 the new MR line (MR_3) intersects MC at Point A_3 and so Q_3 units are produced. When the firm's output is Q_3 , AR_3 and AC are equal, so no supernormal profit is made (but normal profit is earned because this is one of the firm's costs). The absence of supernormal profit means that new firms no longer enter the market and so P_3 is the long-run equilibrium price.

The implications of large numbers of producers

In a perfectly competitive market there are so many producers (sellers) that no single firm (seller) has any influence on the equilibrium price in the market. Although each seller contributes to the overall market supply, the sheer volume of firms within the market, and thus the small scale on which these firms operate, means that no single firm supplies sufficient goods to have a significant impact on the market supply. Consequently, no single firm has an influence on the equilibrium market price. In

effect, each firm is a price taker. In the long run, firms will make only normal profits. However, this situation provides an incentive for firms to improve the efficiency with which they use factors of production. If an individual firm finds a better method of production, then it can lower its average costs to a level below that of its competitors. This will allow it to reap supernormal profits because it will be selling goods at the same price as its competitors. However, due to perfect knowledge, competitors will be able to copy this new production technique and hence, in the long run, the firm's supernormal profit is eroded. From an industry perspective, this opportunity for short-run supernormal profit encourages firms to look for better ways of using scarce resources, and so it encourages greater productive efficiency within the economy as a whole.

The high number of competing firms also encourages firms to spot market changes. Any firm that can anticipate a change in market demand before its competitors will be able to take advantage of the opportunity to sell more goods and probably at higher prices, if this change has led to an increase in demand. Once again, perfect knowledge means that competitors will react to this short-run opportunity to gain supernormal profits and the industry will just return to normal profits again. However, it does mean that there is an incentive to strive to find new ways of satisfying consumers by identifying changes in their tastes.

The implications of identical products

In a perfectly competitive market there is no product differentiation. This means that all firms produce identical (homogeneous) goods. There is no branding or loyalty to a particular firm's goods. Agriculture is often taken to be a close proximity to a perfectly competitive market, but because agricultural produce can vary in quality, product differentiation is often the reason why prices may vary from one supplier to another.

Identical products in a market mean that each product is a perfect substitute for the other products in the market. This leads to the cross elasticity of demand between products being infinite. If a firm increases its price, it will lose all of its sales to its competitors. Furthermore, it will mean that if a firm decreases its price, it will not gain any extra sales volume. Every good can be sold at the market equilibrium price and so a firm that charges one pence less for each good will be reducing its revenue by one pence for each good sold. Since suppliers are rational and aim to maximise profits, they will all charge the market equilibrium price.

Identical products also mean that there is no real advantage to be gained from marketing, and competition is not focused on price because it is set by the market. This can lead to a situation in which markets stagnate because a perfect market is not dynamic – the product does not develop.

The implications of freedom of entry and exit

This is an important assumption of the perfectly competitive market. Rational behaviour would suggest that firms will wish to enter a market in which supernormal profit is being made by suppliers and will wish to exit a market in which losses are being made. This assumption leads to the logical conclusion that scarce resources within an economy will shift from unprofitable uses to profitable purposes.

In Figures 13.1 to 13.4 we looked at situations in which firms would be encouraged

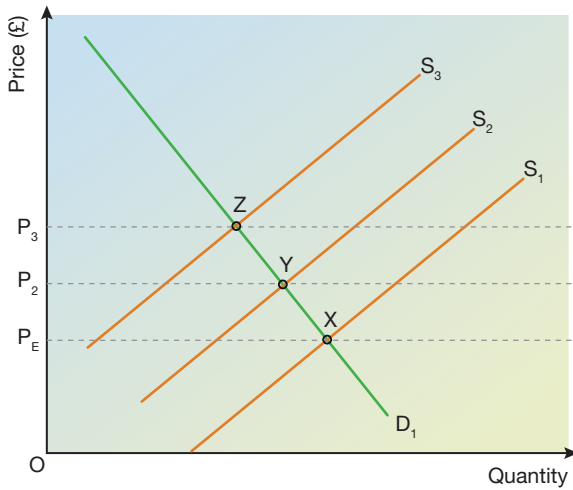


Figure 13.5 Perfect competition – market equilibrium with firms exiting the industry

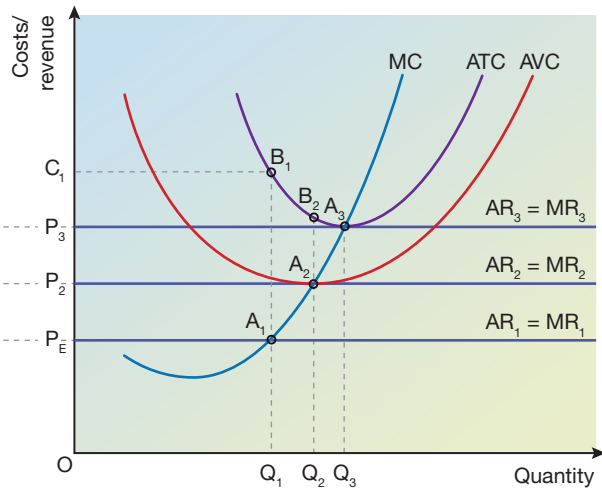


Figure 13.6 Perfect competition – the single firm's equilibrium when losses are made

to enter an industry because supernormal profits are made. Figures 13.5 and 13.6 show the situation in which firms might exit an industry.

In Figure 13.5, D_1 and S_1 are the original demand and supply lines. The original equilibrium for the *market* is shown by Point X, giving a market equilibrium price of P_E . Figure 13.6 shows the resultant equilibrium situation for an individual firm. Profit is 'maximised' at Point A_1 , where MC cuts MR_1 from below. The equilibrium output for the firm is OQ_1 .

$$\text{Profit} = \text{Total revenue (TR)} - \text{Total costs (TC)}$$

$$\text{TR} = p \times q$$

Therefore $\text{TR} = OP_E \times OQ_1$. This is shown by area $OP_EA_1Q_1$.

$$\text{TC} = \text{ATC} \times q$$

Therefore $\text{TC} = OC_1 \times OQ_1$. This is shown by area $OC_1B_1Q_1$.

This means that a *loss* is made. The loss per unit of output is the distance B_1A_1 . The total loss is shown by the area $P_EC_1B_1A_1$.

Key term

Sunk costs are costs that have already been incurred and cannot be recovered. An example is an advanced rental payment for a property that has been paid, even if the property is vacated before the rental contract expires.

A loss means that firms will *not* be reaching their normal profit level. Since perfect competition assumes that there are no barriers to exit, such as 'sunk costs', firms will leave the industry in significant numbers. Although an individual firm is too small to affect the market supply, the exit of a large number of firms will decrease the market supply. In Figure 13.5 this is shown by the decrease in supply from S_1 to S_2 . As a result, the equilibrium point changes from X to Y and so the equilibrium price rises from P_E to P_2 . As price takers, individual firms accept this new, higher price. Figure 13.6 shows the impact of this rise in price on the individual firm. The firm in the market is now receiving a higher average revenue (AR), shown by AR_2 , and also the same, higher marginal revenue (MR), shown by MR_2 . At price P_2 , the firm's equilibrium point is A_2 (where MC cuts MR_2 from below), and so OQ_2 units are produced by the firm. At this output, ATC is B_2 and so a loss is still made (the loss per unit of output is now A_2-B_2). The loss has fallen but still exists and so firms will continue to exit the market, causing a further decrease in market supply. In Figure 13.5 this is shown by

line S_3 . When supply shifts as far as S_3 , the new equilibrium point in the market is Z, giving an equilibrium market price of P_3 . In Figure 13.6 it can be seen that the individual firm's AR line is now AR_3 . At price P_3 the new MR line (MR_3) intersects MC at Point A_3 and so Q_3 units are produced. When the firm's output is Q_3 , AR and AC are equal, so no loss (or supernormal profit) is made. However, normal profit is earned because this is one of the firm's costs. The existence of normal profit means that firms will no longer exit the market and so P_3 is the long-run equilibrium price.

Key note

An assumption of the perfect competition model is that there are no 'sunk costs', such as fixed costs that are paid in advance or already committed. However, this is often not true in real life. For example, a firm ceasing production would still need to pay rent on its factory for the period of time stated in the rental contract. This has an impact on whether firms exit an industry. If a firm has already paid its fixed costs, then it should continue to produce as long as its average revenue is equal to or greater than average variable costs. In Figure 13.6, when price is P_2 the individual firm's average revenue (AR_2) is equal to its average variable costs (AVC). Thus, if fixed costs have already been paid, it would be sensible for this firm to stay in the industry (but only until the fixed costs are due for payment). If the price reaches P_3 before the fixed costs are due, the firm will stay in the industry because at P_3 it is making normal profit.

The implications of perfect knowledge

The model of perfect competition assumes that buyers and sellers have perfect information and knowledge. As the products are identical (homogeneous), the main significance of this assumption is that it means all buyers and sellers are aware of prices being charged and costs of production. If a firm attempts to charge more than the market price, then all buyers will know that they can purchase the products more cheaply elsewhere and so this firm will not sell any goods. If the firm sets its price below the market price, then it will sell all the products that it supplies. However, as firms in a perfectly competitive market can sell all of their output at the market price, it is irrational for a firm to charge a lower price when the result of this action is lower revenue (selling the same quantity of goods but at a lower price).

The effect of perfect information means that a market has only one price at any particular moment in time – the market equilibrium price.

Perfect knowledge also means that, in the long run, competitors copy changes in production techniques. Any short-run benefits of lower production costs disappear in the long run. However, this does mean that the industry as a whole improves its productive efficiency.

Firms operating in perfectly competitive markets are price takers

In a perfectly competitive market, no single seller has any influence on the equilibrium price in the market. Each firm contributes to the overall market supply, but the volume of sellers within the market means that no single firm supplies sufficient goods to have a significant impact on market supply. Consequently, no single seller (firm) has an influence on the equilibrium market price. All the firms supplying the goods are price takers.

Key term

A **price taker** is a firm that has to accept the equilibrium price set by the market in which it operates. Any quantity of goods that it supplies will be sold at the market price.

Perfect competition results in an efficient allocation of resources

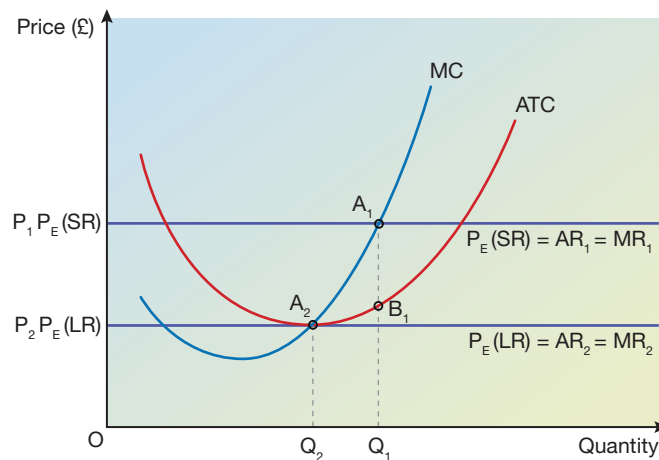
Economics is concerned with using scarce resources to satisfy unlimited wants. Two ways to measure whether this task is undertaken efficiently are productive efficiency and allocative efficiency.

Productive efficiency occurs when production is at the lowest possible ATC.

Allocative efficiency measures whether the goods and services produced are those that consumers want. The demand schedule shows the value placed on a product by consumers. The marginal cost line shows the value of resources needed to produce each additional product. If $P > MC$, then society benefits because the value placed on the product by the consumer is more than the cost of the resources used to make it. However, if $P < MC$, then the resources used to make a good were more valuable than the benefit to the consumer. Thus allocative efficiency occurs where $P = MC$.

Figure 13.7 can be used to illustrate the question: ‘Does perfect competition lead to efficiency?’

Figure 13.7 Perfect competition and productive and allocative efficiency



In Figure 13.7, P_1 shows short-run equilibrium price, with AR_1 and MR_1 showing short-run revenue. The equilibrium point is A_1 , with output Q_1 . At A_1 , $MC = MR_1$ and so allocative efficiency is achieved in the short run. However, average cost is B_1 , which is slightly higher than the lowest ATC, shown by A_2 . Therefore productive efficiency has not been achieved.

In the long run, price falls to P_2 , with the equilibrium at A_2 . Allocative efficiency is achieved because $MR_2 = MC$ at this point. Productive efficiency is also achieved because A_2 is the minimum point of ATC. However, this conclusion is based on assumptions, such as the absence of externalities.

In theory, perfect competition does lead to both productive and allocative efficiency in the long run. However, small firms are unlikely to achieve the economies of scale of large firms. The ATC for a small firm is likely to show much higher levels of costs than the ATC for a much larger firm, and so A_2 does not show the minimum possible cost of production within an economy, unless there is no scope for internal economies of scale.

Most critically, individual demand ignores externalities and does not consider whether goods are merit goods or demerit goods. Furthermore, demand is based on

income and so it encourages the allocation of resources to the rich at the expense of the poor. Allocative efficiency should be based on social benefits and social costs, which consider these externalities as well as the individual consumers' demands and market supply. The 'best' allocation of resources is therefore not where $P = MC$, but where social benefits and social costs are considered. The ideal allocation of resources, where marginal social benefit equals marginal social cost ($MSB = MSC$), will be dealt with diagrammatically in Topic 8.

Conclusion

Perfect competition, in both product and labour markets, provides a yardstick for judging the extent to which real world markets perform efficiently or inefficiently, and the extent to which a misallocation of resources occurs. If the conditions above apply ($P = MC$ and ATC is minimised), then we have the potential for allocative efficiency and productive efficiency. However, allocative efficiency will only be achieved if there are no market externalities (which is highly unlikely). Similarly, productive efficiency only occurs if the industry as a whole is minimising costs (which is unlikely in an industry with many small firms).

REALWORLD ECONOMICS 13.1

Is Billingsgate Market a perfect market?

Economists use the model of perfect competition as a template. It is an idealistic model that is considered to produce a perfect allocation of resources (no market failure) if it operates correctly. However, in practice it is highly doubtful that there is such a thing as a perfectly competitive market, particularly in a complex economy such as the UK.

Local markets in rural communities are perhaps the

markets that approximate most closely to perfect competition. The growth of the supermarkets has meant that most farmers sell directly to large buyers, and so the traditional weekly market in the nearest town no longer sells all of the local agricultural products.

Some large-scale markets have survived, with two notable examples being New Covent Garden Market, which specialises in fruit and vegetables, and Billingsgate Market,

which specialises in fish.

Both of these markets are located in London, and so transport enables traders to come from a large geographical area. Most of the UK's ports supply some of the fish traded in Billingsgate Market. The buyers tend to be knowledgeable retailers of fish, with many London restaurants buying their fish directly from Billingsgate. It is also open to individual customers. In order to ensure that the produce is fresh, Billingsgate only trades on five days a week, from 4am to approximately 8am.

Perfect competition exists where the following conditions exist:

- 1 There are many buyers and sellers, none large enough to influence market price.
- 2 All products are identical (homogeneous).
- 3 There are no barriers to firms wishing to enter or exit the market.



Billingsgate Market – an example of perfect competition in practice?

- 4 Buyers and sellers have perfect knowledge.
- 5 Firms aim to maximise profit; consumers aim to maximise satisfaction.

Do these conditions exist at Billingsgate?

1 There are over 100 stalls or stores, operated by 40 different traders. Although this may not be considered to be a large number, the close proximity of the 40 different traders in one trading hall means that no single trader can influence the market price. Most large-scale buyers of fish, such as supermarkets, buy directly from fishing fleets. Therefore, in Billingsgate there are very many buyers, none of whom has a

significant impact on the market's annual sales of £200 million of fish.

- 2 Fish are not a homogeneous product and so this condition is not met by Billingsgate market. However, it could be argued that variances in quality are fairly small when buying in large quantities. Over 150 different species of fish and shellfish are traded in the market.
- 3 The market is open to any buyer that wishes to purchase fish (although the opening times may restrict some buyers). There are a limited number of stalls and so there are barriers to new entrants.
- 4 Knowledge is a feature of this market. The sellers are specialist companies and the buyers are often

fish restaurants that have a very keen understanding of the fish they are buying.

- 5 Sellers aim to make profit and buyers seek to satisfy their needs, although it is difficult to ascertain whether maximisation of profit or satisfaction is being targeted.

Exercise

- 1 To what extent does Billingsgate meet the conditions of a perfect market? (10 marks)

Discussion point

As a group, choose an alternative market and show why you believe it to be a better (or poorer) example of perfect competition than Billingsgate. (Aim to find a better example, if possible.)

Review questions

Total: 45 marks

- 1 There is a difference between a firm's short-run and long-run profit-maximising position in perfect competition because of:
 - A Freedom of entry and exit
 - B Homogeneous products
 - C Many buyers and many sellers
 - D Perfect knowledge

(1 mark)

Questions 2–5 are based on Figure 13.8.

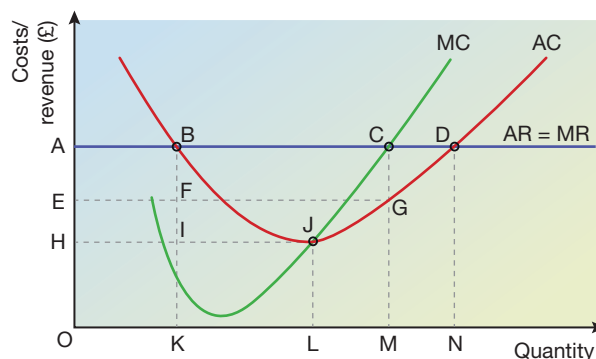


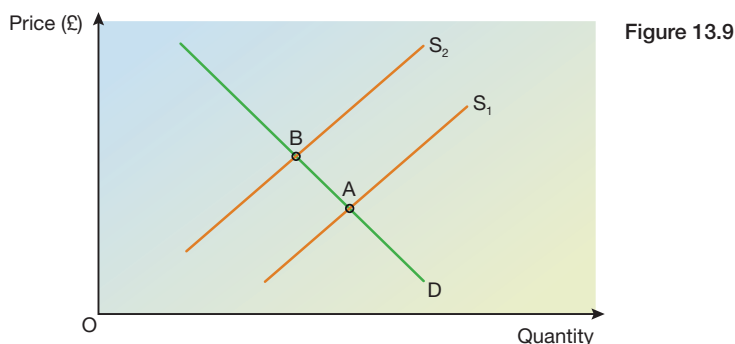
Figure 13.8 Short-run profit maximisation for a firm in perfect competition

- 2 In Figure 13.8, the profit-maximising output is:
 - A OK
 - B OL
 - C OM
 - D ON

(1 mark)

- 3** In Figure 13.8, average total costs are minimised at output:
 A OK B OL C OM D ON **(1 mark)**
- 4** In Figure 13.8, the firm's abnormal profit is shown by the area:
 A ACMO
 B EGMO
 C ACGE
 D BJD **(1 mark)**
- 5** In Figure 13.8, only normal profit is earned when output is:
 A OK and OL
 B OK and ON
 C OL and OM
 D OM and ON **(1 mark)**

Figure 13.9 applies to Question 6.



- 6** In Figure 13.9, the shift from Point A to Point B shows the change in:
 A Market equilibrium when firms enter the market
 B Market equilibrium when firms exit the market
 C The individual firm's equilibrium when firms enter the market
 D The individual firm's equilibrium when firms exit the market **(1 mark)**
- 7** What is meant by the term 'homogeneous products'? **(3 marks)**
- 8** Define the term 'price taker'. **(3 marks)**
- 9** What is the difference between a firm and an industry? **(4 marks)**
- 10** Explain *one* implication of the large number of producers in a perfect market. **(4 marks)**
- 11** Explain why firms will not continue to experience financial losses in the long run in a perfectly competitive market. **(4 marks)**
- 12** Explain why a firm's marginal revenue and average revenue are identical in a perfectly competitive market. **(5 marks)**
- 13** Draw a diagram to show a firm's long-run equilibrium position in a perfectly competitive market. **(6 marks)**
- 14** Analyse why perfect competition might result in an efficient allocation of resources, but only if it is assumed that there are no externalities and no inequality. **(10 marks)**

Monopolistic competition

Key concepts from Year 1

The concept of monopolistic competition was introduced in Chapter 17 of the Year 1 companion textbook. The key characteristics of monopolistically competitive markets were described alongside the features of other market structures.

This chapter commences by summarising the key assumptions and characteristics of monopolistically competitive markets because a clear grasp of these characteristics is essential in order to gain an understanding of the diagrammatic analysis. The diagrammatic analysis of the monopolistically competitive market will contrast the short-run equilibrium and the long-run equilibrium. The chapter concludes by examining the reasons why non-price competition is a feature of monopolistically competitive markets.

The main characteristics of monopolistically competitive markets

In the case of **monopolistic competition**, there are many firms, each supplying similar but differentiated goods. Consequently, individual firms have a minor impact on overall price levels in monopolistically competitive markets because the differentiation may mean that individuals value one firm's version of a good more highly than another firm's version of the same good. However, with no barriers to entry, product differentiation tends to be small because a popular variation of a product can be easily copied by competitors.

Certain factors are used to distinguish between different market structures. These are outlined below and applied to monopolistic competition.

- **Number of firms:** Monopolistic competition has *many* firms. Consequently, each firm has limited influence on market price (although some influence is possible because of differentiated products).
- **Nature of products – product differentiation:** Monopolistic competition has many small firms selling *similar*, but not identical, products. In order to gain a competitive advantage, individual firms try to differentiate their product from competitors. However, the high levels of competition tend to lead to low profit margins in the short run, and only normal profit in the long run. As a result, it is difficult for firms in monopolistic competition to achieve the funds and the market share required to support activities that might create barriers to entry. Product differentiation can be derived from different sources:
 - *product differences* – variables such as design, packaging, colour, size, ingredients/components, aesthetics and quality can be adjusted to give a certain 'uniqueness' to a product in a monopolistically competitive market;
 - *marketing/promotional differences* – advertisements or special offers can be

Key term

Monopolistic

competition is a market structure in which there are many buyers and many sellers. Firms supply similar, but differentiated, goods. There is freedom of entry into the market and freedom of exit out of the market. Knowledge amongst buyers tends to be widespread but is not perfect.

used to encourage consumer loyalty;

- *service differentiation* – the quality of service, such as politeness of staff, prompt delivery, flexibility in making changes or good quality after-sales service, can be a critical element in achieving product differentiation.
- **Ease of entry/barriers to entry:** It is difficult for existing firms to create barriers to entry because individual firms are free to create their own versions of goods in monopolistically competitive markets. The small scale of firms in this market structure tends to exclude factors that might act as barriers to entry, such as high capital costs, internal economies of scale, sunk costs and legal or marketing barriers. Ease of entry into this market structure is therefore high, and short-run supernormal profits will be competed away by new entrants.
- **Extent to which knowledge is perfect:** In monopolistic competition there are many firms and so knowledge will not be perfect. However, firms and consumers will have a good understanding of alternative goods and their prices, and knowledge is therefore widespread.
- **Firms' influence on price:** Since there is differentiation of products, each firm is able to set its own price. Technically, firms in this market are price makers rather than price takers. However, firms' influence on price levels is severely restricted because there are so many similar, competing products – demand tends to be very price elastic. Thus it is difficult to increase price.

Table 14.1 summarises the main characteristics of monopolistically competitive markets.

Characteristic	Feature of monopolistic competition
Number of firms	Many firms
Nature of products	Some differentiation
Ease of entry/ barriers to entry	Very easy to enter/ few barriers to entry
Information/ knowledge	Widespread knowledge, but not perfect
Influence of firms on price	Low/restricted

Table 14.1 *The main characteristics of monopolistically competitive markets*

Examples of monopolistic competition in the UK include bed and breakfast hotels and guest houses, female fashions in larger cities and towns, hairdressing, personal services such as gardening or painting and decorating, and restaurants and cafes.

The formal diagrammatic analysis of the monopolistically competitive model in the short and long run

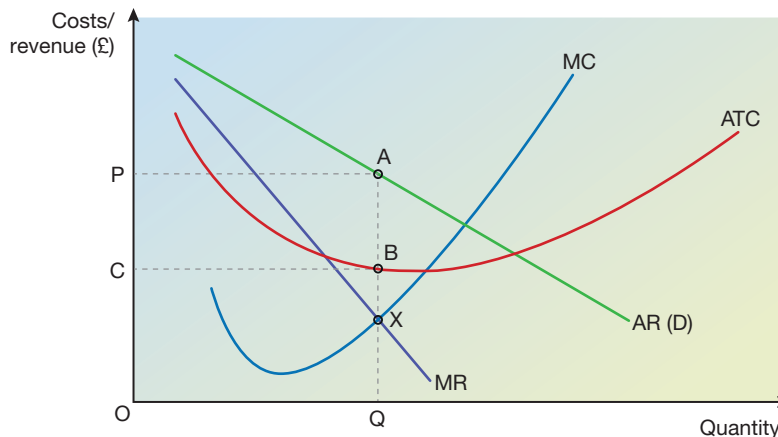
The way in which a monopolistically competitive market operates in the short run can be explained through Figure 14.1.

In monopolistic competition, each firm produces its own, slightly differentiated product. There is therefore no demand curve for the market as a whole. Instead, there are a series of individual demand curves for each differentiated product in the monopolistically competitive market.

In Figure 14.1, the line AR (D) shows the demand curve (and therefore the average revenue line) for a particular firm's product within the market. As a rule, demand

curves in monopolistically competitive markets are price elastic because there are many close substitutes for any given product. The marginal revenue line (MR) slopes downwards twice as steeply as the average revenue line (as previously explained in Chapter 8).

Figure 14.1 The equilibrium of the monopolistically competitive firm in the short run



There is no supply curve as such in a monopolistically competitive market. Each firm is free to choose its level of output. However, since a firm maximises profit when $MC = MR$, the chosen level of output will be determined by the combination of MC and MR. In Figure 14.1, marginal cost cuts marginal revenue from below at Point X. This shows the output that maximises profit for the firm in the short run and so it will produce OQ units. Since OQ units are supplied onto the market, the line QXBA in effect shows the firm's supply onto the market. With OQ units being supplied, demand is met at Point A and so the equilibrium price for the firm's product is P. Point B shows the average total cost of producing OQ units of output. Since B is lower than A, supernormal profit is made. Supernormal profit per unit of output is shown by $A - B$.

$$\text{Profit} = \text{Total revenue (TR)} - \text{Total costs (TC)}$$

$$\text{TR} = p \times q$$

Therefore $\text{TR} = OP \times OQ$. This is shown by area OPAQ.

$$\text{TC} = \text{ATC} \times q$$

Therefore $\text{TC} = OC \times OQ$. This is shown by area OCBQ.

Total profit is shown by area CPAB.

(Note that as normal profit is included in costs, this profit is 'supernormal profit' or 'abnormal profit'.)

Since there are no or few barriers to entry into a monopolistically competitive market, new firms will enter the market in order to gain a share of these supernormal profits.

Figure 14.2 shows the impact of these new entrants into the market on the individual firm.

Since there is no market supply in a monopolistically competitive market, the impact of new entrants is different from perfect competition. These entrants will bring many close substitutes to the firm's slightly differentiated product. Since these products are all very close substitutes, the demand for the firm's product will fall, as overall

demand for this type of product will be spread amongst a much greater number of firms. This will lead to a fall in demand.

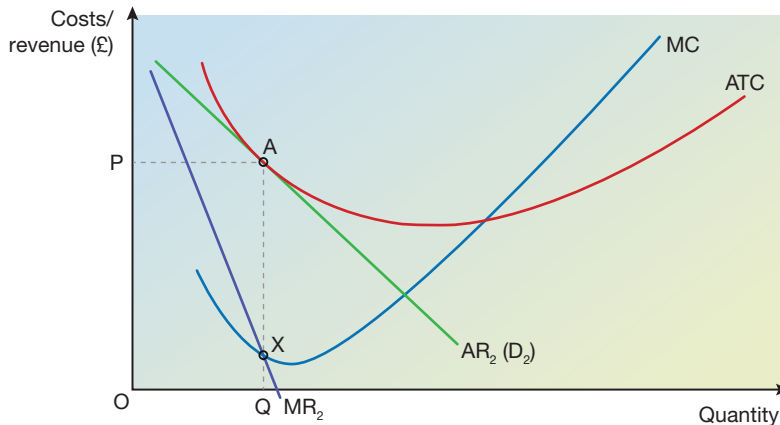


Figure 14.2 The equilibrium of the monopolistically competitive firm in the long run

In Figure 14.2, we see that the demand has fallen from AR (D) to AR₂ (D₂). This will be the culmination of a series of decreases in demand, each of which will reduce the supernormal profit that the firm is earning. As in perfect competition, the long-run equilibrium in a monopolistically competitive market occurs when only normal profit is made. When only normal profits are made, there is no further incentive for new firms to enter the market.

Normal profit is made when AR = ATC. In Figure 14.1, AR > ATC and so supernormal profit was earned. The decrease in demand for the firm's product shifts the AR line to the left and thus the MR line (which falls twice as steeply as the AR line) also shifts to the left. Eventually, AR will shift until there is just one point at which AR = ATC. This is shown by Point A in Figure 14.2. Point X (which is vertically below Point A) is where MC = MR, confirming that this is the profit-maximising output. Thus, the firm will produce OQ units. This is the long-run equilibrium because, with supernormal profits no longer in existence, new entrants into the market will cease and the firm will be making just normal profit. Both TR and TC can be measured by the rectangle OPAQ.

Is monopolistic competition efficient? Figure 14.1 shows the short-run equilibrium. The profit-maximising output is Q. This is slightly lower than the point at which ATC is minimised and so productive efficiency is *not* achieved, although ATC is quite close to its minimum point. Price is much higher than MC. The distance AX shows the difference between P and MC. This indicates that the market is a long way from achieving allocative efficiency in the short run. Figure 14.2 shows the long-run equilibrium. The profit-maximising output is Q. Point A shows that ATC is much higher than the lowest point of ATC. Productive efficiency is *not* achieved and, in fact, the long-run equilibrium is less productively efficient than the short-run equilibrium. Price is also much higher than MC. The distance AX shows the difference between P and MC. This indicates that the market is a long way from achieving allocative efficiency in the long run.

Monopolistically competitive markets will be subject to non-price competition

How do firms in monopolistically competitive markets compete? The existence of

many close substitutes means that demand is very price elastic. Since only normal profits are made in the long run, firms are unlikely to be in a position where they can cut prices. Competition must therefore be based on non-price factors.

Monopolistically competitive firms can achieve greater product differentiation by modifying their product in a way that attracts more consumers. However, as knowledge is widespread in this market structure and there are no barriers to entry, these modifications can be copied by competitors and thus any advantage is quickly lost.

Non-price competition usually takes the form of marketing activities. Successful marketing can help to differentiate a product from its competitors and this brings two main benefits: an increase in demand and a change in price elasticity of demand so that demand becomes more price inelastic. The latter benefit is particularly important in monopolistic competition because it can enable a business to resist the efforts of new entrants to capture their share of the market. Some of the key marketing activities in monopolistic competition are:

- increased expenditure on advertising;

REALWORLD ECONOMICS 14.1

Hairdressing

There are 29,415 hairdressing and beauty salons in the UK, with total annual sales revenue of £4 billion – approximately £140,000 per salon. Whereas some monopolistically competitive industries, such as restaurants, feature dominant market leaders, there is no dominant firm in the hairdressing and beauty market.

In addition to hairdressing and beauty salons, there are also many freelance (or mobile) hairdressers who conduct their business in customers' homes.



The large number of hairdressers in the UK reflects low barriers to entry

The high number of hairdressers indicates that there are relatively few barriers to entry into this market. Qualifications, such as NVQs, are important but not a necessity. Opening a salon can cost anywhere between £3000 and £35,000, depending on the suitability of the property. Many hairdressers enter the industry as mobile hairdressers because there is less financial risk if the business fails. Mobile hairdressers are also able to access new revenue streams, such as visits to day nurseries and care homes to access customers who might not come to a salon. However, many customers have brand loyalty to their existing hairdresser and so building up clientele can be difficult for a newcomer.

Two recent trends that have increased revenue for those hairdressers that anticipated these trends have been an increase in male grooming and the growing

popularity of hair extensions. Chloe Zumeris, a London-based hairdresser, believes that 'the location of the salon is the most important thing to consider'. Other factors cited as important by customers include:

- ancillary services, such as head massages, and the availability of coffee and magazines;
- a sense of being pampered;
- the appearance of the salon;
- the quality of the cut and styling;
- customer service and friendliness.

Sources: Article entitled 'How to start a hairdressing business' by Suzanne Berne in the *Guardian*, 21.5.15; article entitled 'Cutting it as a freelance hairdresser' by Alison Coleman in the *Guardian*, 16.8.13; and other sources.

Exercises **Total: 15 marks**

- 1 Explain two possible barriers to entry into hairdressing. **(6 marks)**
- 2 Analyse two ways in which hairdressers use non-price competition. **(9 marks)**

- product development to update and make products more attractive;
- better and more flexible distribution, such as quicker deliveries and making goods available 24–7;
- widening distribution channels, such as making the goods available from both shops and online;
- special offers such as three for the price of two.

Review questions

Total: 40 marks

- 1 Which *one* of the following features does not apply to monopolistic competition?
 - A Differentiated products
 - B Ease of entry
 - C Few firms
 - D Widespread, but not perfect, knowledge

(1 mark)
- 2 Which *one* of the following statements is true? Monopolistic competition:
 - A Achieves productive efficiency in the short run
 - B Achieves allocative efficiency in the long run
 - C Achieves both productive and allocative efficiency in the long run
 - D Does not achieve productive efficiency or allocative efficiency

(1 mark)
- 3 Which *one* of the following assumptions does *not* apply to *both* perfect competition and monopolistic competition?
 - A Ease of entry
 - B Homogeneous products
 - C Many buyers
 - D Many sellers

(1 mark)
- 4 Define the term 'monopolistic competition'.

(3 marks)
- 5 Explain *two* ways in which product differentiation can be achieved.

(6 marks)
- 6 Draw a fully labelled diagram to show a firm's short-run profit maximisation in monopolistic competition.

(7 marks)
- 7 Explain why demand tends to be price elastic in monopolistic competition.

(5 marks)
- 8 In perfect competition, long-run equilibrium follows an increase in market supply. In monopolistic competition, long-run equilibrium follows a decrease in the individual firm's demand. Explain the reasons for this difference.

(6 marks)
- 9 In the long-run equilibrium in monopolistic competition, *two* of the following equations are correct. Select these *two* equations.
 - (i) $ATC = AR$
 - (ii) $ATC = MC$
 - (iii) $ATC = MR$
 - (iv) $MC = MR$
 - (v) $MC = AR$

(2 marks)
- 10 Explain why there is non-price competition in monopolistically competitive markets.

(8 marks)

Oligopoly

Key concepts from Year 1

The concept of oligopoly was briefly introduced in Chapter 17 of the Year 1 companion textbook, and concentration ratios in Chapter 20. Since this chapter provides a detailed study of oligopoly, there is no need to revisit these chapters.

This chapter introduces the concept of oligopoly and describes the main characteristics of oligopolistic markets and their differences in relation to the number of firms, product differentiation and ease of entry. The chapter then examines the definition of oligopoly in terms of both market structure and market conduct (behaviour). Concentration ratios and their calculation as a means of measuring monopoly power are considered. The chapter considers the difference between collusive and non-collusive oligopoly, the difference between cooperation and collusion in oligopoly, and the kinked demand curve as one model to explain oligopolistic behaviour. An examination is undertaken of the reasons for non-price competition, the operation of cartels and features of oligopoly markets, such as price leadership, price agreements, price wars and barriers to entry. A study of the factors that influence price, output, investment and spending on research and advertising is then undertaken. The chapter concludes by examining the significance of interdependence and uncertainty in oligopoly and by summarising the advantages and disadvantages of oligopoly as a market structure.

The main characteristics of oligopolistic markets

Certain factors are used to distinguish between different market structures. These are outlined below and applied to oligopoly.

- **Number of firms:** In oligopoly there are few firms. This can vary from three firms, upwards. A market with two firms is known as a duopoly. Duopoly markets operate in a similar way to oligopoly, although collusion is more likely to bring benefits in a duopoly market.
- **Product differentiation:** Product differentiation means that consumers may have brand loyalty for their particular product. In extreme cases, product differentiation may be so significant that consumers will insist on a particular brand as they perceive it to possess qualities that competitors' products lack. For example, the market for cola-based carbonated drinks is an oligopoly that features two main brands – Coca-Cola and Pepsi Cola. These two firms have managed to differentiate their products so successfully that many consumers do not consider alternatives to be suitable. They are thus able to sell greater quantities *and* charge higher prices than other cola-based carbonated drinks. The strong brand loyalty also makes it very difficult for new competitors to enter this market. In many of these industries, there is a clear distinction between products and/or high brand loyalty.
- **Ease of entry/barriers to entry:** Barriers to entry restrict the level of competition in

Key term

Oligopoly is a market structure in which the supply of goods is dominated by a few firms. There are barriers to entry into the industry and may be barriers to restrict exit from the market by firms. Although some knowledge may be widespread, firms in oligopoly markets are likely to possess some knowledge that is not available to others.

many industries. These barriers can take many forms. Some of the main examples of barriers to entry are as follows:

- *capital costs* – in industries such as water supply, the initial cost of capital equipment needed to supply the good or service is very high. This means that it is not easy for new firms to enter the market and provide competition;
- *internal economies of scale* – in large industries, existing firms may operate on a very large scale and enjoy considerable internal economies of scale. New entrants are unlikely to operate on a scale that reaches the minimum efficient scale and therefore are not able to compete on price;
- *sunk costs* – these are costs that are not recoverable if a firm ceases to exist. Firms such as car manufacturers will have spent a lot of money on assets that are worthless to other industries. These high ‘sunk costs’ discourage new firms from entering the market, but discourage existing firms from ‘exiting’ the market;
- *legal barriers to entry* – patents, copyright and trademarks give firms a guaranteed monopoly of supply for a good or design that they have invented for a period of years. In some of the privatised industries in the UK, such as rail transport, a firm is given exclusive rights to operate a particular route for a number of years;
- *marketing barriers* – successful product differentiation often requires considerable marketing to create a brand that customers perceive to be differentiated or even unique. Firms such as Apple have created such a strong brand reputation that new firms find it very difficult to enter the market.

As a market structure, oligopoly markets can have considerable barriers to entry, mainly based on the economies of scale and marketing barriers created by existing competitors. However, these barriers vary considerably between different oligopoly markets.

- **Extent to which knowledge is perfect:** In oligopoly, some knowledge may be widespread, but firms in oligopoly markets are likely to possess some knowledge and information that is not known by other firms or customers in the market. Asymmetric information, with firms possessing far more knowledge than consumers, is common in oligopoly markets.
- **Firms’/suppliers’ influence on price:** Individual oligopolists can have considerable influence on price, especially if there is clear differentiation between products in the market. This influence will be greater if there are few competitors and/or the product has a large market share.

Table 15.1 summarises the main characteristics of oligopoly.

Characteristic	Feature of oligopoly
Number of firms	Few firms
Nature of products	Relatively high levels of differentiation possible
Ease of entry/barriers to entry	Many barriers to entry
Information/knowledge	Restricted knowledge and asymmetric information
Influence of firms on price	Quite high

Table 15.1 *The main characteristics of oligopoly*

Examples of oligopoly in the UK include brewing, cars, fast food restaurants, mobile phone networks, soft drinks, sports clothing and supermarkets.

Variation in oligopolistic markets

Oligopolistic markets are commonplace in the UK. In part this is because the definition of oligopoly covers a wide spread of markets. The differences can be quite significant. Some examples are noted below.

- **Number of firms:** Over 50 different car manufacturers sell cars in the UK. With most manufacturers offering a range of models, the number of different cars can be measured in the hundreds. Furthermore, individual models can have significant differentiation through differences in fuel, engine size, fixtures and design. In contrast, the UK market for detergents is dominated by two firms (Procter & Gamble and Unilever), which produce 80% of the detergents sold in the UK.
- **Degree of product differentiation:** The pharmaceutical industry features high levels of product differentiation because it is based on patented drugs that guarantee the supplier a monopoly of supply for about 20 years. However, for non-patented drugs (generic drugs) there is little product differentiation because the active ingredients tend to be the same.
- **Ease of entry:** The minimum efficient scale for car production is estimated to be 2 million cars per annum. This makes it very difficult for new entrants to get into the global market. The trend in recent years has been for firms to exit the industry or integrate with competitors in order to increase their scale. In contrast, it is very easy for a firm to set up a new coffee shop because capital costs are low and the convenience of the location is a crucial factor in gaining customers. Even though three main firms (Costa, Starbucks and Caffè Nero) dominate the coffee shop market, with over 50% of the market for branded coffee shops, this has not prevented small independent coffee shops from entering the market.

Oligopoly can be defined in terms of market structure or market conduct (behaviour)

As indicated in the previous section, the concept of oligopoly covers a broad range of markets. Oligopoly can be defined in terms of market structure or in terms of market conduct (behaviour), although the two are interconnected.

Key term

Market conduct

describes the behaviour of a firm in terms of its actions in its market and its responses to actions taken by competitors.

Oligopoly in terms of market structure

- Supply in the market should be concentrated in the hands of relatively few firms. There may be many small firms, but oligopoly requires some firms to be large enough to influence the market as a whole.
- Barriers to entry must exist. A key feature of oligopoly is the existence of long-run supernormal profit because new entrants cannot readily enter the market to compete for any supernormal profits.
- Interdependence between firms must exist. In oligopoly, the actions of one firm will impact upon other firms in the industry. For example, if an oligopolist reduces price or introduces a successful new marketing campaign, then its competitors are likely to experience a decrease in demand for their products. This contrasts with perfect competition in which firms' actions have no impact on other firms and so there is no interdependence.

Oligopoly in terms of market conduct (behaviour)

The behaviour of firms can be used to define oligopoly. Some examples of oligopolistic behaviour are identified below.

- **Product differentiation:** Various ways are found to create a more unique product, such as a modified design or new functions.
- **Non-price competition:** This concept was introduced in Chapter 14 and occurs when firms use techniques such as marketing strategies in order to compete. In oligopoly, price competition may take the form of a price war, with different oligopolists forcing the price down in order to increase their market share. Ultimately, this can result in much lower profit margins for all of the oligopolists. As a consequence, oligopolists will often try to keep prices quite high but gain market share by persuading customers that their product has unique qualities (i.e. trying to create product differentiation).
- **Price leaders:** A firm with a large market share – often the market leader – has a significant influence on the market price. Smaller competitors will tend to follow the price changes of the market leader.
- **Collusion:** This occurs when oligopolists agree to certain behaviour, such as increasing prices, so that they can all increase their profits. The interdependence between firms encourages this type of collusion, although it is not legal if the government deems it to be ‘anti-competitive’.
- **Agreements to create barriers to entry:** If there is a danger of a new firm entering the industry, oligopolists may get together in order to discourage this new entrant. This often takes the form of much lower prices, so that the new firm cannot make a profit. Oligopolists that anticipate a new entrant can build up retained profits from earlier years, so that they can survive a period of losses. In the long run, the prevention of competition will lead to higher profits.
- **Price rigidity:** This is a feature of oligopoly markets. In a competitive oligopoly, if the firm increases its price, competitors will not react so that they can gain market share. If a firm decreases its price, competitors will react to prevent the loss of market share. As a consequence, there is little to be gained by changing price, and so price rigidity exists.

Concentration ratios and how to calculate a concentration ratio

Concentration ratios were introduced on pp. 194–195 (Chapter 20) of the Year 1 companion textbook. That extract is reproduced here alongside an updated pie chart and commentary to show the changes in the market shares of UK supermarkets.

Concentration ratios provide a good indication of the potential monopoly power of the largest firms in a concentrated market.

There is no set number of firms used to calculate the concentration ratio for a market or industry, although four-firm and five-firm concentration ratios are probably the most commonly used measures. Ideally, the number of firms used should be the one that gives the clearest picture of the degree of dominance by large firms.

In the PC ‘search engine’ market (Table 15.2), the degree of dominance is best shown by a one-firm concentration ratio because Google’s 88.4% market share means that it is quite close to being a pure monopoly. The three-firm concentration ratio adds together the market shares of Google, Bing and Yahoo in order to give

Table 15.2 UK market shares of PC ‘search engines’

Google	88.4%
Bing	6.6%
Yahoo	3.9%
Ask Jeeves	0.4%
AOL	0.3%
Others	0.4%

Source: *Forbes*, December 2014

$88.4\% + 6.6\% + 3.9\% = 98.9\%$. This ratio also gives a clear indication of monopoly power in this market. However, a four-firm concentration ratio would add the market share of the fourth biggest search engine – Ask Jeeves. The four-firm concentration ratio would be $98.9\% + 0.4\% = 99.3\%$. This could be deemed to be a misleading measure because it is highly unlikely that Ask Jeeves’s 0.4% market share gives it any degree of market power.

Calculating concentration ratios

In October 2014, the market shares of supermarkets in the UK were as shown in Figure 15.1. The pie chart shows that for UK supermarkets, the four-firm concentration ratio was $28.7 + 17.2 + 16.4 + 11.1 = 73.4\%$, and the five-firm concentration ratio was $28.7 + 17.2 + 16.4 + 11.1 + 6.2 = 79.6\%$.

In December 2015, the market shares of supermarkets in the UK were as shown in Figure 15.2. The pie chart shows that, for UK supermarkets, the four-firm concentration ratio was $28.0 + 16.7 + 16.2 + 11.0 = 71.9\%$, and the five-firm ratio was $28.0 + 16.7 + 16.2 + 11.0 + 6.2 = 78.1\%$. Both ratios have fallen by 1.5%, suggesting that the growth of Aldi and Lidl is making the market more competitive by reducing monopoly power in the market.

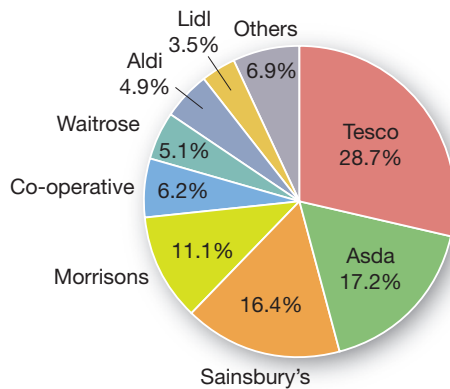


Figure 15.1 Market share of UK supermarkets, October 2014

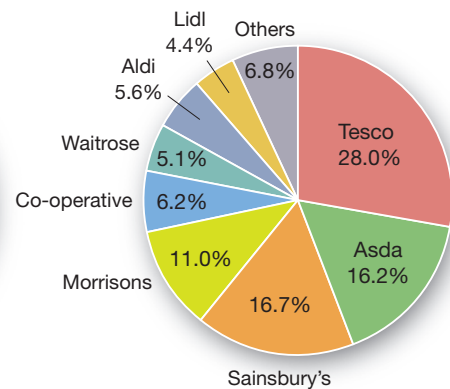


Figure 15.2 Market share of UK supermarkets, December 2015

Conclusion

Concentration ratios are a very good indicator of monopoly power because they provide a clear indication of the number of competitors in a market, the strength of those competitors and an implied indication of whether there are barriers to entry. However, they do not give any indication of the level of product differentiation and price inelasticity of demand, and therefore they do not necessarily provide an overall indication of market power.

The difference between collusive and non-collusive oligopoly

Collusive oligopoly means interdependent oligopolists agree not to compete. This is known as a **cartel**. Cartels are illegal in most countries. In the UK the Competition and Markets Authority (CMA) can investigate possible cartels and recommend that the government takes action to stop them operating, but it can be difficult for the

CMA to prove that a cartel exists. Often cartels make **price agreements** where all firms set a price that means supernormal profits are made. However, cartels may agree to limit competition by allowing each member firm to have a regional monopoly or by agreeing who bids for contracts. Collusive oligopoly depends on barriers to entry, as high profits can attract new competition.

A difficulty in controlling cartels is that their consequences often parallel that of a perfect market. For example, in a perfect market all firms charge the same price; in a collusive oligopoly all firms charge the same price. In competitive markets, firms tend to have an area in which they are more dominant; in collusive oligopoly they also tend to have an area in which they are more dominant. The key difference is that cartels can make long-run supernormal profits, but as the CMA investigates current activities it cannot foresee whether profit is being made until the companies publish their end of year accounts. However, CMA decisions must be made within a certain timescale, and so discovery of supernormal profits often occurs too late because it is after the judgement has been made.

Collusion can be formal or tacit:

- **Formal collusion:** This exists when the firms make an agreement that restricts competition. By agreeing not to compete, oligopolists act as a monopoly because they all provide the same offer (such as the same price). Without competition, the oligopolists can charge a higher price than normal and therefore make supernormal profits. However, this type of collusion can be restricted because of:
 - *difficulties in reaching an agreement* – each oligopolist must be convinced that it is gaining from the agreement;
 - *government or EU action* – the CMA or EU may investigate and break up any cartel, possibly having long-term negative consequences for the firms involved (such as decisions to split them into competing units). For example, an EU investigation into banks concluded that Lloyds needed to reduce its scale and so a new bank (TSB) took over many of its branches and now provides competition;
 - *a lack of trust* – cartels are only effective if each member keeps to the agreement. Breaking an agreement can be profitable for a single firm; it can also lead to fiercer competition and consequently lower profits than non-collusive oligopoly;
 - *the potential for new entrants* – if collusion leads to excessively high supernormal profits, this may make the industry more attractive to new entrants. Despite barriers to entry, the high profits available may make it worthwhile for new firms to enter the market and undermine any cartel;
 - *new technology* – this can change the nature of a product and enable a new firm to overcome barriers to entry. Collusion can discourage research and development and so there is a possibility that firms from outside the industry may develop suitable new technology while the efficiency of existing firms stagnates.

Key terms

Collusive oligopoly exists when firms in an oligopolistic market work together in order to reduce the level of competition.

Non-collusive oligopoly exists when firms in an oligopolistic market do not work together to reduce the level of competition.

A **cartel** is an agreement to restrict competition between firms.

- **Tacit collusion:** This exists when there is no formal agreement but where firms are reluctant to take actions that might lead to fiercer competition and lower profits. Where tacit collusion exists, firms will tend to follow the lead of any price change by the price leader, so that there is no price war.

If collusion works effectively, oligopolists can act as a monopoly and maximise their joint profits. However, such an action may bring about intervention by the Competition and Markets Authority (CMA).

Non-collusive oligopoly occurs when oligopolists compete. Their products are similar and so may have price elastic demand. However, **price wars** lead to low prices to the benefit of consumers rather than oligopolists, and so price competition is not the norm. **Non-price competition** is a feature of many oligopoly markets, such as competing through advertising and branding. Oligopoly also tends to encourage dynamic efficiency because non-price competition acts as a spur to improvements in quality and greater spending on research and development to invent new products or reduce costs.

Price leadership is another feature of oligopoly. The largest firm or firms often initiate price changes, with smaller firms then following suit.

The difference between cooperation and collusion

Collusion is intended to restrict competition so that oligopolists can earn supernormal profits. However, as noted in the previous section, this type of collusion can be undermined by new firms entering the industry with more technologically advanced products.

Many oligopolists attempt to overcome this potential problem by cooperating in certain activities, so that their individual and collective efficiency is improved. This **cooperation** can take different forms:

- *Sharing research and development* – for example, different pharmaceutical companies will often share research and development because of the high spending and high levels of risk involved. By cooperating, they can also share one another's expertise and gain synergy. By sharing expertise, firms can cooperate to discover new products that they would not have developed on their own.
- *Joint technology* – in the car industry, many parts are common to different models of car. By allowing a single supplier to concentrate on providing a particular part, such as light bulbs, the supplier can gain economies of scale and pass the cost savings on to the car manufacturers that use those particular parts.
- *Training of staff* – firms can share expertise so that employees are well trained and more efficient. An industrywide training scheme can share costs and act as a barrier to entry if it reduces the unit costs of the existing manufacturers.

The kinked demand curve model

Competitive oligopoly occurs when firms do not collude. Since firms in oligopoly markets tend to have some monopoly power, in a competitive oligopoly the actions of one oligopolist will affect the outcomes of other oligopolists. This occurs because oligopolists are interdependent and so they must take into consideration the actions of competitors alongside their own actions. This creates uncertainty. An oligopolist can never be sure of the reaction or actions of its competitors, although previous experience will provide some indication of likely outcomes.

Prices in competitive oligopoly are frequently ‘sticky’ and so firms often focus on non-price competition. This ‘stickiness’ arises because in oligopoly the demand curve can be kinked. This kink is based on two assumptions:

- If a firm increases price, competitors keep their prices the same. This means that there is price elastic demand for price increases. If competitors do not increase their prices, then the assumption of ‘ceteris paribus’ does not exist and the firm’s price will become much more uncompetitive.
- If a firm decreases price, competitors also decrease prices. Therefore additional demand will not be gained from competitors, although demand may extend. This means that demand is more price inelastic for price cuts because the percentage change in quantity demanded will be relatively low.

In Figure 15.3, the original price is P, with Q units sold. The kink in demand is at Point X. For price increases above P, demand is more elastic than for price decreases below. As MR slopes downwards at twice the gradient of AR, MR becomes discontinuous between Points A and E. Profit maximisation occurs where $MC = MR$. MC cuts MR from below at Point B, so the profit-maximising output for the oligopolist is OQ units.

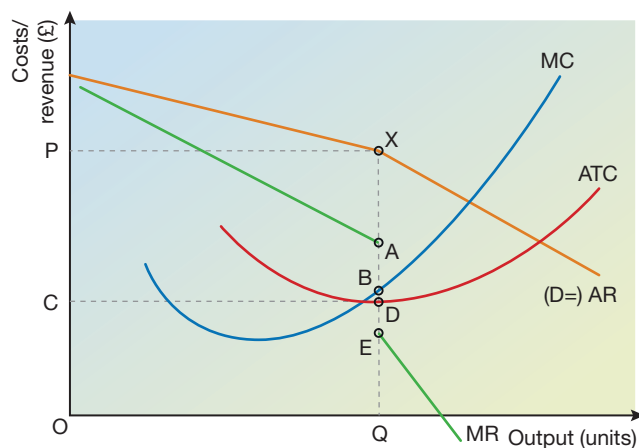


Figure 15.3 Oligopoly – profit maximisation with a kinked demand curve

Since $TR = OPXQ$ and $TC = OCDQ$, supernormal profit = $CPXD$. Given that there are barriers to entry in oligopoly, supernormal profit does not encourage new entrants and so it is both short-run and long-run profit.

If demand does not change, the price will stick at P. In other market structures, changes in costs will affect the market equilibrium because marginal costs will change, and so the point at which $MC = MR$ will change too. However, because of the discontinuity in the MR line in Figure 15.3, there is a high probability that the profit-maximising output will remain at OQ. Marginal cost (MC) could fall to a level as low as Point E and MC will still equal MR at output OQ. Similarly, if MC rises to Point A, MC will still cut MR at OQ, so profit-maximising output and price are unlikely to change.

The interdependence and uncertainty in oligopoly markets can also be shown using game theory, which models interdependent behaviour. Two competing firms may find that an action by one and a reaction by the other produce the best results for both. Alternatively, there may be equilibrium because a given action (or reaction) may always be most beneficial for one firm and so it always chooses this strategy.

However, the best outcome may be outside a firm's control, as it depends on actions of consumers and competitors.

Key note

Since oligopoly markets operate in different ways, the kinked demand curve is just one way of illustrating how an oligopoly market works. The kinked demand curve assumes that there is no collusion between oligopolists and that oligopolists will react to another oligopolist's price changes.

If there is no collusion and no reaction to price changes, the short-run profit maximisation will be similar to that shown in monopolistic competition in the previous chapter (although it is likely that price elasticity of demand will be more inelastic). However, as there are barriers to entry into oligopoly, the short-run diagram also applies to the long run because supernormal profits are not eroded by new entrants.

If there is collusion, oligopoly may act in the same way as a monopoly. This is represented diagrammatically in the next chapter.

The reasons for non-price competition, the operation of cartels, price leadership, price agreements, price wars and barriers to entry

Cartels operate in order to benefit their members. In most industries, the members are the competing firms, but arguably the most well-known cartel in the world is OPEC (the Organisation of the Petroleum Exporting Countries). This group agrees to take decisions to the mutual benefit of its member countries, usually by agreements to limit supply and thus increase price.

The operation of a cartel between oligopolists can take many different forms, such as:

- *price fixing* – by agreeing to charge the same price, oligopolists can remove the threat of price competition and ensure that each member of the cartel enjoys a healthy profit margin;
- *restricting output* – by reducing market supply, oligopolists can increase the market price. However, the success of this action will depend on how much cooperation there is between the oligopolists and the elasticity of demand for the product. For example, restricting output of essential goods, such as energy and water, will lead to a much larger increase in price because demand is price inelastic;

Key terms

Non-price competition occurs when firms agree to compete through activity, such as marketing, but not through changes in price. It usually takes the form of marketing activities such as advertising and promotions. (This topic was covered in the previous chapter on

monopolistic competition – see p. 89.) However, it is a more significant characteristic of oligopoly markets because non-price competition can create product differentiation. Along with barriers to entry, successful product differentiation can lead to long-term supernormal

profits, whereas in monopolistic competition product differentiation only leads to short-term supernormal profits.

Price leadership exists when the market leader sets a price that smaller competitors follow.

Price agreements are when oligopolists agree not to compete on the basis of price.

A **price war** is when oligopolists compete by cutting prices in order to increase their market share.

- *dividing up the market* – this usually takes the form of oligopolists agreeing to focus on different geographical areas, so that each oligopolist can act as a local monopoly;
- *agreeing not to compete when bidding for contracts* – in industries such as construction, firms must tender (bid) for specific building contracts. Through collusion, oligopolists can agree on the prices that they each bid for different contracts. This can be organised in a way that ensures that each of them gets a share of the contracts at higher prices than the prices likely to occur if there were competitive bidding.

In oligopoly markets, prices can be set in different ways. Three examples are described below:

- *price leadership* – this occurs when a large company, usually the market leader (the firm with the largest market share), sets a market price that smaller competitors will tend to follow. However, price leadership may be shared between different firms. For example, observations suggest that the three major oil suppliers in the UK – Esso, BP and Shell – tend to follow one another's lead when one of them changes price, and then smaller suppliers follow suit;
- *price agreements* – these are often tacit agreements that oligopolists will not compete on the basis of price; instead the focus will be on non-price competition. There may be price fixing, where a set price is agreed, or there may be an agreement to restrict output by an agreed amount, allowing the market to set the price. However, the price will be higher if output is restricted;
- *price wars* – these occur when oligopolists lower prices. They are often triggered by an attempt by one oligopolist to try to gain market share at the expense of the others. If the remaining oligopolists enter into a price war, this tends to benefit consumers through lower prices, rather than benefiting the oligopolists.

Barriers to entry are an essential feature of oligopolistic markets. Through activities such as high expenditure on marketing, a firm can persuade consumers of the unique qualities of its products and thus make it difficult for new firms to enter and compete. Legal barriers, such as patents, can prevent or restrict competition. Large-scale production can result in economies of scale, which mean that new entrants cannot possibly compete with existing oligopolists. All these barriers mean that even if supernormal profits are made, new firms cannot enter the market and compete away these profits. Consequently, oligopolists can retain supernormal profits in the long run.

The factors that influence prices, output, investment, expenditure on research, and advertising in oligopolistic industries

In this chapter we have studied the relationship between oligopolists and features such as price and output. This section summarises the main factors influencing these features in oligopolistic markets.

The factors that influence prices in oligopolistic industries

Prices depend on:

- whether the oligopolist is a price taker or price maker – in oligopoly, the price leaders are price makers and have significant influence on the price in the market as a whole. However, smaller oligopolists will normally be price takers and will follow the lead of the price leaders when setting price;

- the interdependence of firms in oligopoly markets – if oligopolists react to the actions of other oligopolists, then these reactions can affect the price set by an individual oligopolist. For example, an oligopolist would find it difficult to avoid a price war if other oligopolists pursue a price war in order to gain market share;
- the price elasticity of demand for the product – the more price inelastic the demand for the oligopolist's product, the higher the price. Firms that can successfully differentiate their products are more likely to have products with price inelastic demand;
- the output of the industry – if market output is high, price is likely to be lower;
- barriers to entry – high barriers of entry allow oligopolists to maintain higher prices because the barriers prevent new firms from entering the market and forcing the price down.

The factors that influence output in oligopolistic industries

The key influences are:

- the size of the market and the market share of the individual oligopolist;
- the scope for economies of scale – greater opportunities for economies of scale are likely to lead to oligopolists with high output; if diseconomies of scale occur at low levels of output, the industry is likely to contain more firms producing lower levels of output;
- the minimum efficient scale (MES) – the greater the MES, the more difficult it will be for new firms to achieve the scale necessary to compete with existing suppliers;
- the shape of the ATC curve – an L-shaped ATC will encourage firms to produce high levels of output; a U-shaped ATC will limit the optimum output.

The factors that influence investment in oligopolistic industries

Investment in oligopoly is used to try to improve the efficiency of production or generate new or improved products. The key influences are therefore:

- the importance/potential benefits of new technology in the industry;
- the need for new product development/the pace of change in consumer tastes;
- the scope to improve the existing productive efficiency of the firm through capital investment;
- the ability of the firm to access funds for investment – this will be closely linked to its profit level;
- interest rates – high interest rates will discourage investment.

The factors that influence expenditure on research in oligopolistic industries

Research and development is intended to assist product differentiation and innovation. Key influences are:

- the level of risk and uncertainty – high risk will discourage expenditure on research and development;
- the need for new product development/the pace of change in consumer tastes;
- chances of achieving protection for new inventions, such as patents, which will give a firm some monopoly power;

- the skills of the workforce with respect to developing new ideas.

The factors that influence advertising in oligopolistic industries

Advertising is also used to assist product differentiation. Key influences are:

- the degree of product differentiation in the market;
- the value placed on branding and uniqueness in the market;
- the level of profits available to fund advertising;
- the cost effectiveness of advertising in achieving product differentiation.

The significance of interdependence and uncertainty in oligopoly

The section on the kinked demand curve illustrates how oligopolists can be interdependent. If oligopolists are interdependent, they must take into consideration the actions of competitors alongside their own actions. For example, if a price decrease received no reaction, a firm can gain market share by cutting price. However, if it leads to other firms reacting by cutting their prices, it will probably lead to a price war. In the first case the oligopolist will probably be successful in increasing profit, but if there is a price war it will probably be unsuccessful. This creates uncertainty.

The advantages and disadvantages of oligopoly

It is evident that oligopoly is beneficial to the oligopolists, although it can make firms complacent and vulnerable to new firms that might use new technology to enter the market with a radically different product or more efficient production methods. However, when examining oligopoly markets, the CMA looks at the situation from the perspective of the consumer.

The advantages of oligopoly

- High levels of supernormal profits can lead to more expenditure on research and development. Consequently, more innovative new products will be launched, thus benefiting consumers in the market.
- High profits can also lead to the introduction of new technology and better production techniques. These will lower the costs of production and therefore could lower prices.
- The lack of firms in oligopoly markets tends to mean that oligopolists produce on a large scale. This allows them to reduce costs because they benefit from internal economies of scale.
- With relatively few firms, it is relatively easy for consumers to make comparisons between prices and other features of oligopolists, such as the differences between their products. This improves information and is likely to reduce the possibility of market failure.
- The analysis of interdependence and the kinked demand curve shows that price stability is a feature of oligopolies. This is beneficial for consumers as it helps them to plan their expenditure with some certainty.
- Oligopolists compete through product differentiation. This provides a much greater choice of products and can therefore cater for many different tastes, such as those of customers in niche markets.

The disadvantages of oligopoly

- Barriers to entry lead to a restricted number of firms and therefore restricted choice; if mergers occur, the number of firms in the market is reduced still further.
- Since both short-term and long-term profit are supernormal, price remains above marginal cost. Oligopolists do not therefore achieve allocative efficiency.
- Barriers to entry can prevent new ideas from being developed, particularly if these ideas originate from small firms that are unable to enter the market.

REALWORLD ECONOMICS 15.1

BT takeover of EE

In 2015 BT PLC indicated that it wished to take over EE Limited. Both of these oligopolists operate within the communications market, but BT argued that there was little overlap between the two firms and that the merger would not cause a substantial lessening of competition (SLC). SLC is the prime consideration of the Competition and Markets Authority (CMA) when deciding on whether to allow the merger or takeover to proceed. The main criterion used by the CMA, when judging a merger, is: 'Does a lessening of competition adversely affect consumers?'

BT wanted to take over EE so



BT's proposed takeover of EE would not cause a substantial lessening of competition, according to the CMA

Providers	Market share (%)
BT	37
Virgin Media	13
Others	50

Table 15.3 Fixed landlines – market shares (2015)

Providers	Market share (%)
BT	32
Sky	22
Virgin Media	20
Talk Talk	14
EE	4
Others	8

Table 15.4 Fixed broadband providers – market shares (2015)

Providers	Market share (%)
EE	29
O2	27
Vodafone	21
3	12
Others	11

Table 15.5 Mobile phone networks – market shares (2015)

that it could improve its presence in mobile communications. It believed that the merger would also bring about cost savings that would benefit customers.

Objections to the merger were mainly based on the market

domination that BT and EE enjoy in their respective markets (see Tables 15.3, 15.4 and 15.5) and that it would reduce the number of firms competing in the fixed broadband market.

The CMA's provisional finding was that the two firms' functions had very little overlap and efficiencies could be gained by a merger. Any loss of competition was deemed to have no substantial lessening of competitiveness. However, the CMA did conclude that its judgement could be undermined by the complexity and dynamic and innovative nature of this industry as well as the significant changes in consumer demands for data. However, it believed that existing competitors were sophisticated and able to ensure high levels of competition.

Sources: CMA report – provisional findings on proposed merger, October 2015; data on market shares from Ofcom

Exercises Total: 10 marks

- 1 Calculate the four-firm concentration ratio for the mobile phone market. (2 marks)
- 2 Based on the evidence provided, do you agree with the CMA's decision to allow this merger? Justify your view. (8 marks)

- Oligopoly markets can become complacent because of an absence of new competition and they may therefore lack dynamic efficiency. This complacency can lead to higher costs because firms can make profit even if they are productively inefficient.
- Oligopolists promote inequality because they can take advantage of both consumers and smaller competitors within the market.
- Oligopoly encourages high levels of expenditure on non-price competition, such as marketing. This adds to the cost and probably the price of the finished products.

Overall, there are benefits to oligopoly and it may be dynamically efficient in terms of developing modified products. However, it is inefficient in terms of both productive and allocative efficiency.

Review questions

Total: 60 marks

- 1 Oligopoly has all of the following features *except*:
 - A Barriers to entry
 - B Few firms
 - C Perfect knowledge
 - D Product differentiation

(1 mark)

- 2 A market with only two firms is known as a:
 - A Biopoly
 - B Duopoly
 - C Totopoly
 - D Twinopoly

(1 mark)

- 3 Which *one* of these factors defines oligopoly in terms of market structure rather than market conduct?
 - A Collusion
 - B Non-price competition
 - C Number of firms
 - D Price rigidity

(1 mark)

Figure 15.4 applies to Questions 4, 5 and 6.

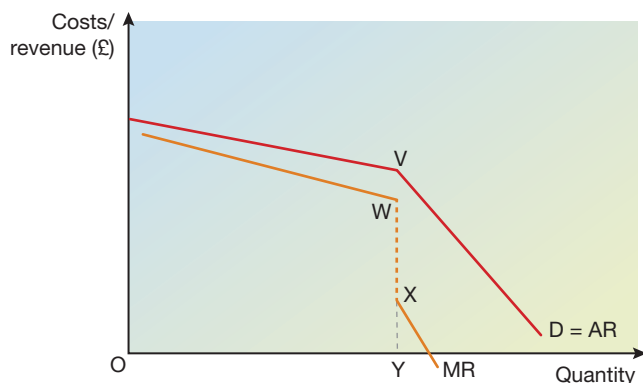


Figure 15.4 Oligopoly

- 4** Price rigidity occurs if marginal cost at output OY lies between:
A V and W
B V and X
C W and X
D W and Y (1 mark)
- 5** Total revenue is shown by:
A VX
B VY
C VX × OY
D VY × OY (1 mark)
- 6** The firm should exit the industry if its average total costs (ATC) at OY units of output is greater than:
A VY
B WY
C WX
D XY (1 mark)
- 7** Concentration ratios are used to show:
A Allocative efficiency
B Levels of supernormal profit
C Monopoly power
D Productive efficiency (1 mark)
- 8** Define the term 'sunk costs'. (3 marks)
- 9** What is meant by the term 'market conduct'? (3 marks)
- 10** Define the term 'concentration ratio'. (3 marks)
- 11** What is meant by the term 'cartel'? (3 marks)
- 12** Define the term 'price leadership'. (3 marks)
- 13** Define the term 'price war'. (3 marks)
- 14** Explain the difference between 'collusive oligopoly' and 'non-collusive oligopoly'. (4 marks)
- 15** Explain why research and development expenditure is likely to be higher in oligopoly than perfect competition. (5 marks)
- 16** Explain two barriers to entry. (6 marks)
- 17** Explain why the demand curve for an oligopolist may be 'kinked'. (6 marks)
- 18** Explain two actions a cartel might take to assist oligopolists. (6 marks)
- 19** Explain two ways in which oligopoly can benefit consumers. (8 marks)

Monopoly & monopoly power

Key concepts from Year 1

Monopoly and monopoly power were introduced in Chapter 20 of the Year 1 companion textbook. You should read that chapter in conjunction with this one in order to have a full understanding of monopoly for the A-level. The Year 1 book discussed the concepts of monopoly and monopoly power and explained how monopoly power is influenced by factors such as barriers to entry, the number of competitors, advertising and the degree of product differentiation. Potential benefits of monopoly were explained along with its negative impact on the allocation of resources.

This chapter will focus mainly on the formal diagrammatic analysis of the monopoly model, allowing it to be compared to the models relating to other market structures. A summary of some further advantages and disadvantages of monopoly will be provided to support the points developed in Chapter 20 of the Year 1 book.

Author tip

This chapter focuses on pure monopoly. It is important to recognise that firms operating in monopolistically competitive and oligopolistic markets are price makers and thus have varying degrees of monopoly power. The extent to which these firms have monopoly power can be seen through the models of monopolistic competition and oligopoly.

In monopolistic competition, firms can make supernormal profits in the short run. These profits will depend largely on the degree of competition, which in turn is influenced by the degree of product differentiation. Since there are many competitors with similar products, monopoly power tends to exist but at a lower level than in oligopoly. In the long run, there is no monopoly power because the model of monopolistic competition suggests that competitors will enter the market until supernormal profits are eroded.

In oligopolistic markets, there is usually more significant product differentiation because there are only a few firms. In some industries, legal protection, such as patents, can create significant product differentiation, even if there are many competitors, and so monopoly power can be quite high. Where there are only a few firms with distinctive products, consumers may feel that certain brands are unique – this situation also leads to significant monopoly power. In oligopoly, there are barriers to entry, and so monopoly power exists in the long run.

Key terms

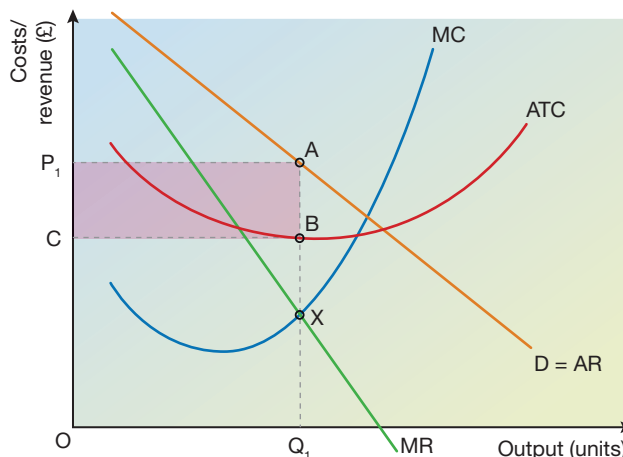
Pure monopoly exists when there is a single supplier in a market.

Monopoly power arises when firms exert considerable influence in a market because of their relatively large size.

The formal diagrammatic analysis of the monopoly model

In monopoly there is a single supplier producing a differentiated product. The demand for the product is therefore the AR line and the MR line slopes downwards with twice the gradient of AR.

Figure 16.1 Monopoly and profit maximisation



In Figure 16.1, MC cuts MR at X and so Q_1 is profit-maximising quantity. The monopolist sells Q_1 units at price P_1 , as A is where demand intersects output Q_1 . In monopoly there is no supply curve as such. The interaction of MC and MR informs the monopolist of the output at which profit will be maximised. The ATC of producing Q_1 units is shown by B, so ATC is OC per unit.

$$TR = OQ_1 \times OP_1 \text{ (or rectangle } OP_1AQ_1\text{)}$$

$$TC = OQ_1 \times OC \text{ (or rectangle } OCBQ_1\text{)}$$

$$\text{Supernormal profit} = OQ_1 \times CP_1 \text{ (or rectangle } CP_1AB\text{)}$$

This shows the *short-run equilibrium*. Barriers to entry mean that supernormal profits do not attract new firms into the market. Thus this diagram also shows the *long-run equilibrium*.

The advantages and disadvantages of monopoly

As with oligopoly, the advantages and disadvantages are considered from the perspective of consumers.

The advantages of monopoly

(Please note that the first three advantages below are explained in detail on pp. 196–198 of the Year 1 companion textbook.)

- Internal economies of scale are possible because a single firm supplies the whole market. This leads to lower ATC and possibly lower prices for consumers.
- Greater innovation is possible because high levels of supernormal profits, in the long run, provide a reliable source of finance for expenditure on research and development. Consequently, more innovative new products will be launched, benefiting consumers in the market.
- High profits can also lead to the introduction of new technology and better production techniques, thus lowering ATC.
- In certain situations, monopoly can widen choice for consumers. Oligopolists tend to focus their attention on the largest market segments, as they offer the highest potential sales revenue. This can mean that minority tastes are ignored. A monopolist only needs to provide one product for the largest market segment

and so can devote resources to minority tastes, if there is the opportunity to make some profit.

The disadvantages of monopoly

- The main disadvantages of monopoly are that it is likely to lead to inefficiency in terms of productive efficiency, allocative efficiency and dynamic efficiency.

The causes of these inefficiencies will be explained fully in Chapter 20.

REALWORLD ECONOMICS 16.1

Royal Mail – monopolist and oligopolist

Postal deliveries consist of two distinct markets – letters and parcels. For most of its 350-year history, the Royal Mail enjoyed a monopoly of both of these markets, but liberalisation of the parcels market in the 1990s led to the creation of oligopoly, with about 20 national firms now competing with Royal Mail. In recent years, alternative communication systems such as email and social media have led to a decline in the letters market, falling by an average of 6.3% per annum between 2008 and 2013. This is in sharp contrast to the parcels market, which grew



Royal Mail lost its legal monopoly of letter deliveries in 2006

by 3.7% per annum from 2008 to 2013, largely through greater use of online shopping.

Until 2005, Royal Mail retained a legal monopoly of the letters market. Governments have always supported the principle that every household in the UK should be entitled to the same delivery service at the same price for letters, however remote their location. This promise is known as the Universal Service Obligation (USO) and was an obligation placed on Royal Mail in return for a guaranteed legal monopoly.

In 2006, Royal Mail lost its legal monopoly of letter deliveries. It argued against this decision, saying that new entrants would target only densely populated urban areas where costs of delivery are much lower. This would leave Royal Mail with the unprofitable, costly rural delivery services that it was obliged to continue under the USO. One advantage that Royal Mail did receive was that its letter service was exempt from VAT, allowing it to save 20% in comparison to competitors. Although private firms were eager to enter the parcels market in the 1990s, the liberalisation of the letters market

in 2006 did not initially attract new entrants.

In 2012, Dutch firm Whistl (TNT) entered the letters market, focusing on deliveries in London, Liverpool and Manchester. By 2015 it was estimated to be delivering 25% of UK mail, despite focusing on this narrow geographical area. However, in 2015 Whistl announced that it had decided to exit the industry, with 2000 delivery staff being made redundant.

Royal Mail is the market leader in parcel deliveries with a market share of over 30%. However, its record on efficiency is not good. In 2007 it estimated that it was 40% less efficient than its competitors in the parcels market. It set itself a target of 2–3% efficiency gains per annum, but has only achieved 1.7% efficiency gains in the last two years. Royal Mail's high market share is largely a result of inertia amongst customers.

Sources: Ofcom, Citizens Advice and other sources

Discussion point

Why is it that the parcel post market has become oligopolistic, whereas in the letters market Royal Mail has a monopoly despite losing its legal monopoly status?

Additional disadvantages are:

- Barriers to entry lead to restricted choice – only one product may be offered by the monopolist because this will reduce costs.
- In turn this can lead to complacency and a lack of new products in the market.
- Monopolists often provide essential services. A profit-maximising monopolist is likely to exploit consumers with the lowest incomes who usually spend a larger percentage of their incomes on essential services.

Overall, there are benefits to monopoly, but it is usually inefficient and thus leads to a misallocation of resources.

Review questions

Total: 25 marks

- 1 Which *one* of the following statements is true? In normal circumstances:
 - A Monopolists earn normal profit in both the short run and the long run
 - B Monopolists earn normal profit in the short run and supernormal profits in the long run
 - C Monopolists earn supernormal profit in the short run and normal profit in the long run
 - D Monopolists earn supernormal profit in both the short run and the long run

(1 mark)
- 2 Define the term 'pure monopoly'.

(3 marks)
- 3 What is meant by the term 'monopoly power'?

(3 marks)
- 4 Draw a diagram to show a profit-maximising monopolist's short-run equilibrium output and price and its profit.

(6 marks)
- 5 Explain why monopoly might lead to:
 - (a) less choice for consumers; *(3 marks)*
 - (b) more choice for consumers. *(3 marks)*
- 6 Other than choice, explain *one* possible advantage of monopoly for consumers and *one* possible disadvantage of monopoly for consumers. *(6 marks)*

Price discrimination

Key concepts from Year 1

Price discrimination is practised mainly by monopolists or firms with monopoly power. An understanding of Chapter 16 of this book is the only foundation required.

This chapter will introduce the concept of price discrimination and outline the conditions necessary for price discrimination to take place. A diagrammatic analysis of price discrimination is provided. The chapter concludes by considering the advantages and disadvantages of price discrimination. (*The impact of price discrimination on consumer surplus and producer surplus will be covered in Chapter 21.*)

Introduction to price discrimination

Price discrimination normally means dividing a market into two or more different groups of customers and setting different prices for each group. This is the most common form of price discrimination and is known as third-degree price discrimination. Examples include rail journeys, telephone charges and cheaper holidays for children or retired consumers.

Second-degree price discrimination occurs when different prices are charged according to the volume purchased. This type of price discrimination may be based on costs, as low-volume tends to mean higher fixed costs per unit for the supplier. Examples include energy supply, telephones and wholesalers selling to retailers.

First-degree price discrimination (also known as perfect price discrimination) occurs when each customer is charged a different price. This is very rare, but might be possible in the case of a service where the supplier knows the demand and ability to pay of each customer. Examples might include private services, such as specialist medical treatment or specialist beauty care.

This chapter will focus on third-degree price discrimination because this is the type that occurs in most markets in which price discrimination is practised.

Price discrimination tends to be based on three factors that can allow a firm to distinguish between different groups of consumers:

- *Time of use* – for example, rail companies charge higher prices during peak times because commuters usually travel at those times. Commuters have more price inelastic demand than casual users because there is no easy alternative to train travel for them to get to their employment.
- *Income* – many services offer lower prices to students because their lack of income means that they are more conscious of price than other groups.
- *Place* – petrol stations in rural areas and on motorways tend to charge higher prices because there is no local competition; in large urban areas there are more petrol stations and so price tends to be lower.

Key term

Price discrimination

occurs when a firm charges different prices to different customers, based on their willingness to pay rather than on differences in the cost of supplying to them.

The conditions necessary for price discrimination

Price discrimination requires certain conditions:

- *The firm must have a certain degree of market power.* In a perfect market there will only be one market price. In monopolistic competition a high price charged to one group of customers can be easily undermined by a competitor. Price discrimination is practised in oligopoly markets, particularly collusive oligopoly, but is most likely to exist in monopoly markets.
- *The price discriminator must be able to identify and separate different groups of customers.* Time of use and place are thus feasible as ways of recognising distinct groups of consumers. However, income is more problematic as it is not usually possible to know a consumer's income.
- *It must be impossible for the product to be sold to another consumer, through what is known as secondary markets.* For example, in 1987 there was a huge surplus of butter in the European Union. The EU purchased this surplus to protect farmers' incomes and sold most of the surplus to non-EU countries, such as Russia. However, people receiving benefits, such as pensioners, were allowed to purchase butter at a heavily discounted price. This led to incidents where people entitled to discounts bought large quantities of butter and then resold the butter in the supermarket car park. In effect, effective price discrimination is limited to services because it can be very difficult to prevent the resale of goods.
- *Different groups of customers must have different price elasticities of demand.* Price discrimination works because those who value the product most highly will tend to have price inelastic demand, whereas those groups that are more price conscious will have price elastic demand. This means that the firm can increase the price to those who have price inelastic demand and only lose a few customers. At the same time it can decrease the price for those who have price elastic demand and gain a significant increase in the number of customers. Overall, this will allow it to increase total revenue by charging a much higher than equilibrium price to people with price inelastic demand and a slightly lower than equilibrium price to people with price elastic demand. If all groups of customers have the same price elasticity of demand, total revenue cannot be increased in this way.

Figure 17.1 shows price discrimination. In the case of railway companies, Figure 17.1(a) shows the situation with commuters, while Figure 17.1(b) shows casual users. The latter have more price elastic demand because rail transport is not essential to them. At off-peak times their journey is less important than getting to work, and, in cities such as London, using other forms of transport is more realistic outside the rush hour.

The MC line is shown as a horizontal line in both diagrams, as this represents the marginal cost of the final unit of rail transport. The market equilibrium price without price discrimination would be slightly above P_E . By cutting price for off-peak customers slightly to P_E , there is a significant increase in off-peak customers because their demand is price elastic. For commuters, the price P_1 is well above the original equilibrium. However, because their demand is price inelastic, this high price only leads to a slight decline in quantity. TR is significantly increased, but as the total output is the same as it would have been without price discrimination, it is assumed that there is no change in costs. Note that the MR for the final commuter is the same as the MR for the final off-peak customer. The assumption that total costs

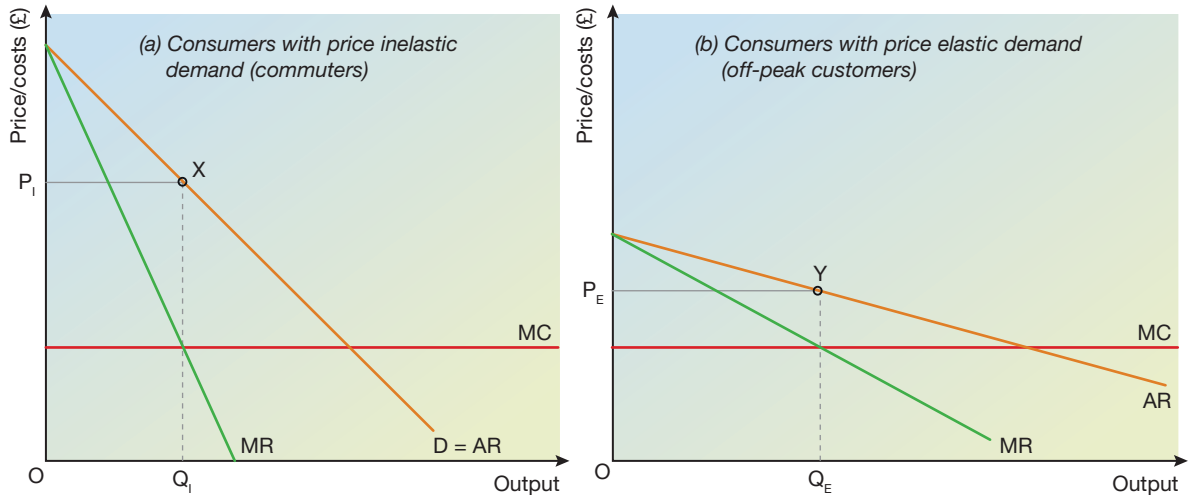


Figure 17.1 Price discrimination

remain the same can be challenged. However, for industries such as rail travel and telecommunications, it is more expensive to supply peak-time consumers because this supply requires most fixed capital, such as railway carriages. Thus it is likely that reducing peak travel and increasing off-peak travel leads to a reduction in total costs because less capital is needed if demand is spread more evenly throughout the day.

The advantages and disadvantages of price discrimination

Price discrimination enables firms to increase their profit and so it is beneficial for firms that practise price discrimination. In certain markets, low demand may mean that it is unprofitable for a monopolist to provide a service. However, the increase in revenue brought about by price discrimination may enable the firm to make an overall profit and therefore provide the service to all groups of customers.

The paragraphs that follow examine the advantages and disadvantages of price discrimination from the perspective of consumers.

The advantages of price discrimination to consumers

- Where price discrimination is the only way in which a firm can stay in business, it will benefit consumers because services will be made available to them that would not be provided if the firm were unable to practise price discrimination.
- Price discrimination benefits those consumers who get lower prices than the usual equilibrium (see Figure 17.1(b)).

The disadvantages of price discrimination to consumers

- Price discrimination means a higher price is charged to consumers who value the good more highly (see Figure 17.1(a)).
- The consumers will be those that have price inelastic demand. For many consumers, this means that they perceive the good to be a necessity. Price discrimination therefore hits those people who are purchasing what they believe to be necessities.
- Price discrimination can take away all or some of an individual's consumer surplus. This is particularly damaging to consumers when first-degree price discrimination occurs. (*This point will be explained in Chapter 21.*)

Author tip

Oligopoly and price discrimination are good topics for real-life analysis and evaluation because they test application, analysis and evaluation. Ensure you are familiar with real-life examples.

REALWORLD ECONOMICS 17.1

Will 'big data' bring more price discrimination?

The term 'big data' describes large pools of data that can be captured, communicated, aggregated, stored and analysed. The increasing use of technology in the business world enables firms to gain much more detailed data on their consumers.

Historically, first-degree price discrimination has not been considered to be feasible because firms do not have the information on their customers to discover their willingness to pay. However, loyalty cards and detailed analysis of customers' visits to websites are now providing firms with much more detailed and relevant data. In the home, Smart TVs are watching people who watch television and feeding data to the manufacturers.

A paper on *Netflix*, by Benjamin Shiller, investigated the extent to which information

on their customers could be used to practise first-degree price discrimination. In his paper, Shiller notes that there is a 16% probability of a customer using *Netflix* in the USA. However, demographic information such as age and income can be used to predict these probabilities more accurately, with certain demographic groups only having a 6% chance of using *Netflix*, while others have a 30% chance. This type of data lends itself to third-degree price discrimination, as the latter group is likely to value *Netflix*'s services more highly, and is thus likely to have more price inelastic demand.

However, big data, such as consumers' web browsing histories, give a much clearer picture of an individual person's

interests. Use of big data allows *Netflix* to identify individuals who are 91% certain to use their services and also individuals for whom there is a 0% chance.

Shiller estimates that using general, demographic information to practise third-degree price discrimination can enable *Netflix* to increase profits by 0.15%, but using big data can lead to a 1.4% increase in profit. The use of Google Glass and Smart TVs would provide more detailed data and almost certainly increase profits further.

Source: Article by Adam Ozimek in Forbes Magazine, 9.1.13

Discussion point

Is the use of big data to practise price discrimination unethical, or is it helping firms to satisfy consumers' needs more fully?

Review questions

Total: 35 marks

- 1 Define the term 'price discrimination'. (3 marks)
- 2 Before price discrimination, a firm sells 600 units of a service at a price of £20. It then splits its market and sells 200 units at £30 and 400 units at £18. Its total costs remain the same. Calculate the increase in its profit as a result of price discrimination. (4 marks)
- 3 Identify and explain one example of a firm/market that uses place as the basis of its price discrimination. (4 marks)
- 4 Identify and explain one example of a firm/market (excluding transport) that uses time as the basis of its price discrimination. (4 marks)
- 5 Explain *two* possible advantages to consumers arising from price discrimination. (8 marks)
- 6 Analyse *three* different conditions that are needed for price discrimination to take place in a particular market. (12 marks)

The dynamics of competition & competitive market processes

Key concepts from Year 1

Chapter 21 of the Year 1 companion textbook introduced the competitive market process by examining how competition is based on price and also studying alternative non-price factors that affect competitive market processes, such as improving products, reducing costs and improving the quality of the service provided. You should familiarise yourself with the material in that chapter before commencing this chapter.

This chapter on the dynamics of competition and competitive market processes examines both the short-run and long-run benefits that are likely to result from competition and the process of creative destruction. For the A-level, you should recognise that firms do not just compete on the basis of price – competition will, for example, also lead firms to strive to improve products, reduce costs and improve the quality of the service provided. This latter topic was covered in the Year 1 book.

The short-run and long-run benefits that are likely to result from competition

Competition brings a variety of benefits to consumers. These benefits vary between the short run and the long run.

Short-run benefits

- *Lower prices* – in competitive markets, especially perfect competition, price is the key feature of competition. Firms will strive to reduce production costs in order to stay competitive in the market, which rewards firms that can charge lower prices. Consequently, the consumer benefits because low prices mean that their incomes can help them to satisfy more of their wants.
- *Product variety* – in non-perfect markets, competition is often based on product differentiation. This encourages firms constantly to introduce new variations of products in order to satisfy consumers. By providing variety, consumers are more likely to be able to purchase the exact product that they prefer rather than having to be satisfied with a standardised, homogeneous product.
- *Innovation and invention* – in the short run, firms in competitive markets can gain high supernormal profits by introducing a new good or service. In the long run, other firms will produce this product, lessening the profits available until only normal profits are made.
- *Improving quality* – product differentiation often takes the form of improvements

in quality in order to attract more consumers. Consequently, consumers benefit from better quality products and better after-sales services.

Long-run benefits

- In perfect competition, competition constantly forces down costs and therefore price, as only normal profits are made in the long run. In this type of market, low prices are the main benefit for consumers.
- In monopolistic competitive or oligopolistic markets, the main benefit for consumers is a greater variety of better quality products, especially in oligopoly markets where supernormal profits remain in the long run.
- Both productive efficiency and allocative efficiency are achieved in a perfect market in the long run. This will be explained in Chapter 20.

The process of creative destruction

Key term

Creative destruction is an incessant process by which innovation and new technology constantly lead to the introduction of new production units that replace outdated ones.

In Chapter 10 we saw how technological change can lead to the establishment and development of new markets. Sometimes the new markets created by technological change and/or innovation challenge existing markets.

The idea of new markets replacing old markets has been established for many years. In 1942, Joseph Schumpeter introduced the term ‘creative destruction’ to describe this process. He argued that innovation and new technology create new products and new markets. Often, these new products and markets completely replace the old products – the process known as creative destruction.

Schumpeter believed that long-wave economic trade cycles were caused by sudden and significant changes in technology, innovation or invention. These innovations gave market power to their owners, enabling them to compete effectively with well-established but technologically inferior existing firms. Eventually the new entrants would gain market power at the expense of the established firms. Schumpeter saw creative destruction as a process of technological or inventive transformation, often forcing out an existing oligopolistic market and transforming it in the short run into a more competitive market, based on product differentiation. However, in the long run the new firms would dominate the industry and so the market would usually revert to oligopoly, but with different market leaders.

In 1995, Christensen and Bower presented the idea that creative destruction is often caused by either ‘disruptive innovation’ or ‘disruptive technology’. Where the innovation is not based on new technology, the term ‘disruptive innovation’ is used to describe the cause of change. The term ‘disruptive technology’ is used to describe a situation in which it is technological advancement that is leading to the displacement of an existing market.

Usually creative destruction results from the actions of a small competitor or new entrants into a market. Even where the new technology is widely recognised and available, the market leaders are reluctant to take actions that would disrupt a market in which they are dominant. For example, when mobile telephony was first developed, market leaders such as BT had high levels of sunk costs, in the form of landline cables. BT was also earning high profits from its dominant market position. Introducing a product that might lead to a completely new market was a very high-risk strategy and so BT was reluctant to take actions that might endanger the future existence of the market that it dominated. As a consequence, mobile telephone networks tended to be formed by new entrants into the market.

REALWORLD ECONOMICS 18.1

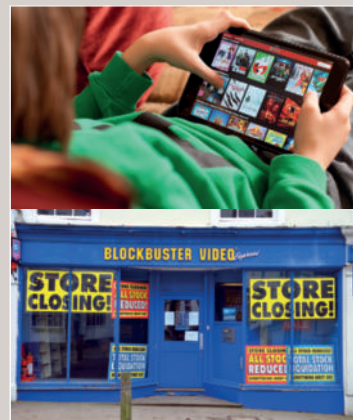
Netflix and creative destruction

Netflix was formed in 1997. Initially, it offered a traditional pay-per-rental model for customers wishing to hire videos. In 1999 it introduced a monthly subscription.

Netflix's early success was partly based on innovation and technology – the firm had developed an algorithm that was accurate in predicting customer ratings of films. In 2007 it introduced 'video on demand' via the internet. 'Streaming' of videos led to a doubling of sales over the next three years.

This success hit competitors, such as Blockbuster, which filed for bankruptcy in 2010. Figure 18.1 shows how employment in the USA in the video and disc rental industry grew from just over 80,000 people in 1985 to a peak of 170,000 people in 2000. In the last decade (2005–15) employment fell by 93%, with only 11,000 people now employed in the industry.

Netflix and other video streaming companies, such as Amazon, have also caused creative destruction to cable networks, such as Time-



The rise of Netflix hit competitors such as Blockbuster

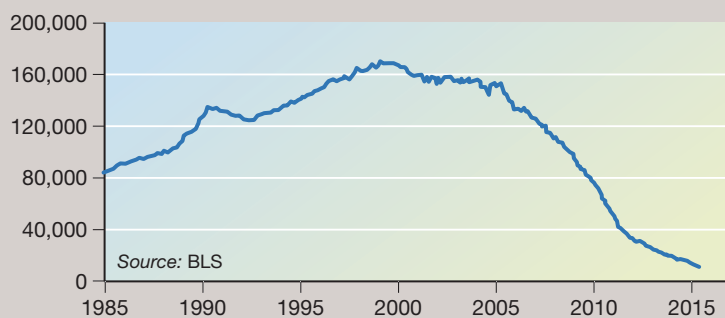


Figure 18.1 US employees in the video and disc rental industry, 1985–2015

Warner and Fox. Pay-TV providers are losing customers as people switch to streaming videos.

Source: American Enterprise Institute, www.aei.org

Discussion points

- 1 Is creative destruction in this industry likely to continue?
- 2 In what other industries do you anticipate creative destruction in the near future?

Review questions

Total: 20 marks

- 1 Competition leads to lower prices from new firms entering the market. In which market structure(s) is this a long-term benefit to customers?
 - A Perfect competition and monopolistic competition
 - B Perfect competition only
 - C Oligopoly and monopoly
 - D Monopoly only
- 2 Define the term 'creative destruction'.
- 3 Explain *one* long-term benefit that might occur in an oligopolistic market.
- 4 Explain *two* short-term benefits of competition.
- 5 Explain how creative destruction can significantly change a market.

(1 mark)

(3 marks)

(4 marks)

(6 marks)

(6 marks)

The concepts of contestable & non-contestable markets

Key concepts from Year 1

The concepts of contestable and non-contestable markets represent an alternative theory to the neoclassical theory of competition which is based primarily on the number of firms in an industry and the existence or non-existence of barriers to entry. An understanding of market structures is all that is required from Year 1.

This chapter introduces the concept of contestable markets and contrasts this theory with the neoclassical theory of market structures. The significance of market contestability for the performance of an industry is examined and concepts such as sunk costs and hit-and-run competition are analysed.

Introduction

The theory of contestable markets was devised by W.J. Baumol in the 1980s. Baumol believed that price and output in a market do not depend on market structure, such as perfect competition or monopoly, but on the potential for new firms to enter or exit the market. This is based on the view that, regardless of their monopoly power in a market, firms behave as if they are actually in a competitive market if there is a threat of competition from outside the market. The logic behind the theory of contestable markets is that firms that believe there is a threat from new entrants will act in a competitive way and make low profits, so as to discourage new entrants.

If there is no fear of new entry, such as a legal monopoly or government-funded natural monopoly, the market will be non-contestable. Similarly, a non-contestable market might be an oligopoly market in which existing firms have high brand loyalty and where barriers to entry are strong.

Key terms

A **contestable market** is one in which behaviour of existing firms is affected by the fear of new competition or takeover.

A **non-contestable market** is one in which existing firms are not affected by the fear of new competition or takeover and so they use any monopoly power that they possess in order to make supernormal profits.

Table 19.1 *The distinction between contestable and non-contestable markets*

Factor	Contestable markets	Non-contestable markets
Ease of entry	No barriers	Many barriers
Ease of exit	No barriers	Many barriers
Collusion	No collusion	Possible collusion
Information/knowledge	Good knowledge	Restricted knowledge
Influence of firms on price	Zero or high	Likely to be high
Number of firms	From one to many	From one to many
Product differentiation	Homogeneous or differentiated products	Homogeneous or differentiated products

Table 19.1 summarises the main distinguishing features used to define market structures and their application to contestable and non-contestable markets. The key differences between the two types of market are the ease of entry and the ease of exit.

The key difference between the theory of contestable markets and the traditional view of market structures is that the number of firms and the nature of products are not considered to be relevant factors in determining contestable markets. The degree to which a market is contestable therefore depends on the level of costs of entry and/or exit. A **contestable market** is one in which behaviour of existing firms is affected by the fear of new competition or takeover:

- *fear of new competition* – firms may reduce price to lower levels of supernormal profit. Thus potential competitors will not enter the market;
- *possibility of takeover* – firms without competition will try to reduce costs to boost share prices, so they are less likely to be taken over by another firm.

A **non-contestable market** is one in which there is little or no fear of new competition or takeover. Many of the former nationalised industries were seen as non-contestable. Similarly, firms with patents, high brand loyalty and economies of scale are likely to operate in non-contestable markets.

Neoclassical economic theory assumes that perfect competition and monopolistic competition have ease of entry and therefore they are contestable. However, oligopoly and monopoly are not contestable because there are barriers to entry. Such markets are more likely to have barriers to entry in the form of legal barriers, economies of scale and marketing/brand loyalty barriers. Contestable market theory also considers the cost of exiting a market. For example, at Christmas many temporary retailers enter the gift wrap market because supernormal profits are available at that time of year. Once Christmas is over, it is easy to exit the market because the firm has a short-term lease on a vacant store and there are minimal capital costs involved.

The significance of market contestability for the performance of an industry

In a contestable market, existing firms will recognise that supernormal profits will encourage new entrants. In the same way as competition acts as a restraint on supernormal profit in perfect competition, contestability will lead to normal profits because existing firms will not want to run the risk of new entrants.

Existing firms will try to ensure that there is limited scope for new entrants. Firms will therefore aim to reduce average total costs to a minimum in order to enable them still to make profit when price is low. If a profit-maximising firm only earns normal profit (in order to eliminate the incentive for new firms to enter the market), then its average revenue will equal its average total costs ($AR = ATC$). In order to prevent it being undercut by more efficient firms, the firm will operate at the lowest point of ATC. Thus it will have achieved productive efficiency. If there is only normal profit and the firm is maximising profit, then $MC = MR$. Since MC cuts ATC at the lowest point of ATC, $MC = MR$ where $ATC = AR$. As average revenue is the price, this means that price p (AR) equals MC. This is the requirement for allocative efficiency.

Conclusion: A fully-contestable market will achieve both productive efficiency and allocative efficiency.

It should be noted that since firms in contestable markets are constantly aware of

potential entrants, they will always try to find ways to stay ahead of competitors, possibly even trying to invent new products that can be patented (because this will prevent easy entry and make the market non-contestable). Thus contestable markets encourage dynamic efficiency too. (Dynamic efficiency will be covered in Chapter 20.)

Although the theory of contestable markets is used to argue that the number of firms in an industry is not necessarily the best determinant of competitiveness, many economists regard the concept of a perfectly contestable market as similar to perfect competition – an ideal that is unlikely to apply fully in real life.

Key term

Sunk costs are costs that cannot be recovered if a firm leaves the market, such as advertising. High sunk costs discourage firms from entering or exiting a market.

Concepts such as sunk costs and hit-and-run competition

Sunk costs tend to take the form of capital equipment and advertising expenditure. If a firm leaves an industry, the advertising expenditure will have been wasted. However, if capital equipment can be transferred to an alternative use, it is not a sunk cost, although some expenditure will have been committed.

High sunk costs represent the cost of exit. If a firm leaves a market, sunk costs will not be recovered. A firm should remain in the market if, and only if, expected revenue exceeds expected costs.

REALWORLD ECONOMICS 19.1

How contestable are UK rail franchises?

In the UK, trains are operated by privately owned franchisees known as the train operating companies (TOCs). These franchises are awarded by government for a given number of years. The franchisees are given sole rights to operate trains on particular routes.

When train travel was privatised in 1993, the TOCs were given seven-year franchises. At present there are over 20 TOCs running services, with some TOCs operating a number of different franchises. In order to avoid problems of sunk costs, the government sold the

rolling stock, such as engines and carriages, to separate rolling stock operating companies (ROSCOs), which lease the stock to the TOCs. In the UK, many of the ROSCOs are banks that can easily access funds to update stock. This means that it is much easier for TOCs to enter and exit the market.

Firms bid to run each franchise. Bids are assessed by the Department for Transport, based on factors such as the track record of existing franchises and promises of capital investment by the TOCs, but the highest bid is usually awarded the franchise.

Over the last 20 years TOCs have earned an average profit of around 5% of sales revenue. Existing TOCs are often successful in their bids to keep a franchise for a second or third term. It is believed that the current number of TOCs is close to the maximum limit of competition because of the importance of previous experience

when bids are judged.

In 2010 the length of franchises was extended to a minimum of 10 years, with the opportunity to extend the franchise to a maximum of 22 years. Franchisees must now meet performance criteria, based on factors such as punctuality and cancellations, to avoid having the franchise taken away. There are also major financial penalties for firms that do not complete the minimum term of their franchise. Another change has been that TOCs must now pay a non-returnable lump sum to make a bid and the successful bidder must also provide a sum of money that will be repaid if the term of the contract is completed.



Virgin Trains has operated the InterCity West Coast franchise since 1997

Exercises **Total: 20 marks**

- 1 Why do ROSCOs make it easier for TOCs to enter and exit the train operating market? **(5 marks)**
- 2 Evaluate the extent to which the train operating market is a non-contestable market. **(15 marks)**

Hit-and-run competition involves firms that are seeking a short-term profit. When profits fall, hit-and-run firms will exit the market. They must decide on whether it is worthwhile to join a market. Usually, hit-and-run competition occurs when sunk costs are low because firms can easily enter and exit a market and take advantage of any short-run supernormal profits available.

The potential for hit-and-run competition helps to keep markets contestable. If existing firms expect hit-and-run competition, they may lower profits in order to make entry unattractive. In this way the predictions of the model of contestable markets are fulfilled because the contestability of the market means that existing firms act as if they were in a perfectly competitive market.

Key term

Hit-and-run competition describes firms that enter markets where supernormal profits exist in order to make a quick profit.

Review questions

Total: 25 marks

- 1 Which *one* of these factors is most likely to apply to a contestable market?
 - A Many barriers to entry
 - B No barriers to exit
 - C Collusion between firms
 - D Restricted knowledge

(1 mark)
- 2 Which *one* of these factors is most likely to apply to a non-contestable market?
 - A Zero influence on price
 - B Perfect knowledge
 - C Barriers to entry
 - D No barriers to exit

(1 mark)
- 3 A non-contestable market may have both:
 - A Ease of entry and ease of exit
 - B Perfect knowledge and many firms
 - C Collusion and no barriers to entry
 - D Many firms and differentiated products

(1 mark)
- 4 Define the term 'contestable market'.

(3 marks)
- 5 What is meant by the term 'sunk costs'?

(3 marks)
- 6 What is meant by the term 'hit-and-run competition'?

(3 marks)
- 7 Explain why firms in contestable markets may only earn normal profit.

(5 marks)
- 8 Why are contestable markets likely to have both productive efficiency and allocative efficiency?

(8 marks)

Market structure, static efficiency, dynamic efficiency & resource allocation

Key concepts from Year 1

This topic builds on the coverage of market structures but is only dependent on work covered in previous chapters of this book.

Each of the market structures outlined in Chapter 11 is analysed in terms of efficiency. You need to understand the diagrammatic analysis of perfect competition (Chapter 13), monopolistic competition (Chapter 14), oligopoly (Chapter 15) and monopoly (Chapter 16). Although the concepts of productive efficiency, allocative efficiency and dynamic efficiency have been noted earlier in this book, this chapter assumes no prior knowledge of efficiency.

This chapter examines the difference between static efficiency and dynamic efficiency. It then considers the conditions required for productive efficiency (minimising average total costs) and allocative efficiency (price = marginal cost). The chapter concludes by studying how dynamic efficiency is influenced by research and development, investment in human and non-human capital, and technological change.

The difference between static efficiency and dynamic efficiency

Static efficiency assesses efficiency at a particular moment. Given the resources available to it, is the firm at its most efficient? If static efficiency exists, resources are being used to produce and allocate resources as effectively as possible. However, static efficiency does not mean that efficiency will improve in the future. The two main measures of static efficiency are productive efficiency and allocative efficiency.

Dynamic efficiency looks at changes over time. If firms are using new technology, innovation and invention to improve the effectiveness of their use of resources in the

Key terms

Efficiency, in terms of economics, measures how well inputs are used to produce an output. The word 'well' refers not just to the physical amount of output but also to the recipients of the output. Overall, efficiency

would not be achieved by maximising the number of units produced if all of these units were consumed by one person or a restricted number of beneficiaries.

Static efficiency measures efficiency at a point in time; productive efficiency and allocative efficiency are both measures of static efficiency.

Productive efficiency arises when production occurs at the lowest possible cost.

Dynamic efficiency looks at changes over time, arising from new technology, innovation and invention.

Technical efficiency is where output is maximised from the available inputs of factors of production.

future, then dynamic efficiency exists. Dynamic efficiency is important if firms are trying to ensure that efficiency is sustained in the long term.

There can be conflict between static and dynamic efficiency. If a firm focuses all of its efforts on producing goods as effectively as possible, it may be ignoring the future potential of a new invention. However, if the firm is aiming for dynamic efficiency, the expenditure required to make these changes may mean that it is unable to maximise the efficiency of its current-day operations.

The conditions required for productive efficiency (minimising average total costs) and allocative efficiency (price = marginal cost)

Economics is concerned with using scarce resources to satisfy unlimited wants. There are two main ways to measure whether this task is undertaken efficiently: productive efficiency and allocative efficiency.

Productive efficiency

Average total cost (ATC) is considered to be the key measure of production costs and so productive efficiency occurs at the lowest possible ATC. First, the production process must be as efficient as possible. This is known as technical efficiency. Productive efficiency is also achieved through the lowering of unit or average costs by other means. The main causes of improved productive efficiency are economies of scale (see Chapter 15 of the Year 1 book). As a consequence, firms try to grow until they are at the optimum size, which allows them to minimise average total costs. If a firm grows too much, it can incur diseconomies of scale, which leads to higher average total costs.

In the short run, productive efficiency occurs at the minimum point of the short-run average cost curve. This is a measure of static efficiency. In the long run, firms can modify their production techniques and ensure that they produce where the long-run average cost curve is at its minimum.

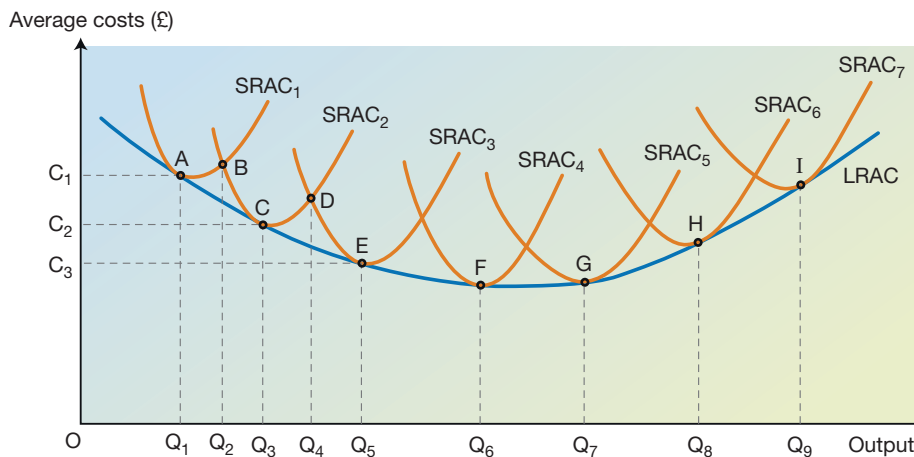


Figure 20.1 The average cost curve and productive efficiency

In Figure 20.1, the minimum points of each short-run average cost curve are productively efficient in the short run. Thus Points A, C, E, F, G, H and I all represent short-run productive efficiency. Points B and D show points where firms are not at their most productively efficient. However, these represent points at which a firm will

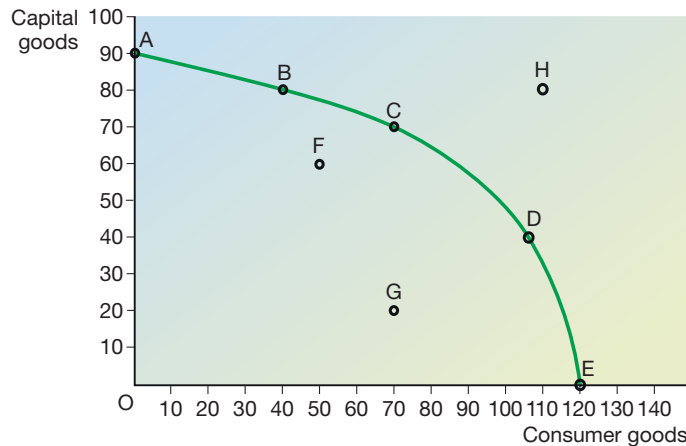
decide to change capacity (moving from an SRAC curve with low capacity potential to one with higher possible capacity).

Points F and G show where the firm is producing at the lowest possible level of average total costs. Thus F and G show long-run productive efficiency, although these points also show productive efficiency in the short run. In order to sustain productive efficiency in the long run, the firm's output should lie between Q_6 and Q_7 .

Conclusion: For productive efficiency to be achieved by a firm, output must be where average total costs (ATC) are at their minimum.

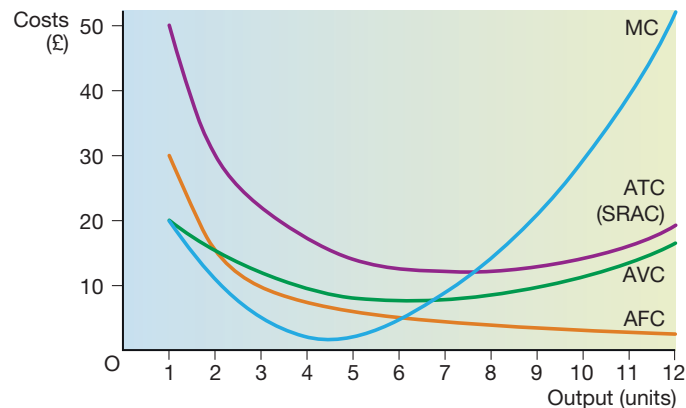
Productive efficiency can also be used to show the efficiency of an economy. In Figure 20.2, the line ABCDE shows the production possibility boundary for a particular economy. Goods produced by the economy are either capital goods or consumer goods. An economy is making effective use of its resources if it produces at the maximum possible level of output. Consequently, Points A, B, C, D and E all represent productive efficiency. Points F and G show productive inefficiency because they are within the boundary and thus output is not being maximised from the available level of inputs. Point H shows a combination of capital goods and consumer goods that it is impossible to reach, given the current resources available.

Figure 20.2 Production possibility diagram of the UK economy



In the long run, the minimum point of the long-run ATC curve shows productive efficiency. If this condition is met by all firms, the economy as a whole will be producing on its production possibility boundary.

Figure 20.3 Average costs and marginal costs



When examining costs in Chapter 6 of this book, we saw that the marginal cost (MC) curve cuts the average total cost (ATC) curve from below at the minimum point of the average total cost (ATC) curve. This is shown in Figure 20.3. Thus, where ATC is minimised, $MC = ATC$. It is therefore the case that productive efficiency occurs where MC equals ATC.

Allocative efficiency

Allocative efficiency occurs when overall economic welfare cannot be improved by reallocating resources between different markets. This arises where price in a market is equal to the marginal cost of the last item (i.e. $P = MC$). Why does $P = MC$ show allocative efficiency?

The demand schedule shows the value placed on a product by consumers. If a consumer is prepared to pay £5 for a good, then it can be seen that society's economic welfare as a whole will increase by £5 if the consumer buys that product. However, the production of that product uses up scarce resources. The costs of manufacturing that product show the value of resources needed to produce a product. Since the optimum situation is shown by the marginal item, the marginal cost of the last item shows the cost to society of making that last item. If $P > MC$, then society benefits because the benefit to the consumer's welfare exceeds the cost to society in terms of the scarce resources used to make that product. As long as $P > MC$, each product brings more benefit than it costs and so society's economic welfare increases. Similarly, if $P < MC$ for a particular product, then the resources used to make that product were more valuable than the benefit to the consumer. Thus society's overall economic welfare has fallen. In these circumstances, resources should be reallocated from goods where $P < MC$ to goods where $P > MC$. Reallocation should cease when the benefit from the good (its price) exactly matches the value of resources used to make it (its MC). Thus allocative efficiency occurs where $P = MC$. This applies to firms and society as a whole.

It should be noted that this analysis is based on the assumption that the consumer's willingness to pay for a product is the extent to which society benefits from it. If all people have the same incomes, allocative efficiency gives an accurate reflection of society's needs. However, where inequality is high, allocative efficiency can occur where some people have no food while others have excess.

Comparing the efficiency of different market structures

When assessing productive efficiency and allocative efficiency it is possible to draw conclusions on the relative efficiency of each market structure, in both the short and long run. The dynamic efficiency of each market structure will be compared too.

This section examines the efficiency of the four market structures that were covered in Chapters 13–16: perfect competition, monopolistic competition, oligopoly and monopoly.

The efficiency of perfect competition

Figure 20.4 shows whether perfect competition leads to allocative efficiency and productive efficiency. Point A_1 shows the initial equilibrium of a firm in a perfectly competitive market. P_1 shows the initial short-run equilibrium price and so AR_1 is the initial average revenue line for the firm. In perfect competition, the marginal

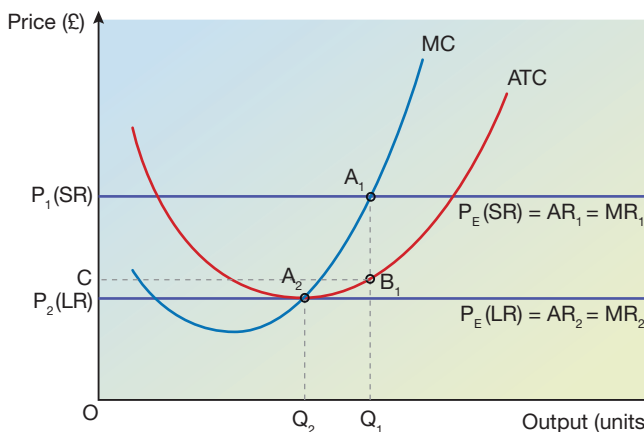
Key term

Allocative efficiency

measures whether the goods and services produced are those that consumers want. It means that no consumer can be made better off without another consumer becoming worse off.

revenue is equal to average revenue and so $AR_1 = MR_1$. A_1 is the initial equilibrium point because $MC = MR_1$ at this point and so the firm is maximising profit. AR exceeds ATC by the distance A_1B_1 , so A_1B_1 represents supernormal profit per unit. The equilibrium output is Q_1 and so supernormal profit is $A_1B_1 \times Q_1$.

Figure 20.4 Perfect competition and productive efficiency and allocative efficiency



In the short run, total revenue = $p \times q$, shown by the area $OP_1A_1Q_1$. Total costs = $ATC \times q$, shown by the area OCB_1Q_1 . Supernormal profit = $TR - TC$, shown by the area $CP_1A_1B_1$. At Point A_1 , price and marginal costs are both equal to P_1 . $P = MC$ is the condition necessary for allocative efficiency. **Thus perfect competition creates allocative efficiency in the short run.** In the short run, average total cost is B_1 , which is slightly higher than the lowest point of the ATC curve, which is shown by A_2 . **Perfect competition does not therefore create productive efficiency in the short run.**

In the long run, the supernormal profits encourage new firms to enter the market. In the market, this leads to an increase in supply and so the market equilibrium price falls. New firms will keep entering the market until there is no supernormal profit available. In Figure 20.4 this is shown by P_2 . At price P_2 , AR_2 shows the new average revenue line; this is identical to MR_2 , the new marginal revenue line. MC equals MR_2 at Point A_2 and so this is the long-run equilibrium point. Therefore, Q_2 shows the firm's equilibrium output. Since average revenue is equal to average total cost at Point A_2 , there is no supernormal profit.

In the long run, total revenue = $p \times q$, shown by the area $OP_2A_2Q_2$. Total costs = $ATC \times q$, shown by the area $OP_2A_2Q_2$. Supernormal profit = $TR - TC$. This is zero because $TR = TC$. At Point A_2 , price is P_2 and MC is also P_2 . Therefore $P = MC$. **Thus perfect competition creates allocative efficiency in the long run.** At Point A_2 , average total costs are at their minimum point. This is the condition for productive efficiency. **Perfect competition therefore creates productive efficiency in the long run.**

In theory, perfect competition does lead to efficiency. However, small firms are unlikely to achieve the economies of scale of large firms. The ATC for a small firm is likely to show much higher levels of costs than the ATC for a much larger firm, and so A_2 does not show the minimum possible cost. Furthermore, demand ignores externalities. It could be argued that allocative efficiency should therefore be based on social benefits and social costs rather than on demand and supply. The 'best' allocation of resources is not where $P = MC$, but where $MSB = MSC$.

Dynamic efficiency arises from a willingness to devote resources to improving

production methods and creating new, differentiated products. Since products are homogeneous, there is no incentive to modify products. Moreover, lowering production costs only brings a short-run supernormal profit because perfect knowledge means that other firms can copy new production techniques and so the cost advantage is soon lost. *Overall, perfect competition does very little to encourage dynamic efficiency.*

The efficiency of monopolistic competition

Figures 20.5 and Figure 20.6 show whether monopolistic competition leads to allocative efficiency and productive efficiency.

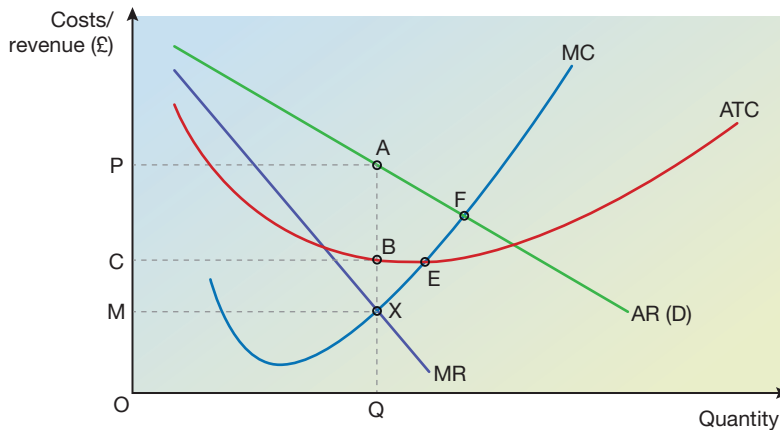


Figure 20.5 Monopolistic competition and efficiency in the short run

In monopolistic competition, each firm produces its own slightly differentiated product. There is therefore no demand curve for the market as a whole, but rather a series of individual demand curves for each differentiated product.

In Figure 20.5, the line AR (D) shows the demand curve (and therefore the average revenue line) for a particular firm's product within the market. The marginal revenue line (MR) slopes downwards twice as steeply.

There is no supply curve as such in a monopolistically competitive market. Each firm is free to choose its level of output. However, because a firm maximises profit when $MC = MR$, the chosen level of output will be determined by the combination of MC and MR. In Figure 20.5, MC cuts MR from below at Point X. This shows the output that maximises profit for the firm in the short run and so it will produce OQ units. Since OQ units are supplied onto the market, the line QXBA in effect shows the firm's supply on to the market. With OQ units being supplied, demand is met at Point A and so the equilibrium price for the firm's product is P. Point B shows the average total cost of producing OQ units of output.

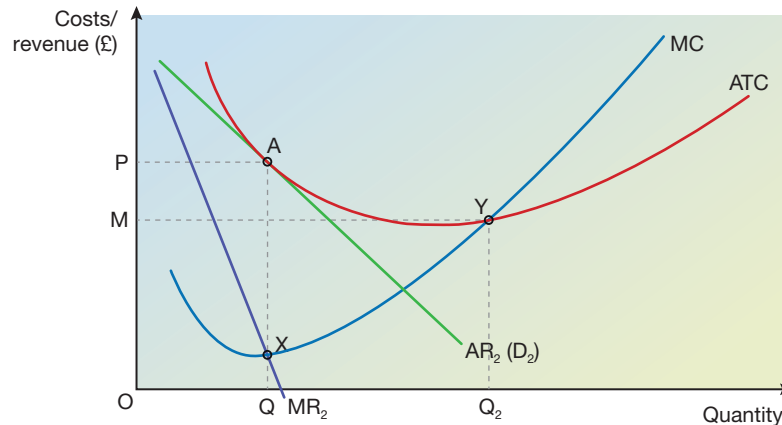
At output Q, the price is P and the marginal cost of the Q^{th} item is M. Therefore $P > MC$ and so *monopolistic competition does not lead to allocative efficiency in the short run*. At output Q, the average total cost (ATC) is shown by Point B. The lowest point of ATC is shown by Point E, where $MC = ATC$. *Monopolistic competition does not therefore lead to productive efficiency in the short run.*

Since B is lower than A, supernormal profit is made. Supernormal profit per unit of output is shown by $A - B$. $TR = p \times q$ and therefore $TR = OP \times OQ$, shown by area OPAQ. $TC = ATC \times q$ and therefore $TC = OC \times OQ$, shown by area OCBQ. Total

profit is $TR - TC$, shown by area CPAB. (Note that as normal profit is included in costs, this profit is ‘supernormal profit’ or ‘abnormal profit’.)

Since there are no or few barriers to entry into a monopolistically competitive market, new firms will enter the market in order to gain a share of these supernormal profits. Figure 20.6 shows the impact of these new entrants into the market on the individual firm.

Figure 20.6 Monopolistic competition and efficiency in the long run



The new entrants provide close substitutes and so demand for the firm’s product will fall, as demand will be spread amongst a much greater number of firms.

Figure 20.6 shows the long-run equilibrium. Demand keeps falling until only normal profits are made. In this case, the lower demand is shown by $AR_2 (D_2)$; MR also shifts to the left. Eventually, AR shifts until there is just one point at which $AR = ATC$. This is shown by Point A in Figure 20.6. Point X (which is vertically below Point A) is where $MC = MR$, confirming that this is the profit-maximising output. Thus, the firm will produce OQ units. This is the long-run equilibrium because with supernormal profits no longer in existence, new firms will no longer enter the market. Both TR and TC can be measured by the rectangle OPAQ.

In Figure 20.6, price is shown by Point A and MC is shown by Point X. Price is much higher than MC. The distance AX shows the difference between P and MC. *This indicates that monopolistic competition is a long way from achieving allocative efficiency in the long run.* The profit-maximising output is Q. Point A shows that ATC is much higher than the lowest point of ATC, which is at Point Y. *Productive efficiency is not achieved by monopolistic competition in the long-run equilibrium. In fact, it is less productively efficient than the short-run equilibrium.*

Dynamic efficiency provides short-run benefits to firms in monopolistic competition because it can both reduce costs and achieve more product differentiation. However, due to perfect knowledge, new firms will match these benefits and so dynamic efficiency will only help if it creates a barrier to entry, such as a patented product.

The efficiency of oligopoly

Please note that the following analysis is based on the assumption of a kinked demand curve. If there is no kinked demand curve and no collusion, the diagrammatic analysis shown in the previous section on monopolistic competition will apply, with two major exceptions. First, there is no change in the long run because oligopoly has barriers

to entry. Secondly, due to restricted competition in oligopoly, it is probable that the demand curve will be much more price inelastic in oligopoly than in monopolistic competition. If there is no kinked demand curve but collusion exists, then the oligopolists are acting as a monopoly. The diagrammatic analysis for monopoly (in the next section) will therefore apply.

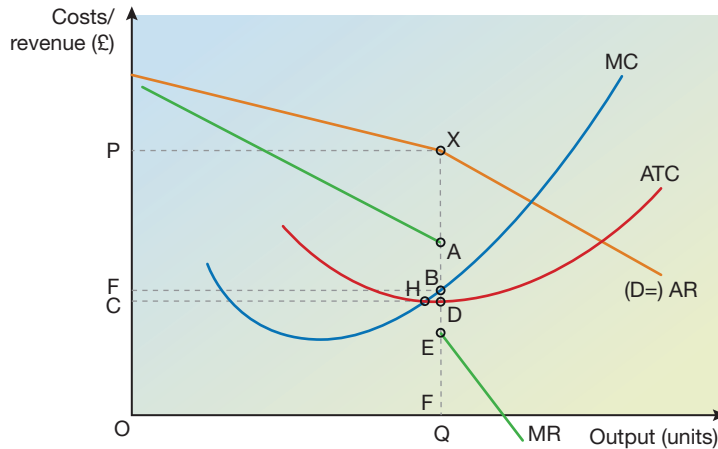


Figure 20.7 Oligopoly and efficiency

Figure 20.7 shows whether oligopoly leads to allocative efficiency and productive efficiency. The original price is P with Q units sold. The kink in demand is at Point X . Profit maximisation occurs where $MC = MR$. MC cuts MR from below at Point B , so the profit-maximising output for the oligopolist is OQ units. Price is OP . At Q units of output, marginal cost is shown by Point B , so MC per unit is OF . Price exceeds MC by a significant margin. *Oligopoly does not therefore achieve allocative efficiency.* At output Q , ATC is shown by Point D . This is quite close to the lowest point of ATC , which is at Point H . However, *oligopoly does not achieve productive efficiency.*

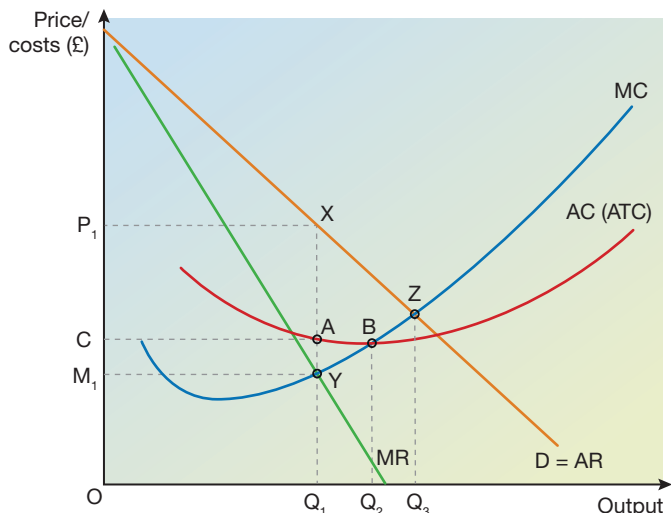
Since $TR = OPXQ$ and $TC = OCDQ$, supernormal profit = $CPXD$. However, there are barriers to entry to monopoly markets and so the short-run equilibrium is identical to the long run. Thus the analysis above applies to both time periods.

Dynamic efficiency may be more likely in concentrated markets, especially oligopoly. In oligopoly, superior innovation and quality create barriers to entry. Thus oligopolists use non-price competition to compete. They are therefore more likely to create dynamic efficiency by focusing on factors such as research and development, investment in human skills and training, investment in capital equipment and encouragement of technological change.

The efficiency of monopoly

Figure 20.8 shows whether allocative efficiency and productive efficiency are achieved by monopoly. In Figure 20.8, $MC = MR$ at Point Y and so the monopolist's equilibrium output is Q_1 . The monopolist will produce Q_1 goods and the demand (average revenue) line shows that the equilibrium point will be Point X . Thus the equilibrium price will be P_1 . ATC is shown by A (or line Q_1A). The lowest point of ATC is B (line Q_2B), so Q_2 is the output that minimises ATC (and thus achieves productive efficiency). *Monopoly does not therefore create productive efficiency.* However, in this case economies of scale might mean that Q_1A is much lower than costs in perfect competition.

Figure 20.8 Monopoly and efficiency



At output Q_1 , price is P_1 but MC is M_1 . P_1 is much greater than M_1 . *Since $P > MC$, monopoly does not deliver allocative efficiency.* Again, due to the potential for lower costs with monopoly, this difference may exaggerate the monopolist’s inefficiency in comparison to perfect competition – with very low costs it is feasible that the monopoly price might be lower than the price in perfect competition. However, there is clear evidence of inefficient allocation of resources. The Q_1^{th} item added Q_1X to society’s welfare but only used Q_1Y of the country’s resources. The production of more units would thus have led to more allocative efficiency. Since allocative efficiency is where $P = MC$, Point Z shows allocative efficiency. Thus the production of Q_3 items would have maximised allocative efficiency.

For the monopolist, $TR = p \times q$, shown by the area OP_1XQ_1 . Total costs = $ATC \times q$, shown by the area $OCAQ_1$. Supernormal profit = $TR - TC$, shown by the area CP_1XA .

There are barriers to entry to monopoly markets and so the short-run equilibrium is identical to the long run. Thus the analysis above applies to both time periods.

Dynamic efficiency may occur in monopoly in order to strengthen barriers to entry. Some goods may also compete with goods in different markets and so there is an incentive to cut costs. Furthermore, cutting costs leads to a larger supernormal profit. However, in comparison to oligopoly there is less likely to be dynamic efficiency in a monopoly market because there is less need to introduce new goods.

How do the different market structures compare?

Table 20.1 provides a summary of efficiency based on market structures.

Table 20.1 Efficiency and market structures

Market structure	Productive efficiency	Allocative efficiency	Dynamic efficiency
Perfect competition: short run	No	Yes	Very limited
Perfect competition: long run	Yes	Yes	Very limited
Monopolistic competition: short run	No	No	Limited
Monopolistic competition: long run	No	No	Limited
Oligopoly: short run and long run	No	No	High
Monopoly: short run and long run	No	No	Variable

REALWORLD ECONOMICS 20.1

The National Lottery and efficiency

Camelot, which runs the UK's National Lottery, claims that it is the most cost effective in Europe. This claim is based on the operating costs. Just over 4% of the cost of a lottery ticket goes on operating costs, with a little under 1% being profit after tax. However, the purpose of a lottery is to use the revenue to fund good causes. Therefore, the full breakdown of costs per ticket in 2015 was as follows:

- Good causes – 25%
- Prizes – 53%
- Lottery tax – 12%
- Retailer's commission – 5%
- Operating costs – 4%
- Profit – 1%



Is the National Lottery efficient?

As the lottery has expanded, operating costs have fallen slightly, but mainly through new technology. The long-run average total cost would appear to be L-shaped. With an L-shaped LRAC, regardless of output Camelot's ATC is likely to be close to or at its minimum point. This suggests that, unlike most monopolists, it is productively efficient. On the other hand, with only one operator and a guaranteed profit, there is little incentive for Camelot to be dynamically efficient, and so it is not certain that its ATC is at the lowest possible level. However, there is an incentive. 5% of a ticket's price goes to Camelot – if it can reduce operating costs, its profit margin will increase.

For allocative efficiency, price must equal marginal cost. This is clearly not the case, as the price of £2 is slightly higher than the additional cost of supplying a ticket, although a fixed 95% of the price is spent on the first four items above.

The small profit margin suggests that P and MC are very close, indicating high (but not perfect)

allocative efficiency. However, this measure is based on the market mechanism. The £2 paid by lottery ticket buyers is for a dream. After purchase, most buyers will receive a zero return.

Analysis by Theos, a theological think tank, suggests that the lottery is a regressive tax. People on low incomes spend a higher percentage of their incomes on lottery tickets. However, lottery funds are distributed in favour of good causes that benefit people on high incomes, in part because funding is often dependent on the ability of a good cause to raise some money itself to go alongside the lottery money. Two opera houses and three London theatres have received a total of £315 million from the lottery.

*Sources: Camelot, www.theosthinktank.co.uk and an article by Liz Hill in *Arts Professional*, 25.4.14*

Discussion point

'In terms of productive efficiency and allocative efficiency, the UK's National Lottery is highly efficient.' Critically evaluate this statement.

Influences on dynamic efficiency

There is no specific measurement of dynamic efficiency. It arises where firms are devoting resources to activities such as research and development, investment in human and non-human capital, and encompassing technological change. All of these actions help firms to increase their efficiency so that, in the long run, static measures such as productive efficiency show improvements.

- **Research and development (R&D):** R&D can lead to improvements in production processes. These improvements will enable firms to achieve lower average costs and therefore earn higher profits. In some cases, these improvements take the form of more efficient use of scarce resources, and so society as a whole can benefit from improved sustainability of its resources. R&D can also improve the quality of products and therefore achieve product differentiation in concentrated

markets. Ultimately, effective research and development can produce new goods and services, which can be protected by patents or copyright. This will enable the firm to retain a monopoly of this good or service for up to 20 years.

- **Investment in human capital:** In many markets the skills of the workforce are crucial to the success of the firms operating within that market. In primary and secondary production, a skilled workforce can increase labour productivity and therefore lower costs of production. In service industries, the quality of the service enjoyed by customers is often closely related to the skills of the workforce. Firms invest in human capital by providing training and development to enhance existing skills. There is close correlation between the skills of the workforce and other factors leading to dynamic efficiency, such as innovation and successful R&D. A highly trained workforce is more likely to be inventive and provide the foundation for dynamic efficiency within a firm or an economy as a whole.
- **Investment in non-human capital:** Productive efficiency is very closely related to fixed capital formation (i.e. investment in capital goods). Capital goods, such as machinery, enable the firm to increase the productive efficiency of its operations and therefore also improve the efficiency of other factors of production, such as labour.
- **Technological change:** New technology can lead to the invention of new products, innovation that brings about more efficient methods and processes, and successful R&D that provides new, more advanced goods.

Review questions

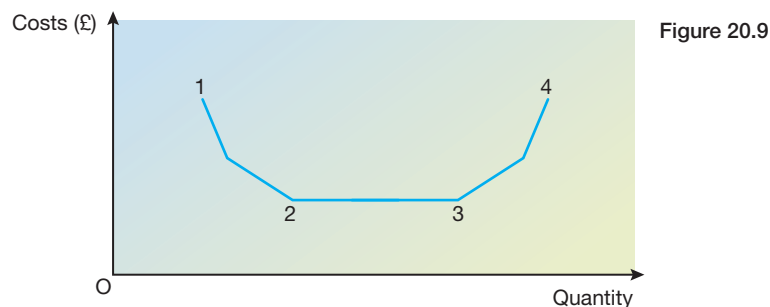
Total: 45 marks

- 1 Productive efficiency occurs where:

A MC = MR	C P = MC	(1 mark)
B ATC = MC	D ATC = AR	
- 2 Allocative efficiency occurs where:

A MC = MR	C P = MC	(1 mark)
B ATC = MC	D ATC = AR	
- 3 Which *one* of the following statements is correct? Productive efficiency:

A For an economy is achieved on the production possibility boundary	(1 mark)
B Involves maximum input from a given level of output	
C Is improved by internal diseconomies of scale	
D Is not affected by external economies of scale	
- 4 Figure 20.9 shows ATC for a firm.



Productive efficiency occurs:

- A Between Points 1 and 2
- B Only at Point 2
- C Between Points 2 and 3
- D Between Points 3 and 4

(1 mark)

- 5** 'No consumer can be made better off without another consumer becoming worse off.' This quote fully describes:

- A Allocative efficiency
- B Dynamic efficiency
- C Productive efficiency
- D Static efficiency

(1 mark)

- 6** Productive efficiency is certain to occur in:

- A Perfect competition in the long run
- B Monopolistic competition in the short run
- C Oligopoly
- D Monopoly

(1 mark)

- 7** Allocative efficiency is achieved in:

- A Perfect competition in the short run
- B Monopolistic competition in the long run
- C Oligopoly
- D Monopoly

(1 mark)

- 8** Define 'productive efficiency'.

(3 marks)

- 9** What is the difference between static efficiency and dynamic efficiency?

(4 marks)

- 10** Explain why monopoly might be more productively efficient than perfect competition?

(4 marks)

- 11** Explain how research and development might influence dynamic efficiency.

(4 marks)

- 12** Explain how investment in human capital might influence dynamic efficiency.

(4 marks)

- 13** Explain why allocative efficiency might not represent a perfect allocation of resources.

(5 marks)

- 14** Draw a diagram to show the levels of productive efficiency and allocative efficiency in perfect competition in both the short run and long run.

(8 marks)

- 15** Draw a diagram to show the levels of productive efficiency and allocative efficiency of a monopoly.

(6 marks)

Consumer & producer surplus

Key concepts from Year 1

Consumer and producer surplus are based on supply and demand. No other concepts from Year 1 are required.

This chapter introduces the concepts of consumer surplus and producer surplus and shows how they can be measured on a demand and supply diagram. The chapter also shows the application of these concepts when discussing economic efficiency and welfare issues, such as price discrimination and the deadweight losses associated with monopoly.

Consumer surplus and producer surplus

Consumer surplus is a measure of economic welfare for consumers, as it shows the value they receive from a good over and above the price that they paid for that good.

Producer surplus is a measure of economic welfare for producers, as it shows the price received over and above the price they would have been prepared to accept.

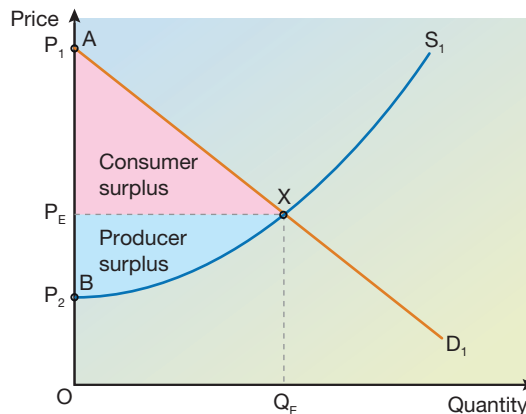
Figure 21.1 shows equilibrium in **perfect competition**. The analysis below therefore shows consumer surplus and producer surplus in a perfectly competitive market.

Key terms

Consumer surplus is the difference between the price a buyer would pay and the price actually paid by the buyer.

Producer surplus is the difference between the price received by a firm and the price it would have been prepared to accept in return for supplying.

Figure 21.1 Consumer surplus and producer surplus



In Figure 21.1, D_1 shows that the first consumer would have paid P_1 . Equilibrium price is P_E and so consumer surplus for the first consumer is $P_1 - P_E$. All consumers have consumer surplus until Q_E because the value that they place on the good (shown by the price which they are prepared to pay) exceeds the price that they actually pay (the equilibrium price of P_E). The consumer at Q_E earns no consumer surplus because they were prepared to pay P_E and they do pay P_E . Therefore, the value they receive from the good is exactly the same as the price they pay. For each consumer between O and Q_E , the consumer surplus is the vertical distance between

the demand curve and the equilibrium price. It can be seen that the maximum individual consumer surplus is earned by the first consumer (who is prepared to pay P_1). Individual consumer surplus falls for each additional consumer because each one places a lower value on the good than the previous consumer. The shaded area of triangle $P_E P_1 X$ is thus total **consumer surplus** for buyers. Beyond output Q_E , each consumer values the good at a lower price than the equilibrium price. Thus they will not purchase this good.

In Figure 21.1, S_1 shows that the first unit supplied would have been offered for P_2 but the price actually paid was the equilibrium price P_E . Thus producer surplus for the first item supplied is $P_E - P_2$. Producer surplus is earned from each product until Q_E , for which it is zero ($P_E - P_E$). The vertical distance between P_E and the supply curve shows the producer surplus for each individual good supplied. Thus the shaded area $P_2 P_E X$ is the total **producer surplus** for sellers. Beyond output Q_E , each unit produced costs more than the price that would be received from supplying the good (the equilibrium price). Thus the firm will not supply this good once output reaches Q_E items.

Application of consumer surplus and producer surplus to economic efficiency and welfare issues

Consumer surplus and producer surplus are measures of welfare because they show how much benefit consumers and producers receive from a good in comparison to the price paid/received.

In Figure 21.1, the first consumer receives a benefit equal to OP_1 from the good. The second consumer receives a benefit of just below OP_1 . The demand curve shows the level of benefit received. The only consumer with no consumer surplus is the consumer at Q_E who values the good at P_E . The total welfare received from all consumers is thus the area shown by $OP_1 X Q_E$.

The price paid by each consumer is P_E and Q_E items are bought. Thus expenditure by consumers is shown by the area $OP_E X Q_E$. Consumer surplus is the amount by which the consumers' welfare exceeds the amount spent. This is the area $P_E P_1 X$.

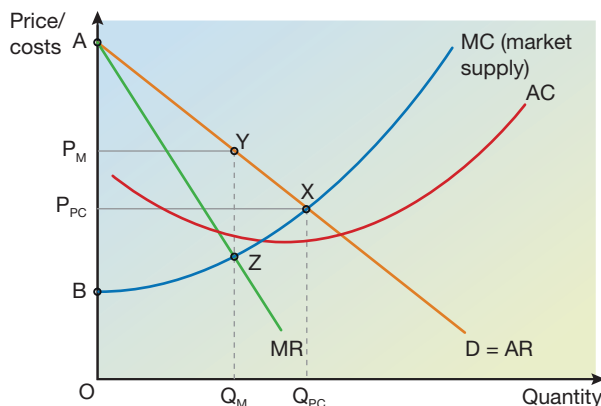
Similarly, the amount of money received by suppliers is the area $OP_E X Q_E$. However, the producer would have been prepared to supply the first item for P_2 , the second item for a price marginally above P_2 , and so on. If each good had been sold separately, the producer would have been satisfied with revenue shown by the area $OP_2 X Q_E$. Producer surplus is the amount by which the producers' revenue exceeds the amount they would have been prepared to accept in return for supplying each good. This is the area $P_E P_2 X$.

In Figure 21.1, the consumer surplus appears to be greater than the producer surplus. This arises because of the shapes of the demand curve and the supply curve. In this diagram, some consumers are prepared to pay a very high price – overall the demand is relatively price inelastic. However, the supply curve is more price elastic and so there is less difference between equilibrium price and the price suppliers are prepared to accept. In general terms, the more price inelastic the demand curve, the greater the level of consumer surplus received. Similarly, the more price inelastic the supply curve, the greater the level of producer surplus received.

As noted earlier, Figure 21.1 applies to perfect competition. How does monopoly affect consumer and producer welfare? Figure 21.2 shows equilibrium, consumer

surplus and producer surplus in a **monopoly**. Please note that market supply is the same as the MC line. Thus the MC line in this diagram is identical to the supply curve in Figure 21.1. Similarly, the demand curve in Figure 21.2 is identical to the demand curve in Figure 21.1. This allows us to compare consumer surplus and producer surplus between perfect competition and monopoly.

Figure 21.2 Deadweight welfare loss from monopoly



In the monopoly market, the monopolist's $MC = MR$ at Point Z and so equilibrium output is Q_M . At output Q_M , Point Y shows where the amount offered for sale equals demand and so the monopolist's equilibrium price is P_M . This is higher than the perfectly competitive price of P_{PC} . Only Q_M units are sold, rather than Q_{PC} . Thus monopoly leads to a lower equilibrium quantity and a higher equilibrium price. In the monopoly market, it can be seen that total consumer surplus is P_MAY , while total producer surplus is BP_MYZ . In relative terms, producers in a monopoly market have a larger share of the surplus than in a perfectly competitive market. The total welfare (consumer surplus + producer surplus) is $BAYZ$ in a monopoly market.

More significantly, there is an overall loss in economic welfare in a monopoly market in comparison to a perfectly competitive market. In Figure 21.2 it can be deduced that in perfect competition the equilibrium point would have been X (where market supply equals demand). This would have led to a consumer surplus of $P_{PC}AX$. The producer surplus would have been $P_{PC}BX$. Thus the combined consumer and producer welfare would have been AXB . AXB represents a higher level of welfare than $BAYZ$. The difference between the two is shown by the area YXZ . YXZ shows overall loss of welfare (known as the **deadweight welfare loss**) arising from monopoly.

The analysis above assumes that the monopolist sells all output at one price. What happens if the market conditions allow the monopolist to use price discrimination? Price discrimination allows the monopolist to earn additional revenue, thus shifting the balance of welfare towards the producer and away from the consumer. The data below apply to a market before and after price discrimination.

Before price discrimination

Price = £20 Quantity demanded = 100 TR = £2000

The quantity demanded equals 100: 50 demanded by consumers who have price inelastic demand and 50 demanded by consumers who have price elastic demand. If these two groups of consumers can be identified and separated, then the monopolist can use price discrimination by splitting them into two separate markets.

After price discrimination

Market A is the group of consumers with price inelastic demand.

Market B is the group of consumers with price elastic demand.

A price discriminating monopolist will increase price in Market A and decrease price in Market B. Since demand is price inelastic in Market A, a £10 price increase from £20 to £30 only leads to the loss of 10 customers (from 50 to 40). In Market B, where demand is price elastic, a £2 decrease in price from £20 to £18 attracts an additional 10 customers (from 50 to 60). Consequently, the same total number of customers are using the service provided, but the monopolist has increased total revenue.

$$TR = (40 \times £30) + (60 \times £18) = £1200 + £1080 = £2280$$

The monopolist's total revenue has increased by £280, from £2000 to £2280, but total costs remain the same because the number of consumers has not changed. In effect, the monopolist has increased its producer surplus at the expense of the consumer surplus of its customers.

REALWORLD ECONOMICS 21.1

Bidding pricing and consumer surplus and producer surplus

Where items are in limited supply, such as antique furniture, firms often use auctions to sell individual items or ask prospective buyers to put in a sealed-bid written offer to pay a certain sum for a good. By selling items individually, the firm can avoid facing a situation in which the market price is determined by the marginal consumer.



Selling at auction – one method by which suppliers seek to minimise consumer surplus and increase producer surplus

If an individual item is sold at an auction, the price paid tends to be marginally higher than the value placed on it by the second-highest bidder. In an auction, a good is sold to the consumer who places the highest value on that good. However, to secure ownership of the good that person must pay more than the person who places the second-highest value on the good. The consumer surplus is thus approximately the difference between the valuation of the first consumer and the valuation of the second consumer. This leads to a small consumer surplus.

If more versions of this item are available, they can be sold at subsequent auctions. In the next auction, the person who places the second-highest valuation on the good will secure its ownership by outbidding the person who has the third-highest valuation. Again, this

will lead to a very small consumer surplus. Auctions can thus be a very effective way for a supplier to minimise consumer surplus and increase the level of producer surplus.

This analysis assumes that each consumer is aware of the existence of an auction but unaware that there will be a subsequent auction. The more a firm uses this technique, the less effective it will become because consumers will defer bidding for the good and this is likely to force the price down. For sealed bids, consumers may bid high (the full value that they place on it) to secure the item or they may collude with other buyers to ensure a low price.

Discussion point

How feasible is it for firms to eliminate or reduce consumer surplus through the use of techniques such as auctions and sealed bids?

Review questions

Total: 25 marks

- 1 Which *one* of the following would lead to a higher consumer surplus?
 - A Demand being price elastic
 - B Demand being price inelastic
 - C Supply being price elastic
 - D Supply being price inelastic

(1 mark)

- 2 Which *one* of the following would lead to a higher producer surplus?
 - A Demand being price elastic
 - B Demand being price inelastic
 - C Supply being price elastic
 - D Supply being price inelastic

(1 mark)

- 3 Price discrimination would be expected to:
 - A Increase consumer surplus and increase producer surplus
 - B Increase consumer surplus and decrease producer surplus
 - C Decrease consumer surplus and increase producer surplus
 - D Decrease consumer surplus and decrease producer surplus

(1 mark)

- 4 Define the term 'consumer surplus'.

(3 marks)

- 5 Define the term 'producer surplus'.

(3 marks)

- 6 Define the term 'deadweight welfare loss'.

(3 marks)

- 7 Draw a diagram to show consumer surplus and producer surplus in a perfectly competitive market.

(7 marks)

Based on Figure 21.3, answer Questions 8, 9 and 10.

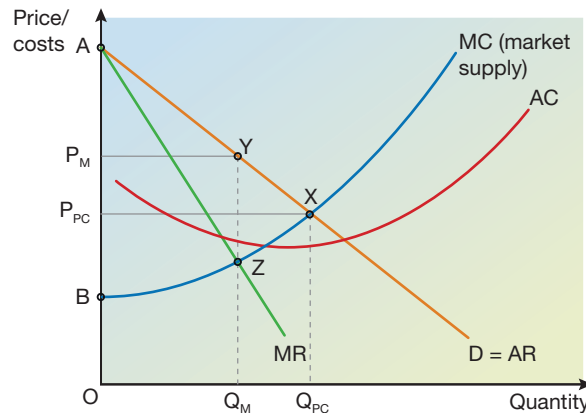


Figure 21.3 A profit-maximising monopoly

- 8 Which area shows the consumer surplus under monopoly?

(2 marks)

- 9 Which area shows the producer surplus under monopoly?

(2 marks)

- 10 Which area shows the deadweight welfare loss arising from monopoly?

(2 marks)

Topic 5 Exam-style questions

A-LEVEL PAPER 1

SECTION A Context – Competition in the banking industry

Extract A Market shares and profit

Bank	Market share in UK (%)	Profit 2014–15 (£ billion)
Lloyds	27	7.8
Barclays	18	5.5
RBS	18	3.5
HSBC	12	14.9
Santander	10	4.5
Others	15	N/A

Note: Profits relate to group profits, not just UK.

Extract B The Competition and Markets Authority (CMA) report on competition in the banking industry, October 2015

Although UK banks offer similar services, there has been no evidence to suggest that there is collusion in this market.

In its recent in-depth investigation into UK banks, the CMA concluded that banks do not have to work hard enough to compete for customers. The CMA identified a number of competition problems in this market:

- *low levels of customer switching* – customers are loath to change bank accounts because of concerns about possible complications arising, the potential for errors when switching and the time taken to switch accounts;
- *new products and new banks do not attract customers quickly enough* – customer reluctance to switch accounts and limited availability of information mean that accounts offering more favourable terms to customers are not recognised by customers;
- *limited information* – there is a lack of information available to allow customers to make an effective comparison between different bank accounts. This encourages inertia and therefore a lack of switching bank accounts. The CMA found many examples of customers who could save £260 a year from switching current accounts. It concluded that the average customer could save £70 a year by switching to other suppliers. In other industries, such cost savings would have a much greater impact on customers switching between suppliers;
- *customer inertia* – more than half (57%) of UK consumers have the same personal bank account as they had 10 years ago. Around 37% of consumers have been with the same bank for more than 20 years;
- *market structure* – the four largest banks in the UK account for approximately 75% of UK bank accounts. This makes it difficult for new competition to break into the market.

Source: CMA, October 2015

Extract C **CMA proposals to improve competition in the UK banking industry**

The CMA investigation found some positive developments and indications of dynamic efficiency. Technology was well used to enhance certain services and both new entrants and existing firms provided innovative products. However, a lack of information and an inability to compare accounts meant that fewer customers benefited from these changes than would normally be the case.

Potential remedies suggested by the CMA include:

- requiring banks to prompt customers to review the service they receive through individual messages at certain ‘trigger points’, such as when there is a change in interest rates;
- making it easier for consumers and businesses to compare bank products and recognise the costs incurred through their existing bank products – other bodies have suggested using opportunity cost to show how ‘free current account banking’ can lead to lost opportunities to earn interest from other savings accounts;
- requiring the creation of a new price comparison website, so that customers can compare the cost and quality of services offered by different banks;
- requiring banks to raise public awareness and confidence in switching bank accounts;
- requiring better sharing of information between financial organisations, particularly so that small businesses can shop around for loans without having to complete multiple application forms.

Source: CMA, October 2015 and other sources

Questions

Total: 40 marks

- 1 Using the information in Extract A, calculate the five-firm concentration ratio for the UK banking industry. (2 marks)
- 2 Explain how the data in Extract A might indicate the existence of barriers to entry into this market. (4 marks)
- 3 ‘Although UK banks offer similar services, there has been no evidence to suggest that there is collusion in this market.’ With the help of a diagram, explain how collusion between banks could affect the price for consumers and the level of profitability for banks. (9 marks)
- 4 To what extent is the low level of competition in the banking industry a result of poor and asymmetric information? (25 marks)

SECTION B Essays

Total: 40 marks

The efficiency of a market at a particular point in time can be measured in different ways, such as productive efficiency and allocative efficiency. However, the long-run improvements in efficiency are more likely to result from dynamic efficiency. In competitive markets, firms have little control over price. However, in concentrated markets, firms can more significantly influence the price of their goods and therefore the revenue they receive.

- 1 Explain why a firm’s average revenue and marginal revenue differ, according to whether it is in a competitive or concentrated market. (15 marks)
- 2 Critically evaluate the view that perfectly competitive markets lead to an efficient allocation of resources. (25 marks)



Topic 6

The labour market

The demand for labour & marginal productivity theory

Key concepts from Year 1

This chapter builds on Chapters 6 and 7 of the Year 1 companion textbook, which introduced the concepts of demand and elasticity of demand.

This chapter develops the first year topics of demand and elasticity of demand in the context of the labour market. The section on marginal productivity theory builds on the law of diminishing returns, which was explained in Chapter 5 of this book. We explain why the demand for a factor of production, such as labour, is derived from the demand for the product. We then examine how marginal productivity theory determines the demand for labour. The demand curve for labour and the relationship between the wage rate and number of workers employed is considered. The chapter concludes by studying the causes of shifts in the demand curve for labour and the determinants of the elasticity of demand for labour.

The demand for a factor is derived from the demand for the product

The demand for all factors of production is a derived demand. In the case of labour, the demand for baristas is derived from the demand for coffee from coffee shops; and the demand for car mechanics is derived from the demand for cars. Similarly, for capital, the demand for coffee machines is derived from the demand for coffee; and the demand for machinery needed to make cars is derived from the demand for cars. In the case of land, the demand for natural raw materials, such as coffee, is derived from the demand for coffee; and the demand for rubber, metal and other materials needed in car manufacturing is derived from the demand for cars. Finally, high demand for any good leads to more demand for entrepreneurs.

The marginal productivity theory of the demand for labour

The demand for labour is a derived demand, as labour is demanded because it produces goods and services. How is the demand for labour derived?

Figure 22.1 shows marginal physical productivity (MPP) of labour for a particular firm. The marginal physical productivity (MPP) reflects the law of diminishing returns (explained in Chapter 5). At low levels of output, MPP increases as output increases because fixed factors of production are used more effectively. However, after a certain level of output MPP falls and continues to fall as output increases.

In a perfect market, a firm sells all of its output at the market equilibrium price. This production's monetary value – its marginal revenue productivity (MRP) – is calculated by the formula: $MPP \times \text{price (p)} = MRP$. If the n^{th} worker adds 5 units of

Key term

Derived demand occurs when the demand for a good or service is determined by the demand for another good or service.

Key term

Marginal revenue productivity (MRP) is the addition to total revenue arising from the employment of one additional unit of labour.

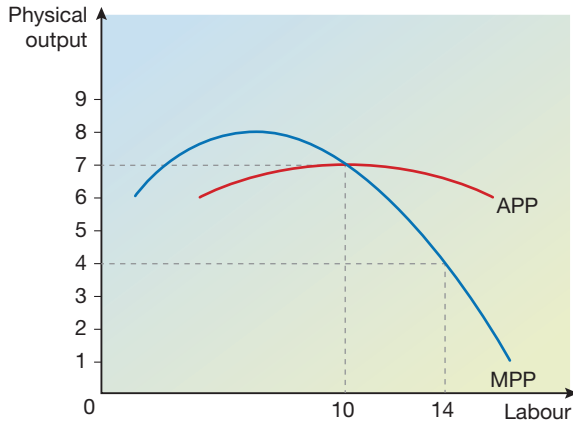


Figure 22.1 Marginal physical productivity

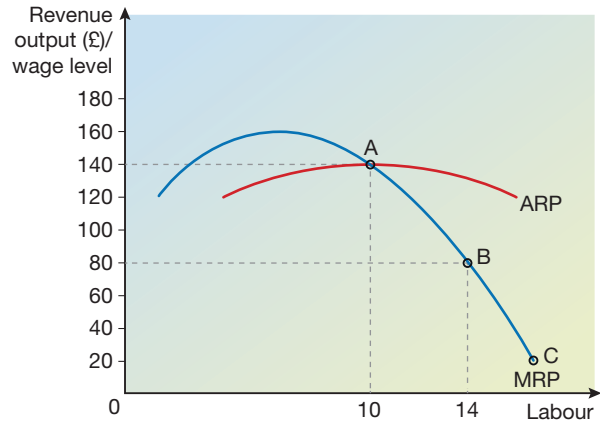


Figure 22.2 Marginal revenue productivity

output and the market equilibrium price is £16, then the MRP of the n^{th} worker is £80.

Author tip

Do not confuse MR and MRP. MR applies to the additional revenue from one extra unit of output. MRP applies to the additional revenue from employing one extra worker.

It is assumed that in this market the equilibrium price is £20. Consequently, each additional unit of output that labour produces generates an additional £20 of revenue. Thus the shape of MPP and MRP are identical. However, units for the 'y' axis change from output to revenue: one unit of MPP equals £20 of MRP. Figure 22.2 shows marginal revenue productivity (MRP) based on the MPP shown in Figure 22.1.

The demand curve for labour shows the relationship between the wage rate and number of workers employed

The marginal cost of employing one additional worker is the cost of employing the additional worker. This will be the wage paid to the worker plus additional payments such as employer's national insurance contributions. In Figure 22.1, the tenth worker's MPP is 7 units of output. Thus in Figure 22.2, Point A shows that the tenth worker brings in £140 ($7 \times £20$) of revenue. Therefore the firm could afford to pay the tenth worker up to £140 because any wage below £140 would lead to an increase in profit. From the perspective of demand for labour, if the price of labour (wages) were £140, the firm would demand 10 workers. Beyond Point A, MRP is falling, so the firm will only employ more than 10 workers if the wage level is lower than £140.

In Figure 22.1, the fourteenth worker has a marginal physical productivity (MPP) of 4 units of output. In Figure 22.2, Point B shows that the MRP of the fourteenth worker is £80 ($4 \times £20$). Thus 14 workers would be demanded if the price of labour (wages) were £80.

In effect, the line ABC in Figure 22.2 becomes the demand for labour (with the 'y' axis measuring 'wages' rather than revenue). No workers are demanded if wages exceed £140 because ARP is at its maximum at A (£140). Thus the firm will lose money if it employs workers at a wage level above £140.

Author tip

The analysis above applies to a perfectly competitive market, where all of a firm's output is sold for the same price.

In concentrated markets, the firm's price must fall if it is to sell more. For a concentrated market, marginal physical productivity (MPP) is the same as for perfect competition. However, in a concentrated market, additional units of output can only be sold if price is lowered. Thus, as output increases, MRP slopes downwards more steeply than MPP, because price (p) is constantly falling. Table 22.1 provides a brief extract to illustrate MRP in a monopoly.

Table 22.1 MPP and MRP in a monopoly

Units of labour	Output	MPP	Price per unit in a monopoly market	Total revenue product (TRP) in a monopoly market	Marginal revenue productivity (MRP) in a monopoly market
1	6	6	£13	£78	£78
2	15	9	£12	£180	£102
3	28	13	£11	£308	£128
4	42	14	£10	£420	£112
5	52	10	£9	£468	£48
6	58	6	£8	£464	-£4
7	60	2	£7	£420	-£44

In Table 22.1 it can be seen that for a monopoly market, MRP declines after the third unit of labour has been employed. Although the fourth unit of labour increases MPP, the lower price at which these additional units are sold leads to MRP declining. The final column also shows how quickly MRP declines in a monopoly, as more workers are employed. The sixth worker adds 6 units of output but the lower price means that MRP is negative (i.e. total revenue falls by £4).

Author tip

MPP and MRP are short-run concepts. They show how output and revenue change as variable factors of production (such as labour) change but fixed factors (such as capital) remain fixed. **MRP therefore shows the short-run demand for labour.** In the long run, greater revenue may be earned by changing the balance between different factors of production, say by making output more capital intensive by replacing labour with capital.

The causes of shifts in the demand curve for labour

In the short run the demand curve is determined by marginal revenue productivity (MRP), which is determined by marginal physical product (MPP) and price (p).

$$MRP = MPP \times p$$

In effect, this means that changes in the demand for labour in the short run can be caused by one of two factors – a change in MPP and/or a change in the price of the good that the workers are producing. If labour productivity increases, MPP will increase and so MRP will also increase. However, an increase in the price of

the good being produced by the workers also increases revenue productivity. Thus labour produces more revenue for a firm if it produces the same quantity of goods as before but those goods are sold at a higher price.

Table 22.2 shows how MRP (and thus the short-run demand curve for labour) increases because of (a) a 50% increase in MPP and (b) a 50% increase in price.

1	2	3	4	5	6	7	8
Units of labour	Q	MPP	p*	TRP	MRP	MRP after 50% increase in MPP	MRP after 50% increase in price
1	14	14	£10	£140	£140	$21 \times £10 = £210$	$14 \times £15 = £210$
2	36	22	£10	£360	£220	$33 \times £10 = £330$	$22 \times £15 = £330$
3	60	24	£10	£600	£240	$36 \times £10 = £360$	$24 \times £15 = £360$
4	84	24	£10	£840	£240	$36 \times £10 = £360$	$24 \times £15 = £360$
5	100	16	£10	£1000	£160	$24 \times £10 = £240$	$16 \times £15 = £240$
6	108	8	£10	£1080	£80	$12 \times 10 = £120$	$8 \times £15 = £120$
7	110	2	£10	£1100	£20	$3 \times £10 = £30$	$2 \times £15 = £30$

Table 22.2 Effects of 50% increases in MPP and price on MRP

Note: * This table applies to a perfectly competitive market, so price is constant.

Column 5 shows total revenue productivity (TRP). This is total output (Column 2) multiplied by price (Column 4). Column 6 shows marginal revenue productivity (MRP). This is MPP (Column 3) multiplied by price (Column 4).

If there is a 50% increase in labour productivity (Q), MPP and MRP will increase by 50%. Thus MRP for one unit of labour rises from £140 ($10 \times £14$) to £210 ($15 \times £14$). Column 7 shows the new MRP after a 50% increase in labour productivity.

Column 8 shows the impact of a 50% increase in price on MRP. The first unit of labour still has an MPP of 14 units, but these units are sold for £15 and so MRP is £210. Column 8 shows the new MRP after a 50% increase in price.

Conclusion: An increase in price has the same impact on the demand for labour as an increase in MPP.

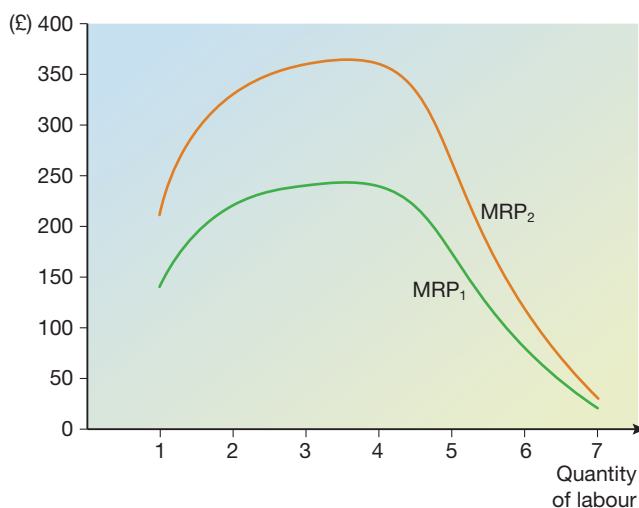


Figure 22.3 Factors changing marginal revenue productivity (MRP)

Figure 22.3 shows the original MRP line (MRP_1). The line MRP_2 shows marginal revenue productivity after the 50% increase in MPP or after the 50% increase in

price. It can be seen that marginal revenue productivity has increased. Therefore the demand curve for labour, which is that portion of the MRP line that is below the average revenue productivity (ARP) line, will have increased.

Table 22.2 and Figure 22.3 show factors leading to an increase in demand for labour. **A decrease in demand for labour** is caused by the opposite changes (i.e. a decrease in productivity, as measured by MPP, and/or a decrease in price).

When considering factors influencing the demand for labour it should be noted that MPP and price have many determinants, all of which therefore affect the demand for labour.

MPP will be affected by the efficiency of production. Thus factors such as technology, the level of capital goods, and the skills and training of the workforce will all affect productivity and thus the demand for labour. For example, improvements in training will increase productivity and hence increase the demand for labour.

Price is determined by demand and supply in a perfect market. Thus any factor that increases the demand for a particular good will increase the demand for labour that is derived from the demand for the good. Therefore, factors such as an increase in the price of substitutes, an increase in consumers' incomes or individual preferences

REALWORLD ECONOMICS 22.1

Wages in UK car manufacturing

Between 2012 and 2015, prices of cars manufactured in the UK rose by 15% in the case of mainstream manufacturers, such as Nissan, Mini, Honda and Toyota. For luxury models, such as Range Rover and Jaguar, price rises averaged 8% over the three years.

Measured in terms of added value per employee, the UK car industry showed significant increases in productivity. In 2012, the average employee in car manufacturing achieved added value of £66,500. By 2015, added value per employee had risen



UK car manufacturing – high wage costs but high productivity

to £70,400. This compares very favourably to the UK average added value in manufacturing, which is just over £35,000 per employee. UK labour productivity in motor manufacturing is the highest in Europe, although on a global scale the UK is the seventh most productive country. However, labour costs are high in the UK car industry, with wages being the fourth highest out of the 10 major European car-producing countries.

Car production in the UK has risen steadily over the last two decades. Although wage costs are high, productivity is high too. The UK car industry also scores highly in terms of labour flexibility. It is ranked highest in Europe and second highest in the world in its ability to cope with changes in the motor vehicle industry – this is a key measure of dynamic efficiency.

In 2015, wages in UK car manufacturing increased by an

average of 3% – well above the rate of inflation. However, between 2012 and 2014 wages stagnated and, in real terms, car workers suffered a slight decrease in wages.

Overall, employment has declined as more capital-intensive methods of production have been employed, although in the last year there has been growth in the demand for labour.

Source: Automotive Council UK, UK Automotive International Competitiveness Report 2015 and other sources

Exercises Total: 15 marks

- 1 Explain why the information in the opening two paragraphs would support the view that there has been an increase in demand for labour in the UK car industry. (6 marks)
- 2 Analyse possible reasons why actual wage rises in the UK car industry have been so low over the last few years. (9 marks)

changing in favour of the good will all lead to an increase in the demand for labour to make that good.

The determinants of the elasticity of demand for labour

In this section we will study factors that influence the elasticity of demand for labour and thus the shape of the demand curve. The main determinants are:

Substitute factors of production: If it is easy to substitute labour with another factor of production, such as capital or land, then elasticity of demand for labour will be more price (wage) elastic. If wages increase, firms will seek alternative methods of production in order to save costs. If it is easy to substitute another factor of production, then there will be a significant decrease in demand for labour as a result of a wage increase.

Time: In the short run, it can be difficult to modify methods of production and so a change in wages will only have a slight effect on the demand for labour. In the long run, firms can change methods of production, and so an increase in wages will lead to a more significant decrease in the demand for labour.

The price elasticity of demand for the good being produced: Other things being equal, the more price elastic the demand for the good, the more price elastic is the demand for labour. If a good has price elastic demand, then any increase in costs (and therefore the price charged) will have a significant impact on the amount of the good sold. A small wage increase might therefore lead to a major decrease in sales of the good and thus a major decrease in the demand for labour to make that good.

Labour costs as a percentage of total costs: If wages constitute a significant percentage of total costs, then an increase in wages can have a significant impact on the total costs of the good. Therefore, a firm will react to wage increases by cutting the use of labour, where possible. However, if wages only represent a small percentage of total costs, a firm is more likely to accept a wage increase without seeking to replace labour because the wage increase has very little impact on total costs.

Key note

The factors noted above apply to the elasticity of demand for labour in any market structure, but the market structure itself can influence the elasticity of demand for labour. In perfect competition, the price is constant and so the MRP is the same shape as the MPP. However, Table 22.1 shows how in monopoly (or imperfect competition) the MRP falls more steeply because as more workers are employed and output rises, the MRP declines more rapidly. Thus increases in wages will lead to a larger decrease in the employment of labour in monopoly than in a perfectly competitive market.

Review questions

Total: 35 marks

- 1 Other things being equal, the demand for labour by a breakfast cereal producer will decrease if:
 - A Breakfast cereals become more popular
 - B Income tax falls
 - C Labour productivity falls
 - D The price of a substitute for the breakfast cereal rises

(1 mark)

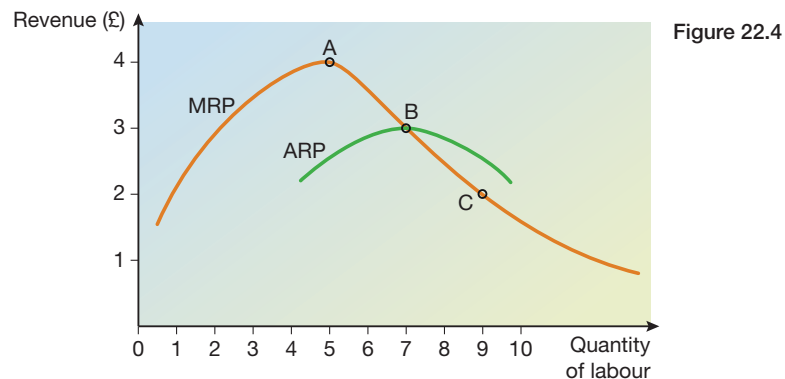
- 2 What is the formula for calculating marginal revenue productivity? (2 marks)
- 3 What is meant by the term 'derived demand'? (3 marks)

Question 4 is based on the Table 22.3, which shows data for the workforce of a manufacturer of kettles. Each kettle is sold for £25.

Number of workers	Output	MPP	MRP (£)
1	10		
2	25		
3	52		
4	72		

- 4 (a) Complete Column 3. (2 marks)
 (b) Complete Column 4. (2 marks)

Question 5 is based on Figure 22.4, which shows the MRP and ARP of labour in a match factory.



- 5 (a) What is the demand for labour at a wage rate of £3? (2 marks)
 (b) If nine workers are employed, what is the equilibrium wage? (2 marks)
 (c) Why is Point A *not* on the demand curve for labour? (4 marks)
- 6 In perfect competition, explain why the MRP increases and then decreases as more workers are employed? (6 marks)
- 7 Explain why MRP slopes downwards more steeply in monopoly than in perfect competition. (5 marks)
- 8 Explain two factors that influence the elasticity of demand for labour. (6 marks)

Influences upon the supply of labour to different markets

Key concepts from Year 1

This chapter builds on Chapter 8 of the Year 1 companion textbook, which introduced the concept of supply of goods and services.

This chapter develops the first year topic of supply in the context of the labour market. It explains how the supply of labour to a particular occupation is influenced by monetary considerations, mainly wages, and non-monetary considerations, such as job satisfaction and dissatisfaction and working conditions. The chapter examines how the supply curve for labour shows the relationship between the wage rate and number of workers willing to work in an occupation and concludes by analysing the causes of shifts in the market supply curve for labour.

Influences on the supply of labour to a particular occupation

Author tip

The supply of labour can be viewed in different ways, such as the supply to a single firm, the supply of labour to an industry or supply of labour to the economy as a whole. Supply of labour can also be applied to the factors that influence an individual when deciding on their occupation, firm, industry or economy. The AQA A-level economics specification states that students will not be required to understand the determinants of an individual's supply of labour.

The supply of labour to a particular occupation is derived from the individual workers' willingness and ability to undertake that particular work and macroeconomic factors that influence the overall supply of workers in the economy as a whole.

The supply of labour shows how many workers offer to work in a particular job at different wage rates. It is independent of market conditions as it is determined by the workers' perceptions of the suitability of the job. Workers' willingness to offer their services are influenced by:

- monetary considerations, such as wages and bonuses;
- non-monetary considerations, such as job satisfaction and working conditions.

Monetary considerations

If wages increase, the supply of workers extends as that job becomes more attractive in comparison to alternative occupations. Workers who are already employed are also likely to work more hours if the hourly wage rate is higher. Although monetary

Key term

The **supply of labour to a particular occupation** is the number of hours people are willing and able to supply to that particular occupation at any given wage rate.

considerations are primarily the wage rate, there are other monetary factors that influence the supply of labour:

- **Wage rate/salary:** The wage rate is normally expressed as a sum of money per hour worked. This is often used for manual work and part-time work. A salary is a fixed sum per annum. A high wage or salary for a particular occupation will attract more workers to offer their services. However, most occupations require certain skills, experience and qualifications and so a high wage may not always attract a large rise in employees offering to work.
- **Bonuses:** Many jobs provide a basic salary or wage but with added opportunities to earn further money. These bonuses can take many forms, such as profit share, commission and productivity bonuses.
- **Opportunities for overtime:** Overtime is an additional payment for working more hours than those stated in the employee's contract. For workers on part-time or temporary contracts, the opportunity for additional hours of work may be an attraction to some employees, as it can be used flexibly to suit their lifestyles.
- **Wage rates in alternative occupations:** A rise in the monetary value of each of the previous factors will lead to a rise in supply of labour. However, if alternative occupations have higher wages, this will lead to a decrease in the supply of workers in the original occupation. This factor is particularly important for unskilled jobs, where workers can more easily switch between different occupations.

Non-monetary considerations

Workers' willingness to take a job or work longer hours depends on factors such as job satisfaction and working conditions. Furthermore, the ability to take a job depends on the education and skills of the workforce. Some of the main examples of non-monetary considerations are explained below.

- **Job satisfaction and dissatisfaction:** This factor is, arguably, the most important non-monetary consideration when selecting a job. Many theorists argue that workers are more motivated in jobs that they find satisfying than those which pay well. In selecting an occupation, most workers will try to find a job that suits their particular characteristics. If this aim is achieved, many workers should find that they gain job satisfaction and may be reluctant to change occupation, even if the pay is less satisfactory than they would wish. However, job dissatisfaction has the opposite effect. This dissatisfaction may arise from the job being different than expected, or external factors such as relationships with a manager, relationships with colleagues and the degree of independence offered by the job.
- **Working conditions:** Working conditions may dissuade potential employees. Examples include coal mining and jobs on North Sea oil rigs. In contrast, many jobs have very pleasant working conditions that will encourage more applicants. Other things being equal, good working conditions mean that workers are prepared to offer their services at lower wage rates; poor working conditions mean that workers must be offered higher pay to attract sufficient numbers.
- **Barriers to entry:** Some occupations, such as professions, have barriers to entry that limit the number of potential applicants. For example, lawyers must achieve certain qualifications and gain certain experience before they can become solicitors.
- **The length of training required:** Jobs that require high levels of training, such as dentists, are likely to have a lower supply of labour because few people will want

to complete the training or be able to complete it successfully.

- **Immobility of labour:** This can take two forms – occupational immobility and geographical immobility. Occupational immobility occurs because most jobs require certain skills. If relatively few people possess the skills required for a job, then there will be occupational immobility and so the supply of labour will be low. Geographical immobility occurs when workers are reluctant to move between different geographical areas. The high price of property in the Southeast restricts geographical mobility and so workers with relevant skills may be reluctant to apply because of difficulties in obtaining property. They may prefer to stay in a lower paid job in an area where living costs are lower.
- **The opportunity cost of work:** Working means sacrificing leisure time. Some individuals may value their leisure time highly and will want to restrict the number of hours they work. For families, the cost of childcare and a reluctance to limit contact with children are likely to mean that parents may work fewer hours, often with only one parent working full-time. Welfare benefits made by government can also influence people: high welfare payments will discourage people from working; low welfare payments will encourage people to work more hours.
- **Non-monetary characteristics of the job:** The factors below may also affect the supply of labour:
 - the number of hours of work and whether these hours are anti-social, such as night shifts;
 - any risk involved in the occupation (this risk can be physical danger);
 - job security (is there a high risk of losing the job?);
 - opportunities for promotion;
 - training provided, and the opportunity to develop new skills;
 - benefits provided by the employer, such as health insurance and gym membership.
- **The working population as a whole:** The supply of labour in the UK has risen in recent years because of increases in population. This may arise from a higher birth rate, a lower death rate or net immigration. The opposite of these factors will cause a decrease in labour supply. Demographics can also be an important influence on labour supply, notably the age distribution of population. Government policy, such as retirement age and school leaving age, will impact upon the potential supply of labour as well. Finally, social factors can influence labour supply. In recent decades, the UK working population has shown a significant increase in female employment through social changes.
- **Trade unions:** Trade unions can affect the supply of labour and the ability of workers to gain pay increases. This factor will be dealt with in Chapter 26.

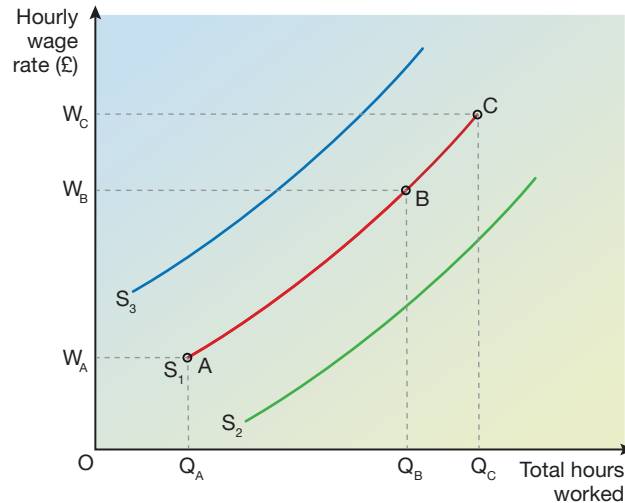
The supply curve for labour shows the relationship between the wage rate and number of workers willing to work in an occupation

The supply of labour shows the amount of labour offered for sale at any given price. The price of labour is the wage rate.

In Figure 23.1, the price of labour is ‘Hourly wage rate (£)’ and the quantity of work supplied is measured by the total number of hours worked. Line S_1 shows the initial supply of labour in this particular occupation. The impact of monetary factors

related to this occupation is shown by a movement along the supply curve. Thus an increase in the wage rate, salary, opportunities for overtime or bonuses will all lead to a higher supply of labour. The supply curve S_1 shows that higher wages will extend supply along the supply curve. If the original wage rate was W_B , Point B shows that the quantity of labour supplied will be Q_B . An increase in wages from W_B to W_C will lead to Point C, and so Q_C will be the quantity supplied. Similarly, if the wage rate falls from W_B to W_A , Point A shows that Q_A units of labour will be supplied.

Figure 23.1 Supply of labour



REALWORLD ECONOMICS 23.1

Training and wage levels

One of the key influences on labour supply for a particular occupation is training. The need for training limits supply and thus increases wages. However, a long period of training should also impact on the value of the good or service being provided. Therefore, the marginal



The work of a surgeon requires extensive training, which limits supply and raises wage levels

revenue productivity (MRP) of highly trained labour should exceed that of workers with lower skill levels. This factor will also lead to higher wage levels.

In the UK, the average annual wage in 2014 was £26,500. The three occupations with the lowest annual salaries were:

- lollipop ladies (£3187);
- theme park attendants (£6011);
- bar staff (£7317);

All of these occupations require limited training, although the salaries can be affected by the number of hours worked. Lollipop ladies are very unlikely to be full-time.

Most of the individuals that are very highly paid are self-employed.

However, amongst the highest paid occupations were:

- airline pilots (£78,000) – training of 3–5 years, although often self-financed;
- surgeons (£67,000) – training of 10–14 years;
- lawyers (£65,000) – training of 6 years.

The only high-paid occupation that does not need formal training is that of a Member of Parliament. In 2014, MPs were paid £67,000 a year and in 2015 their salaries increased to £74,000.

Discussion point

To what extent is training the most important influence on wage levels?

It should be noted that only monetary considerations related to that particular occupation will cause a movement along the supply curve. The final monetary consideration – ‘wage rates in alternative occupations’ – will make the original occupation more/less attractive at any wage rate, and so it will shift the supply curve for labour.

The causes of shifts in the market supply curve for labour

As with any supply curve, the shift in market supply will occur because of a change in the willingness and ability to supply at any given price (wage). Thus, it is the non-monetary considerations that cause changes to the market supply of labour because changes in wages show movement along the supply curve.

Non-monetary considerations affect the overall attractiveness of a particular occupation. If these non-monetary considerations improve the willingness and ability of workers to choose this occupation, there will be an increase in supply of labour. This is shown by the shift from S_1 to S_2 . However, if non-monetary considerations reduce the willingness and ability of workers to choose this occupation, there will be a decrease in supply of labour. This is shown by the shift from S_1 to S_3 .

Table 23.1 summarises some of the main non-monetary considerations, alongside wage rates in alternative occupations, to show their impact on the supply curve of labour.

Factor causing a shift in supply	Cause of an increase in supply	Cause of a decrease in supply
Job satisfaction/dissatisfaction	Improved satisfaction	More dissatisfaction
Working conditions	Better conditions	Poorer conditions
Barriers to entry	Removal of barriers	More barriers
Length of training required	Shorter training period	Longer training period
Mobility/immobility of labour	More mobility of labour	More immobility of labour
Opportunity cost of work	Fall in the opportunity cost of work	Rise in the opportunity cost of work
Non-monetary characteristics of job	Improvement in characteristics	Worsening of characteristics
Working population	Rise in working population	Fall in working population
Wage rates in alternative occupations	Lower wage rates	Higher wage rates

Table 23.1 Factors causing a shift in the market supply curve for labour

When deciding on how many hours of work to offer at a particular wage rate, an individual worker will consider the advantages and disadvantages of work. The main advantage is the economic welfare that the individual obtains from the goods and services that they can purchase with their wages. Non-monetary factors will also be considered – the main disadvantage is the time consumed by work, and the potential lost opportunities from not having this time available for other activities, such as leisure pursuits. Non-monetary factors can provide advantages, such as social aspects of employment, good working conditions and feelings of achievement, and disadvantages, such as boring work, unpleasant working conditions and anti-social working hours. Overall, these non-monetary factors can create job satisfaction or job dissatisfaction. A worker primarily experiencing job satisfaction will be prepared to work more hours, while a worker mainly experiencing job dissatisfaction will

wish to work fewer hours. The monetary rewards plus the level of job satisfaction/dissatisfaction together constitute the **net advantage** of the job – this is the value that will determine the number of hours that a worker will wish to work.

For the economy as a whole, labour supply will be higher if wages are higher and jobs offer satisfaction rather than dissatisfaction.

Review questions

Total: 20 marks

- 1 The supply of labour to a particular occupation will *increase* if:
 - A There are barriers to entry to an occupation
 - B There is greater immobility of labour
 - C There is net emigration
 - D Working conditions improve

(1 mark)
- 2 The supply of labour to a particular occupation will *decrease* if:
 - A Job security becomes higher
 - B Opportunities for training improve
 - C Leisure time becomes more highly valued
 - D Wage rates in alternative occupations decrease

(1 mark)
- 3 Which *one* of the following factors will cause a movement along the supply curve for labour?
 - A A decrease in wages
 - B Anti-social working hours
 - C Better working conditions
 - D Improvements to the occupation's pension scheme

(1 mark)
- 4 Which *two* of the four factors listed below explain why the supply of labour for a particular occupation will rise as wages rise?
 - i Existing workers will extend their hours of work
 - ii Firms will want to employ more workers
 - iii More workers will want to transfer to that occupation
 - iv Workers can afford to take more leisure time

(2 marks)
- 5 Define the term 'supply of labour to a particular occupation'.

(3 marks)
- 6 Explain *one* monetary factor that might influence the supply of labour to a particular occupation.

(4 marks)
- 7 Explain *two* non-monetary factors that might influence the supply of labour to a particular occupation.

(8 marks)

The determination of relative wage rates & levels of employment in perfectly competitive labour markets

Key concepts from Year 1

No prior knowledge from Year 1 is required.

This chapter builds on Chapters 22 and 23 in this book, by combining the demand for labour and the supply of labour to construct the economists' model of wage determination in a perfectly competitive labour market. It also examines the role of market forces in determining relative wage rates.

The economists' model of wage determination in a perfectly competitive labour market

As with any good or service, equilibrium in a perfectly competitive labour market occurs where demand equals supply. Figure 24.1 shows wage determination in such a market.

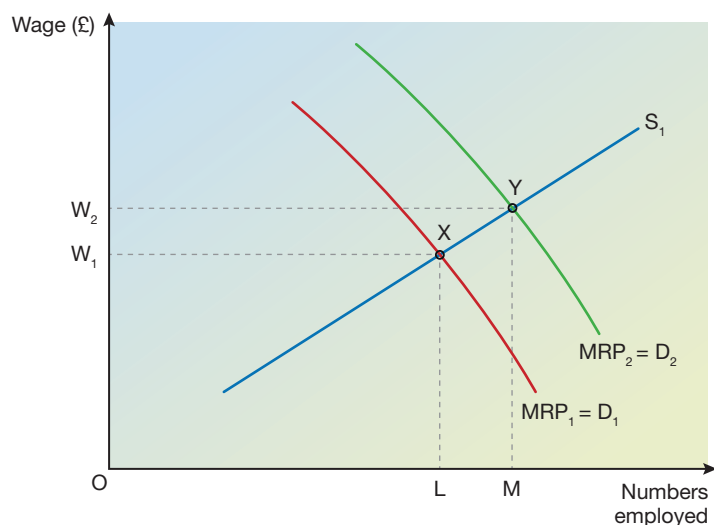


Figure 24.1 Wage determination in a perfectly competitive labour market

The original demand curve (D_1) is determined by the marginal revenue productivity (MRP) of labour in that occupation. The original supply curve (S_1) is influenced by the factors that affect the willingness and ability of workers to supply their labour.

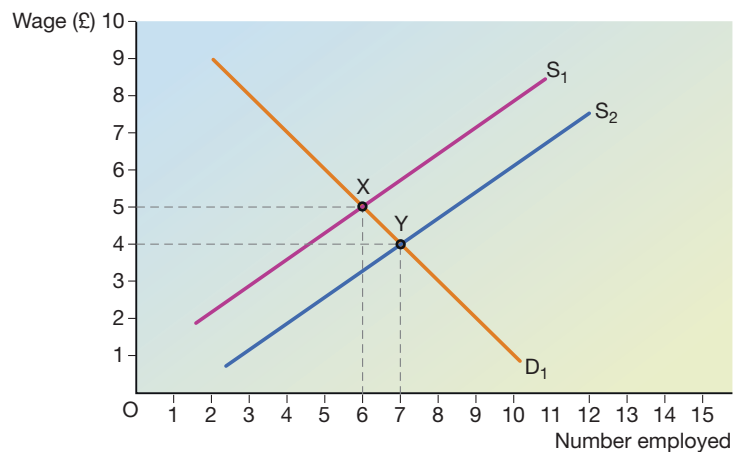
The market equilibrium occurs where the demand for labour (D_1) and the supply of labour (S_1) meet. This is at Point X. The initial equilibrium wage is therefore W_1 and OL shows the number of workers employed.

As with any market equilibrium, this is not a static situation. A change in any factor that can influence the demand and/or supply of labour will lead to a change in the equilibrium wage and the equilibrium level of employment.

In Figure 24.1, D_2 shows an increase in the demand for labour. (This could be caused by increased labour productivity and/or a higher price for the good being produced.) Other things being equal, this change in demand will lead to a new equilibrium at Point Y. Thus there will be an increase in the equilibrium wage from W_1 to W_2 . There will also be an increase in the number of workers employed, from OL to OM.

If there had been a decrease in labour productivity and/or a decrease in price, this would have caused a decrease in demand. A decrease in demand would lead to lower wages and a lower level of employment. A change in equilibrium from Point Y to Point X could be used to show a decrease in demand, assuming that the original equilibrium was at Point Y.

Figure 24.2 The impact of a change in supply of labour on equilibrium wage and equilibrium level of employment



In Figure 24.2 the original demand for labour is shown by D_1 and the original supply of labour is shown by S_1 . Point X shows the original equilibrium point. The equilibrium wage is thus £5 and the equilibrium number employed is 6 workers.

The line S_2 shows an increase in the supply of labour; this would have arisen because of a change in conditions that led to more workers wishing to offer their labour in this particular market. As a result of this increase in supply, the equilibrium point changes from X to Y. This leads to a decrease in the equilibrium wage from £5 to £4 and an increase in the number employed from 6 workers to 7 workers.

Figures 24.1 and 24.2 show a perfectly competitive market as a whole. How is the equilibrium for an individual firm in a perfectly competitive market determined? In such a market, an individual firm has no influence on market wage but must pay it to attract workers. Since there is perfect information, workers will know that they will only be able to sell their services at the equilibrium wage and firms will know that they will need to pay the equilibrium wage in order to attract workers. Thus the equilibrium for an individual firm is shown in Figure 24.3.

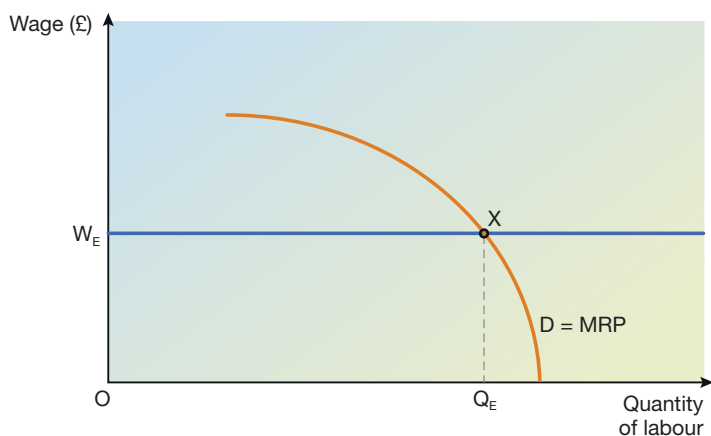


Figure 24.3 The individual firm's equilibrium position in a perfectly competitive market

In Figure 24.3, the demand for labour is shown by the line $D = MRP$. The equilibrium wage in the market as a whole is W_E . All workers will offer their services at the equilibrium wage and so, for the individual firm, this is effectively the supply of labour. Supply and demand meet at Point X and therefore the quantity of labour employed is shown by Q_E . The wage paid by the firm is the market equilibrium wage of W_E because individual firms in a perfectly competitive market are price takers. At Point X, the marginal revenue received from an employee is MRP and the marginal cost is the additional wage paid to that employee. As these two are equal, the firm maximises profit by employing Q_E workers at the equilibrium wage of W_E .

Author tip

You should recognise that a perfectly competitive labour market is unlikely to exist in real life. This model is studied to provide a yardstick when judging the extent to which real labour markets perform efficiently or inefficiently and whether they lead to a misallocation of resources. Where there are many individuals supplying their labour and many small firms demanding labour, the equilibrium number of employees and the equilibrium wage rate will approximate to the perfectly competitive market. The further a market is from this equilibrium, the greater the level of imperfection in that market.

Role of market forces in determining relative wage rates

The characteristics of a perfectly competitive labour market are the same as the characteristics of a perfectly competitive market for any other good or service. Thus in a perfectly competitive labour market the following assumptions are made:

- **Homogeneous goods or services:** For a labour market this would mean that all workers have the same skills and capabilities.
- **Perfect knowledge:** All suppliers of labour (workers) and all buyers of labour (employers) are fully aware of jobs in the economy and their wages.
- **Buyers and sellers are price takers:** The equilibrium wage in the market is determined by the supply and demand of labour.
- **Perfect mobility:** Labour can move freely between jobs and geographical areas.
- **Maximisation:** Workers aim to maximise their wages, and employers aim to maximise their profits by minimising their costs.

If all of these assumptions apply to the UK labour market, then all workers will be paid the same wage. If wages in one occupation were higher, more workers would move into that occupation because they would possess the same skills as the existing labour force. This increase in supply would force down wages and so wages in that occupation would reach the same level as other occupations. The same logic applies to regional wages. If wages are higher in a certain area of the UK, more workers will move to that area (because labour is perfectly mobile) until the wage level equates to wage levels in other areas.

It is evident that the UK labour market is *not* perfectly competitive in this way. In the UK, wages vary considerably between occupations and between geographical areas. Wage levels vary because of differences in demand factors or supply factors. Of particular importance is the variance in skills and capabilities of different people. Immobility of labour is also a major factor that prevents an equalisation of wages. Market forces thus determine relative wage rates in the UK. The key factors determining relative wage rates are outlined below:

- **Differences in MRP:** This is the key influence on the demand for labour. Increases in MRP will lead to an increase in the demand for labour and therefore higher wages. This is a key reason why workers in the same occupation receive different wages. A firm that uses workers efficiently to generate output and profit can afford to pay higher wages, whereas a worker doing the same job in a less efficient firm will receive lower wages.
- **Education, skills and training:** These factors also influence MRP because workers with higher skills should be able to produce more output. However, jobs that require more education, skills and training will often find a restricted number of applicants for jobs, particularly where there is a long period of training needed, such as surgeons and lawyers. Thus, education, skills and training also cause a decrease in supply of labour and therefore higher wages.
- **Natural ability:** For some occupations, natural skills can impact upon employees' effectiveness and therefore their wage level. Occupations such as professional footballers, artists and actors all reward people with natural ability because of the revenue they can generate.
- **Gender:** There are significant discrepancies in wages between males and females in the UK. This can arise from a greater tendency for women to take time out of the labour market, but also from prejudice. At one time, this discrepancy partially arose from a tendency for females to leave education at a younger age, but in recent times female graduates have exceeded male graduates. In the long run, this factor and social change may equalise some of the current inequalities of income between genders.
- **Age:** Younger people tend to have more energy but older people have more experience. These factors can influence the attractiveness of a worker to different employers and occupations.
- **Ethnic background:** Prejudice again prevents the free working of labour markets and can lead to lower levels of employment and income as a result of ethnic factors.
- **Job satisfaction and dissatisfaction:** If a job provides overall job satisfaction, the worker may be prepared to accept a lower wage, whereas job dissatisfaction will usually mean a higher wage to compensate for the negative aspects of the job. Thus jobs such as coal mining and working on oil rigs tend to have higher pay than those requiring comparable skills but with pleasanter working conditions.

It can be seen that perfect competition in labour markets is a theoretical rather than a practical model. All real-world labour markets are imperfectly competitive to a greater or lesser extent. Wage determination in imperfectly competitive markets will be considered in the next chapter.

REALWORLD ECONOMICS 24.1

Wage differences within occupations

It can be seen that wage differences occur between different occupations for a variety of reasons. What is less clear is why wages within a particular occupation can vary considerably.

Chartered accountants are specialist accountants whose education and training beyond A-levels can take between five and seven years; most successful chartered accountants will have a degree plus specialist training in order to gain the Association of Chartered Accountants (ACA) qualification.

Numeracy is a key skill, but

analytical business skills are more important. The average starting salary for graduates is £30,000 a year and career progression can be quick, with ACA students doubling their salary during training. Salaries can depend on location. The high incidence of financial services in London and the Southeast means that many chartered accountants work in that geographical area, with 83% of FTSE companies having a chartered accountant on their board of directors.

The average global salary of a chartered accountant is £111,300 (£90,800 + £20,500 bonus). In

the UK the average salary of a chartered accountant is £102,000 (£84,500 + £17,500 bonus). However, salaries vary from an average of £170,000 in banking to £68,500 for chartered accountants employed by government.

Source: Salary Survey Report 2014, Stott and May

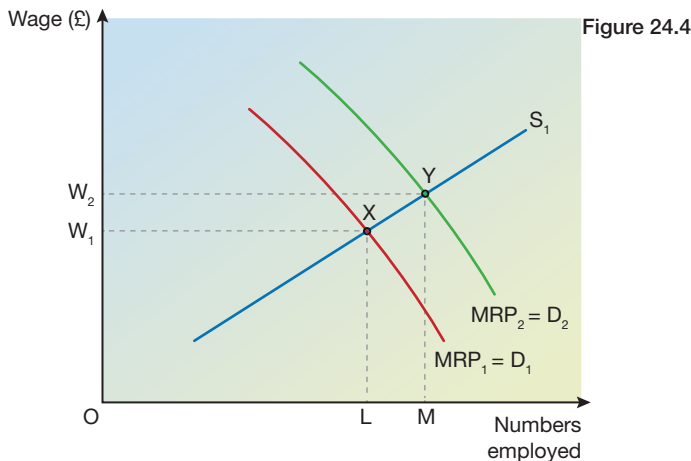
Exercises *Total: 15 marks*

- 1 Explain two reasons why chartered accountants have high salaries. *(9 marks)*
- 2 Explain why salaries of chartered accountants vary between different business sectors. *(6 marks)*

Review questions

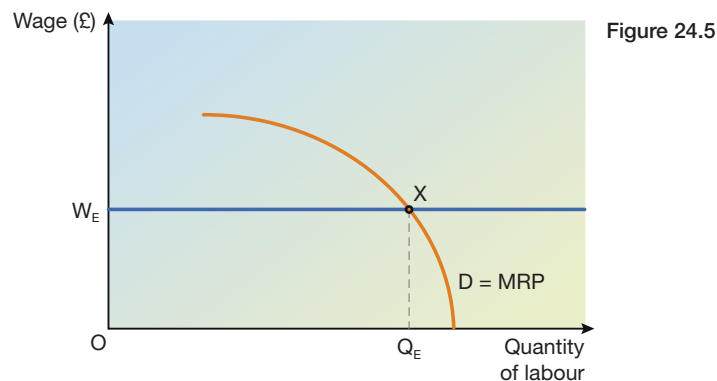
Total: 20 marks

Use Figure 24.4 to answer Questions 1–3.



- 1 In Figure 24.4, Point X is the original equilibrium. At this equilibrium, what is the total level of income received by employees in this market?
 A $OW_1 \times OL$
 B $OW_1 \times OM$
 C $OW_2 \times OL$
 D $OW_2 \times OM$ (1 mark)
- 2 In Figure 24.4, the supply of labour is price elastic. This is most likely to be the result of:
 A A long period of training required by workers
 B Employers demanding a specific qualification for employment
 C The job being unskilled
 D Very unpleasant working conditions (1 mark)
- 3 In Figure 24.4, the change in equilibrium from X to Y may have been caused by:
 A A worsening in working conditions
 B An increase in labour productivity
 C Barriers to entry into the occupation
 D More occupational immobility of labour (1 mark)

Use Figure 24.5 to answer Question 4.



- 4 In Figure 24.5, Point X shows the equilibrium point for:
 A The individual worker in a perfectly competitive market
 B The individual firm in a perfectly competitive market
 C The market as a whole in a perfectly competitive market
 D The market as a whole in all market structures (1 mark)
- 5 Analyse why the assumptions of a perfectly competitive market lead to the conclusion that all workers will receive equal wages. (8 marks)
- 6 Explain two reasons why a software designer might receive a higher wage than an agricultural worker. (8 marks)

The determination of relative wage rates & levels of employment in imperfectly competitive labour markets

Key concepts from Year 1

An understanding of imperfect information is required. This was covered in Chapter 27 of the Year 1 companion textbook and has also been dealt with in Chapter 3 of this book.

After our examination of perfectly competitive labour markets, this chapter studies the determination of relative wage rates and levels of employment in imperfectly competitive labour markets. It considers how factors such as monopsony power, trade unions and imperfect information contribute to imperfections in a labour market and how, in a monopsony labour market, the employer can use market power to reduce both the relative wage rate and the level of employment below those that would exist in a perfectly competitive labour market.

How various factors contribute to imperfections in a labour market

Imperfect labour markets take different forms. They may occur because of:

- monopsony power amongst employers;
- imperfect information;
- monopoly power exerted by trade unions;
- bilateral monopoly (where monopsony and monopoly power combine).

Each factor is discussed below (and the last two will be analysed in more detail in Chapter 26).

Monopsony power amongst employers

A monopsonist means a single buyer. Monopsony tends to occur when there is a monopolist in a market who is the only user of a specialist product or service. For example, in the UK the Royal Mail would be the only organisation that buys post boxes – a firm manufacturing post boxes will only have one buyer for its product.

In a labour market, monopsony means a single employer, such as the Royal Mail being the only employer of postmen. However, if there are a few employers in a market, each employer has some monopsony power. Many large oligopolists therefore have some monopsony power when buying materials and employing workers.

Key term

Monopsony is a market structure with a single buyer.

Imperfect information

Imperfect information arises because employees are unaware of potential jobs and wage rates. As a result, when seeking employment, employees may make decisions that are not rational because they lack the necessary information.

Employers often have more data, particularly as employment databases become more sophisticated. However, it is unlikely that they have perfect information because the supply and demand for jobs will be changing daily.

Time is also a constraint on decision makers in the labour market. A worker who is currently employed is likely to find it difficult to devote sufficient time to researching alternative occupations. A worker who is unemployed will invariably seek to rectify that situation quickly and is thus likely to accept employment even where conditions are unfavourable. Similarly, firms can only devote a certain amount of time to researching and analysing the labour market.

These factors contribute to market failure and so imperfect information can lead to imperfections in the allocation of labour.

Monopoly power exerted by trade unions

Trade unions try to encourage as many workers to join as possible. If all workers in a particular trade or profession, or all of the employees in a particular firm, belong to a specific trade union, then that trade union effectively becomes the only seller of labour in that market. As a sole supplier, it has monopoly power. Relatively few trade unions are sole suppliers of labour.

Where trade union membership is high, the trade union has some monopoly power and can use this to influence wages and the number of workers employed.

Bilateral monopoly (where monopsony and monopoly power combine)

Labour markets have the potential for bilateral monopoly. If an occupation only has one potential employer, it is possible that workers in that occupation will belong to the same trade union. For example, postmen in the UK tend to be members of the Communication Workers Union (CWU), and so the CWU negotiates with Royal Mail on issues related to its workforce, such as wage levels.

Chapter 26 examines the impact of trade unions on wages and the level of employment. The impact of trade unions and bilateral monopoly will therefore be considered in that chapter.

How a monopsonist can use market power to reduce the relative wage rate and level of employment

In a perfectly competitive labour market, an individual firm will need to pay the equilibrium market wage. If this is £24, then the average cost of each employee is £24. Each time an additional worker is employed, this adds £24 to the total labour costs and so the marginal cost of labour is also £24. This logic does not apply to a monopsony buyer of labour.

In a monopoly labour market, the firm/employer can use its market power to reduce wages and the level of employment. However, a monopsonist faces a dilemma. In order to attract additional employees, higher wages will need to be paid. As there will be the same wage for each employee in a particular job, this means that employing

Key terms

A **trade union** is an organised association of workers in a particular trade, profession or industry. It is formed to protect and further the rights and interests of its members.

Bilateral monopoly exists when there is a single buyer and a single seller of a good or service.

additional workers will increase marginal cost quite significantly. This situation is illustrated by the data in Table 25.1.

1	2	3	4	5
Quantity of labour	Wage rate (£)	Average cost (£)	Total cost of labour (£)	Marginal cost (£)
1	20	20	20	20
2	21	21	42	22
3	22	22	66	24
4	23	23	92	26
5	24	24	120	28
6	25	25	150	30
7	26	26	182	32

Table 25.1 Supply of labour to a monopsony market

Columns 1 and 2 show the market supply of labour. When one worker is employed, the wage rate will be £20, as this is the wage needed to attract one unit of labour. This means that the marginal cost and the average cost of the first unit of labour are both £20. However, if a second worker is required, the wage must rise to £21. This will add £21 to the costs of labour plus an additional £1 because the existing worker must also be paid £21. Thus, the marginal cost of the second unit of labour is £22. If a third worker is required, the wage must rise to £22. This will add £22 to the costs of labour plus an additional £2 (£1 to each of the two existing workers, who must also be paid £22). Thus, the marginal cost of the third unit of labour is £24.

Table 25.1 shows the marginal cost and average cost of labour to a monopsony buyer of labour. It can be seen that the marginal cost of employing labour rises twice as steeply as the average cost. Figure 25.1 shows the marginal cost and average cost of labour to the monopsony buyer, based on the data in Table 25.1.

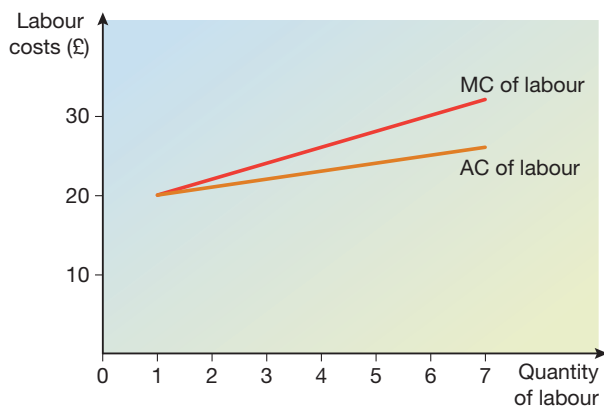


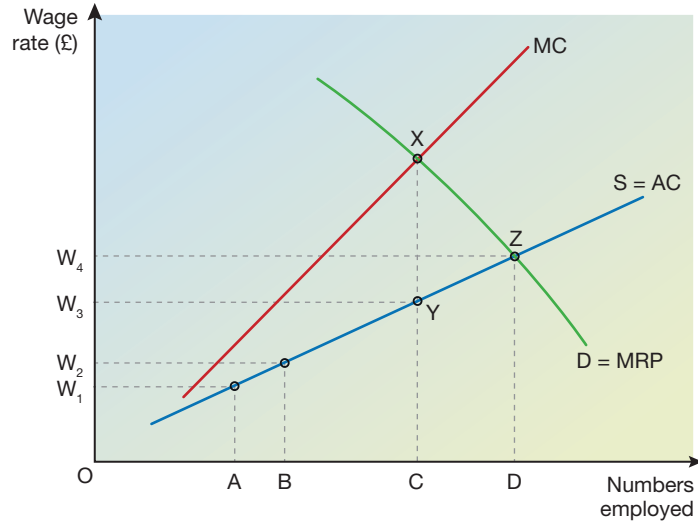
Figure 25.1 MC and AC of labour to a monopsony buyer

In a perfectly competitive market the equilibrium wage is paid to all employees. Thus the AC line shows the supply of labour in a perfectly competitive market.

In a monopsony labour market, the monopsonist will decide on the level of employment. This occurs where the marginal revenue received from employing labour (the MRP) is equal to the marginal cost (MC) of employing labour.

Figure 25.2 shows a monopsony labour market. Demand for labour is MRP, as in a perfectly competitive labour market. However, supply differs. $S = AC$ shows the

Figure 25.2 *Equilibrium wage and level of employment in a monopsony market*



normal market supply for labour. If the firm wants OA workers, it pays wage level of OW_1 . As every worker gets OW_1 , this is the AC of labour. If it needs additional workers (say, OB), then it needs to pay OW_2 . However, it must pay OW_2 to all workers, and so workers previously on OW_1 get a wage rise to OW_2 . Thus additional cost (MC) of labour rises much more steeply than AC .

REALWORLD ECONOMICS 25.1

Nurses wages, 2015

In the UK, 96.9% of health expenditure is government funded. This is the highest level in Europe. However, an OECD survey of European countries revealed that the government accounts for more than 50% of health expenditure in all but three countries out of 45 surveyed.

This gives the government in most countries considerable monopsony power when buying health service workers, such as nurses. From a sample of 19 countries, Table 25.2 shows the ranking of the countries with the highest and lowest percentage of health expenditure that is government funded and the ranking of average wages paid to nurses in each country, relative to the average wage of that country.

Country	Percentage of health spending by government (rank)	Ranking of nurses wages relative to average wage
UK	1 st	10 th
Czech Republic	2 nd	7 th
Luxembourg	3 rd	1 st
Netherlands	17 th	13 th
Greece	18 th	3 rd
Italy	19 th	9 th

Source: OECD database, 2015–16

Table 25.2 *Data on government spending on health and nurses' relative wages*

Conclusion: *The data show very little correlation between the two sets of figures.*

Exercises

Total: 20 marks

1 If governments were to use their monopsony power, what link would you expect to see between these two sets of data?

(5 marks)

2 Which country's data are the most inconsistent with your conclusion in Question 1, and why?

(3 marks)

3 Explain three possible reasons why there appears to be no close link between government monopsony power and nurses' wages.

(12 marks)

The quantity of labour that the monopolist will wish to employ occurs where MR of labour (MRP) equals MC. This is at Point X. Thus OC workers are employed. What wage rate will the monopsonist pay? Point Y shows that OC workers will offer to supply their labour if the wage rate is OW_3 . Thus the monopsonist employs OC workers at a wage rate of OW_3 .

If this had been a perfectly competitive labour market, then the market equilibrium would be where the demand for labour equals the supply of labour. This is shown by Point Z, where $S = D$. In a perfectly competitive market, OD workers would be employed at a wage level of OW_4 . Under monopsony, both wages and employment levels are lower than in a perfectly competitive labour market.

Review questions

Total: 30 marks

- 1 Bilateral monopoly occurs in a market when:
 - A There are only two buyers
 - B There are only two sellers
 - C There is only one buyer and one seller
 - D There are two buyers and two sellers

(1 mark)
- 2 Define the term 'trade union'. (3 marks)
- 3 What is meant by the term 'monopsony'? (3 marks)

Use Table 25.3 to answer Questions 4, 5 and 6.

1	2	3	4
Quantity of labour	Total labour costs (£)	Average labour costs (£)	Marginal labour costs (£)
1	42		
2			48
3		49	

Table 25.3 Data on a monopsonist's labour costs

- 4 Complete the missing figures in Column 2 (total labour costs). (2 marks)
- 5 Complete the missing figures in Column 3 (average labour costs). (2 marks)
- 6 Complete the missing figures in Column 4 (marginal labour costs). (2 marks)
- 7 Explain one way in which imperfect information might reduce the equilibrium wage in a labour market. (5 marks)
- 8 For a monopsony, marginal cost of labour increases more quickly than average cost, as more workers are employed. Explain why this is the case. (4 marks)
- 9 Draw a single diagram to show:
 - (a) the equilibrium wage and level of employment in monopsony; (5 marks)
 - (b) the equilibrium wage and level of employment that would have occurred if this market had been perfectly competitive. (3 marks)

The influence of trade unions in determining wages & levels of employment

Key concepts from Year 1

No prior knowledge from Year 1 is needed for this chapter.

This chapter examines trade unions and the various factors that affect the ability of trade unions to influence wages and levels of employment in different labour markets. Diagrammatic analysis is undertaken to show how wages and employment are likely to be affected by the introduction of a trade union into a previously perfectly competitive labour market and the impact of the introduction of a trade union into a monopsony labour market.

Functions of trade unions

A trade union is an organised association of workers in a particular trade, profession or industry. It is formed to protect and further the rights and interests of its members. Trade unions carry out various functions on behalf of their members:

- **Improving earnings of members:** Individual trade unions use collective bargaining to try to earn higher wages for their members. **Collective bargaining** occurs when a trade union agrees wage rates for its members with an employer or group of employers. In many firms, particularly those employing a large number of workers, firms prefer to negotiate with a trade union rather than lots of individual employees. Trade unions will also try to improve wages through agreements such as productivity bonuses or other rewards based on performance. The trade union movement as a whole will also campaign for improvements in wages for workers, such as the campaign for a living wage. The Trades Union Congress (TUC) acts as a pressure group on behalf of the trade union movement in the UK.
- **Protection of jobs:** A trade union will work to prevent employers from dismissing workers unfairly. It will negotiate with the employer and, if necessary, represent its members in cases of unfair dismissal.
- **Improving working conditions:** Trade unions will negotiate with firms in order to improve features of jobs such as working hours, the safety of the working environment and security of job contracts.
- **Protecting pension rights:** Trade unions will also try to ensure the welfare of their members after they have retired by negotiating secure and valuable pensions.
- **Counteracting the monopsony power of large employers:** Where a large employer can use its buying power to push down wages, a trade union can act as a monopoly seller in order to ensure a fairer wage for workers.

Factors affecting the ability of trade unions to influence wages and levels of employment in different labour markets

The **trade union wage gap** is defined as the percentage difference in average hourly wages of trade union members in comparison with non-trade union members. In 1995, the trade union wage gap was 30.3% in the public sector and 15.3% in the private sector. By 2014, this had fallen to 21.6% in the public sector and 8.1% in the private sector. Although there are many other factors impacting upon these differentials, the trade union wage gap is some indication of the impact that trade unions have on wage levels. It can be concluded that the influence of trade unions on wages has fallen over the last 20 years.

Trade unions have a different impact in different labour markets because the factors that affect their ability to influence wages and levels of employment vary considerably between different occupations. It can also be affected by the market structure.

Factors affecting the ability of trade unions to influence wages and levels of employment

- **The ease with which a job can be substituted:** If workers are unskilled, it will be easier to find other workers to replace them at short notice. However, highly skilled workers cannot easily be replaced because of the time required to train new workers. In many industries, workers can be replaced by other factors of production, most notably by capital equipment.
- **Union density:** This measures the percentage of the workforce or occupation that belongs to a trade union. If a firm has a high percentage of union members, industrial action, such as a strike, will have a much greater negative impact because it will be difficult to have enough workers to continue production.
- **Macroeconomic factors:** In a period of economic growth, firms are likely to be making high profits and can afford to pay higher wages to their workers. At such a time there will also be fewer workers unemployed and so workers will have a much stronger bargaining position because the employer will find it difficult to replace existing workers if the workers are unhappy. The profitability of individual firms is a key influence on workers' wages. In general, highly profitable firms will pay higher wages than less profitable ones, even if workers have the same occupation.
- **Government legislation:** Since 1979, government legislation has restricted the power of trade unions to take industrial action. A trade union may be liable to compensate firms for lost profits as a result of illegal actions by its members. Industrial action by trade unions must now be agreed by a recent secret ballot. Currently the level of industrial action in the UK is low in comparison to previous decades and many other countries.
- **Financial consequences of trade union action:** The employer is much more likely to agree to the requests of the trade union where trade union action is likely to lead to heavy financial losses. Trade unions are reluctant to take action because their members will suffer a loss of wages if they are not fulfilling their work contracts.
- **Public support:** The majority of trade union members in the UK are in the public sector. The success of a trade union is therefore likely to be influenced by public support for the action taken. Government may wish to avoid conflict with a group of workers that have public sympathy. This factor can influence private sector firms too. Recently, Sports Direct received a lot of negative publicity over

its failure to pay the national minimum wage and has now declared that it will ensure that its workers will not be paid below the minimum wage.

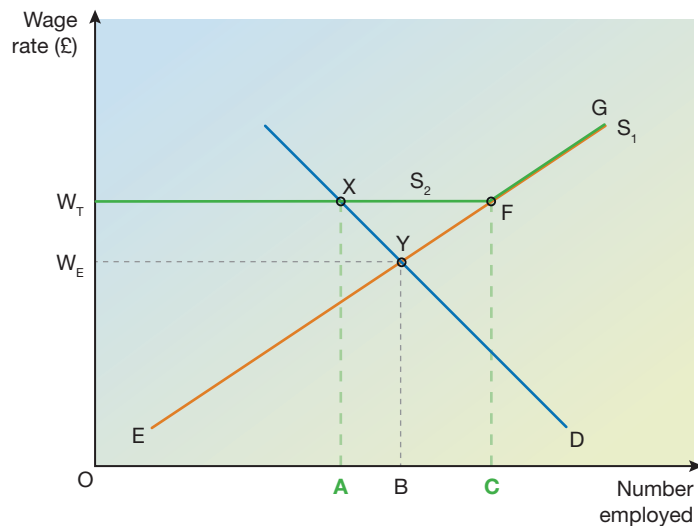
The impact of the introduction of a trade union into a previously perfectly competitive labour market

Monopoly power can be exerted by trade unions in labour markets because they control labour supply. The extent of this monopoly power will vary according to the density of union membership. The analysis in Figure 26.1 assumes that the firm has agreed to collective bargaining with the trade union. Complete control of labour supply is unlikely in practice, so trade union influence on wages and employment is likely to be more restricted than this diagram shows.

In Figure 26.1, the market labour supply without intervention by a trade union would be S_1 (the line EYG). The demand for labour is the MRP line – in Figure 26.1 this is shown as D (the blue line in the diagram). In a perfectly competitive labour market, the equilibrium where $S = D$ would be at Point Y. This gives an equilibrium wage rate of W_E and an equilibrium level of employment of OB workers.

Trade unions aim to improve workers' living standards, usually by agreeing wages above a certain level. In this case the trade union has negotiated a minimum wage of W_T . Trade union members will therefore only work for wages rates of W_T or above. Consequently, the new supply becomes the horizontal line $W_T F$. Above OC workers, employers must pay above W_T , shown by the old S_1 line. This is because these workers are only prepared to offer their labour for a wage rate that is above W_T . The supply of labour after trade union influence is therefore S_2 ($W_T F G$). The new supply curve (S_2) meets the demand curve (D) at Point X. This gives a new equilibrium wage rate of W_T , with OA workers employed.

Figure 26.1 Wages and level of employment in a trade union-influenced labour market



In comparison to a perfectly competitive market, the wage rate has risen from W_E to W_T , but this has been achieved at the expense of job losses. OC workers wanted jobs at W_T , but only OA workers were employed and so AC workers did not get jobs. However, in comparison to perfect competition, job losses are AB, as the equilibrium quantity has fallen from OB workers to OA workers.

Trade unions also seek job security and so this is a dilemma. As a rule, the opportunity cost of higher wages is fewer jobs. From a macroeconomic view, higher wages increase AD and thus jobs increase, so collectively trade unions prioritise higher wage rates. Increasingly, it is recognised that higher productivity increases MRP, securing higher profits and wages. This has led to wage agreements based on worker performance. If a trade union can help its workers to improve productivity, then there will be an increase in the demand for labour. An increase in demand leads to an increase in both the quantity of labour employed and the price of labour (the wage rate).

The analysis based on Figure 26.1 assumes that the trade union's only impact is its ability to negotiate a wage rate that is higher than the previous equilibrium wage. In some occupations, trade unions try to restrict membership and thus the number of workers able to supply labour. If this occurs, it shifts the supply curve of labour to the left, leading to an increase in wages and a decrease in levels of employment. However, this approach has tended to require a 'closed shop', in which all employees of a firm/occupation have to belong to a specific trade union. Such restrictive practices are now illegal and so it is very difficult for a trade union to increase wages by restricting labour supply.

The impact of the introduction of a trade union into a monopsony labour market

Figure 26.1 shows a labour market where there is a single seller of labour (monopsony). In Chapter 25, Figure 25.2 showed a labour market where there was a single buyer of labour (monopsony).

Bilateral monopoly exists when there is a single buyer and a single seller of a good or service. This occurs in a labour market where there is a monopsony employer and a single trade union controlling supply.

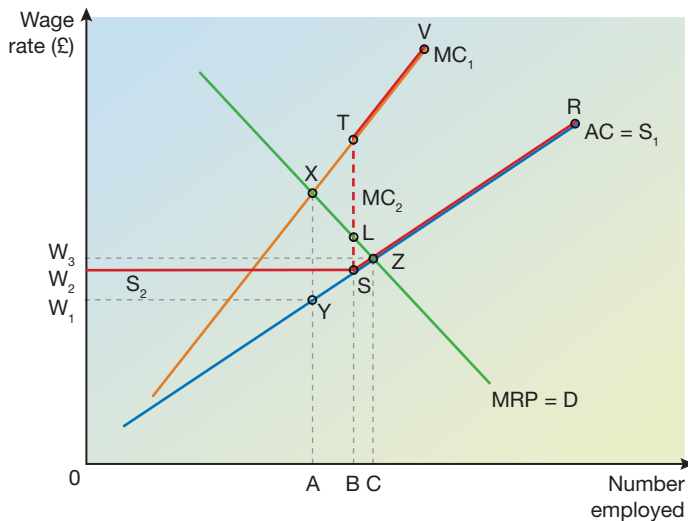


Figure 26.2 Bilateral monopoly in a labour market

In Figure 26.2, average cost (AC) = S₁ shows the perfectly competitive supply of labour and MRP = D shows the demand for labour. In a perfectly competitive market, the equilibrium would be at Point Z. Thus the equilibrium wage rate in perfect competition would be W₃ and OC would be the equilibrium level of employment.

As described in Chapter 25, the MC for a monopsony buyer of labour increases twice as steeply as the AC line. In Figure 26.2, this is shown by the line MC_1 . The monopsony employer maximises profit here ($MC = MRP$). Thus a monopsonist would choose to employ OA workers and pay a wage of OW_1 . In Figure 26.2, it is assumed that the trade union negotiates a minimum wage of OW_2 . The supply of labour is now shown by the line W_2SR .

From W_2 to S, the average cost and marginal cost of labour are now shown by the line W_2S , as each additional worker adds W_2 to the total wage bill. However, beyond Point S, employing one more unit of labour means paying a wage slightly higher than W_2 . Thus all of the existing workers must be paid marginally more than W_2 . This means that MC is now discontinuous at S, when OB workers are employed. The MC of one more worker is now shown by Point T. Thus the new MC of labour is the line W_2STV . With this new MC line, $MC = MRP$ at Point L. The profit-maximising monopsony buyer will therefore employ OB workers at a wage rate of OW_2 .

Conclusion: *The introduction of a trade union to a monopsony labour market leads to an increase in the wage rate from OW_1 to OW_2 and an increase in employment from OA to OB.*

The overall impact of trade unions

Some of the key arguments for and against trade union involvement in labour markets are noted below.

Pros

- Trade unions can help the UK to improve its productivity, through productivity agreements and encouraging and providing training of employees.
- Improvements in working conditions help the satisfaction of workers; this can lead to greater morale amongst the workforce, and therefore greater productivity and better quality goods and services.
- Where employers have monopsony power in the labour market, a trade union can ensure fairer wages.
- Trade unions protect people on low wages and therefore help government and society to achieve one of its primary aims of greater equality of income and wealth.
- Achieving higher wages leads to greater spending power for the workforce. This leads to economic growth and higher employment.

The Governor of the Bank of England has recently cited inequality as a major difficulty in the UK economy because people on high incomes have a much lower propensity to consume, and therefore economic recovery is being threatened because increases in income are not leading to significant increases in spending.

Cons

- Trade unions distort labour markets by acting as a monopoly supplier of labour. This distortion is particularly likely to hinder markets with small employers.
- Increasing wages will lead to inflationary pressure and make it difficult for the government to maintain its inflationary target of 2% CPI.
- Analysis shows that an increase in wages will lead to a decline in the number of jobs in that particular occupation.
- In trying to protect jobs, trade unions will often resist the implementation of new

technology. This will lower productivity and can mean jobs being lost to other countries.

- Trade unions tend to favour fixed working conditions. This reduces the flexibility of the workforce. It is therefore important to encourage flexible working to allow markets to operate freely.
- Intervention by trade unions can lead to market failure because the equilibrium wage level of employment will not be achieved.

Author tip

The topic of trade unions encourages strong opinions, both for and against. When analysing the consequences of trade union involvement, ensure that you support your reasoning with logical economic arguments. It can be easy to allow strong feelings to divert your arguments away from sound economic reasoning.

REALWORLD ECONOMICS 26.1

Trade union membership

In June 2015, the Department for Business, Innovation and Skills published a statistical bulletin on trade union membership in 2014. The data provide an insight into the changing nature of trade union membership and its impact on the UK labour market.

In the UK, trade union membership peaked at 13 million workers in 1979; over half of UK employees belonged to a trade union in 1979. In 2014, this figure was 6.4 million (25% of all employees).

Membership of trade unions is influenced by certain key characteristics:

- **Sector of the economy:** In the private sector, there are 2.7 million trade union members – 14.2% of private sector workers. In the public sector there are 3.7 million trade union members – 54.3% of public sector workers. These percentages have fallen gradually since 1995, following a more dramatic decline in the

1980s, although private sector membership of trade unions has increased in the last two years.

- **Gender:** 28% of female employees are in trade unions, compared with 22% of male employees.
- **Occupation and education:** Employees in professional occupations are more likely to be trade union members than employees in other occupations; while only 20% of UK employees work in the professions, they represent 36% of total trade union membership. 31% of employees with a degree belong to trade unions, whereas only 20% of employees without a formal qualifications are trade union members.
- **Country of birth:** 27% of UK-born employees are in a trade union, whereas 18% of non-UK born employees are members of a trade union.
- **Age:** Older people are more

likely to belong to a trade union. 38% of trade union members are aged over 50, but only 28% of employees fall into this age group.

- **Size of firm:** In larger firms (with more than 50 employees), 33% of employees belong to a trade union. In firms with fewer than 50 employees, 16% of employees belong to a trade union. Pay is decided by collective bargaining with a trade union in 39% of large firms compared with 15% in firms with less than 50 workers.
- **Employment contract:** 26% of employees with a permanent contract belong to a trade union; for employees with temporary contracts, this falls to 15%.

The trade union wage gap is the percentage difference in average hourly wages of trade union members in comparison with non-trade union members. In 1995, the trade union wage gap was 30.3%

in the public sector and 15.3% in the private sector. By 2013, this had fallen to 19.8% in the public sector and 7% in the private sector. However, in 2014 this trend was reversed – the trade union wage gap in 2014 was 21.6% in the public sector and 8.1% in the private sector.

Source: Department for Business, Innovation and Skills, ONS

Exercises

Total: 15 marks

- 1** Select three characteristics that influence trade union membership and explain why they affect trade union membership in the way that they do. **(9 marks)**

- 2** Explain possible reasons for the decline in the trade union wage gap since 1995. **(6 marks)**

Discussion point

Given the existence of a trade union wage gap, why is it the case that trade union membership has halved since 1979?

Review questions

Total: 25 marks

- 1** Which one of the following factors will help a trade union to secure higher wages? **(1 mark)**
 - A A job that can be easily automated
 - B Higher unemployment
 - C Higher wages in competitor countries
 - D Lower profitability
- 2** Define the term 'bilateral monopoly'. **(3 marks)**
- 3** What is meant by the term 'trade union wage gap'? **(3 marks)**
- 4** How does 'union density' affect a trade union's ability to secure higher wages? **(4 marks)**
- 5** Explain *one* possible benefit of trade unions to the UK economy. **(4 marks)**
- 6** Explain *one* possible disadvantage of trade unions to the UK economy. **(4 marks)**
- 7** Draw a diagram to show the impact of a trade union on wages and the level of employment in a perfectly competitive labour market. **(6 marks)**

The national minimum wage

Key concepts from Year 1

No prior knowledge from Year 1 is needed for this chapter.

This chapter introduces the concept of a national minimum wage (NMW) and examines its effects upon labour markets. The introduction of the ‘living wage’ as a replacement for the NMW for workers over 25 is discussed. The chapter concludes by comparing the advantages and disadvantages of a national minimum wage/living wage.

Background

In the UK the national minimum wage (NMW) was introduced in April 1999. The hourly rate was £3.60 for employees aged 22+ and £3 for employees aged 18 to 21. Since then the rate has increased at a faster rate than inflation and, in October 2015, it was £6.70 per hour for workers aged 21+ and £5.30 per hour for workers aged 18 to 20. For 16 to 17-year-olds, it was £3.87 per hour.

In 2015, the Chancellor of the Exchequer, George Osborne, announced the introduction of the ‘national living wage’ (NLW) from April 2016. The national living wage has been initially set at £7.20 per hour and applies to workers aged 25 or over. For workers aged under 25, the national minimum wage (NMW) rates applicable in October 2015 continue to apply. In his Budget Speech, the Chancellor of the Exchequer indicated that the ‘national living wage’ would increase over the lifetime of the current Parliament; he anticipated that the rate would be £9 per hour in 2020.

Key term

The **national minimum wage** is an hourly wage rate set by the government, which represents the minimum wage that an employer must pay workers above a certain age.

The effects of a national minimum wage upon labour markets

(At this point you may wish to refer back to Figures 26.1 and 26.2, from Chapter 26.)

The national minimum wage (NMW) is the hourly rate of pay below which workers must not be paid. In terms of its impact on labour markets, the NMW operates in exactly the same way as a trade union agreed minimum wage rate. Thus Figure 26.1 in Chapter 26 can be used to show the impact of the NMW in a perfectly competitive market. Similarly, Figure 26.2 illustrates its impact in a monopsony market.

Figure 26.1 showed that a minimum wage would only have an impact if it were above the equilibrium wage rate in a particular market; for this reason, most workers in the UK are not directly affected by the national minimum wage. In markets where the national minimum wage is above the equilibrium wage, Figure 26.1 shows that the wage level will increase (to the NMW) but the number of workers employed will fall. Thus, higher unemployment in that particular occupation will be the opportunity cost of the higher wage rate.

In effect, Figure 26.2 illustrated the impact of a NMW in a monopsony market. In this market, the wage level increased and the number of workers employed also

rose. Hence in monopsony markets both wages and employment levels will rise. The opportunity cost will be that the monopsony loses some of its supernormal profits.

The advantages and disadvantages of a national minimum wage

The advantages

- **Reduce poverty/improve equality:** A government aim is to create greater equality of income and wealth. Since the national minimum wage targets those on low incomes, it should help some families to overcome poverty.
- **Encourage more workers to enter the labour market:** Low wages may act as a disincentive to workers to offer their labour, particularly if welfare benefits match the wages they could receive in employment. A national minimum wage should ensure that the level paid to low-wage workers is sufficient to encourage more workers to seek employment.
- **Decreased general/cyclical unemployment:** Studies of the effects of the national minimum wage since its introduction in 1999 indicate that, contrary to many people's expectations, it is likely to have increased levels of employment. Although it may lead to lower levels of employment in occupations in which the national minimum wage applies, it injects more money into the pockets of lower paid workers. This money is then spent, thus increasing GDP.
- **Increased productivity:** This may result because higher wages provide more job satisfaction, which can improve the productivity of workers.
- **Focus on higher added value:** In developed economies, such as the UK, the NMW is likely to encourage firms to focus on providing goods and services with high added value, rather than on products that can only be produced competitively if wages are kept low.
- **Control of monopsony power:** If a monopsony buyer of labour dominates a market, this will lead to lower wages and lower levels of employment in contrast to a perfectly competitive market. The minimum wage reduces some of this monopsony power and leads to a higher wage and more jobs.

The disadvantages

- **Higher unemployment:** Analysis of perfectly competitive labour markets indicates that a minimum wage leads to lower levels of employment in that industry.
- **Impact on costs:** Higher wages will lead to higher costs, which will normally be passed on to the consumer in the form of higher prices. Thus, consumers will be adversely affected by the national minimum wage, particularly if it affects firms that make goods that have price inelastic demand.
- **The minimum wage becomes the norm:** In the UK there are over 1 million workers being paid the national minimum wage and living wage; this is a much higher figure than when the NMW was introduced in 1999. This might suggest that firms resist paying a higher wage and that it may actually be reducing wages in some markets.
- **Regional differences:** A national minimum wage does not take into account major differences in regional wage rates. In places such as London, the cost of living is much higher, but the national minimum wage does not reflect these additional costs.
- **Age discrimination:** The NMW has always led to lower pay for teenagers, even where they are carrying out the same job as an older worker. With the new living

wage being paid at the age of 25, the minimum wage discriminates against far more young people than before.

- **Operational costs:** The NMW imposes additional costs on firms, which must show that they are meeting legal requirements. It also imposes costs on the government, which must enforce the legislation.

The impact of the minimum wage depends on its level in comparison to wages in other occupations. When the NMW was first introduced, it was set at quite a low level and therefore had a limited impact. Over time, the NMW has risen faster than inflation and now has a greater impact (over 1 million workers receiving the NMW and living wage). The government has announced its intention of increasing the NMW, so that eventually the national minimum wage will reach a level that is 60% of the median wage in the UK. In 2015, the median wage was £26,500 and so the national minimum wage would need to be £15,900 to achieve this target. Figure 27.1 shows how the national minimum wage in selected countries compares to their median wage.

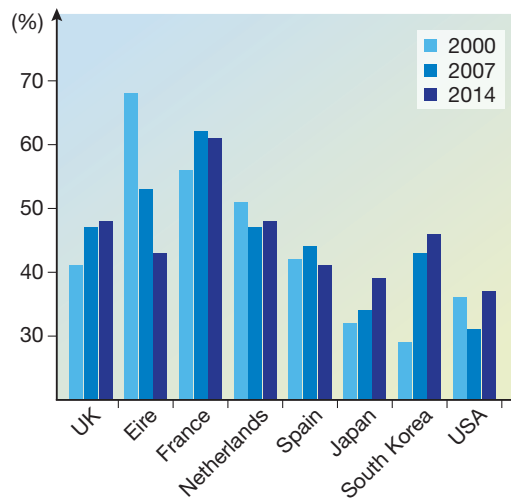


Figure 27.1 National minimum wage as a percentage of median wage in selected countries

REALWORLD ECONOMICS 27.1

The national living wage (NLW)

Since 1 April 2016, the UK has a 'national living wage' (NLW). This



Is the 'national living wage' really a 'living wage'?

is £7.20 per hour and applies to workers aged 25 or over. For younger workers, the national minimum wage (NMW) rates continue to apply. The government indicated its intention to increase the NLW to £9 per hour by 2020.

The use of the title 'national living wage' has caused some disquiet in certain circles. Some charitable pressure groups introduced the idea of a living wage

to describe the pay needed by full-time workers wishing to afford a decent standard of living. These groups use a model devised by Loughborough University, which in January 2016 calculated that the living wage was £9.40 per hour in London and £8.25 in the rest of the country. This independent living wage is a voluntary agreement that some firms use in order to set a rate that guarantees a decent standard

of living for their employees. The Living Wage Foundation argues that the national living wage is not a living wage because it does not consider the cost of living.

The new national living wage is expected by many to lead to higher unemployment. The government's own assessment is that the NLW is likely to increase business costs by £700 million in the first year, and by a higher figure if the rate is increased in future years, as indicated by

the Chancellor of the Exchequer. A survey of 340 companies that employ 1 million workers found that a quarter of firms believe they will hire fewer workers as a result of the national living wage. One-third of firms said that they would increase prices to customers in order to pay higher wages caused by the national living wage. There is serious concern that some low-pay sectors, such as care services, will need to cut back on services and

staff because the higher living wage coincides with a time of cutbacks in government finance to such services.

Discussion points

- 1 Is the introduction of the national living wage beneficial to the UK economy?
- 2 Based on the article and Figure 27.1, is the aim of a minimum wage equal to 60% of the median wage a realistic target?

Review questions

Total: 25 marks

- 1 In 2016, for workers aged 25 and over, the national minimum wage (NMW) was replaced by the:

A LW	B MNW	C MW	D NLW	(1 mark)
------	-------	------	-------	----------

- 2 The introduction of the national minimum wage into a perfectly competitive labour market will lead to:

A Higher wages and higher levels of employment in that market	B Higher wages and lower levels of employment in that market	C Lower wages and higher levels of employment in that market	D Lower wages and lower levels of employment in that market	(1 mark)
---	--	--	---	----------

- 3 The introduction of the national minimum wage into a monopsony labour market will lead to:

A Higher wages and higher levels of employment in that market	B Higher wages and lower levels of employment in that market	C Lower wages and higher levels of employment in that market	D Lower wages and lower levels of employment in that market	(1 mark)
---	--	--	---	----------

- 4 The national minimum wage was £6.50 in 2014. This was 48% of the median wage in the UK. Assuming no change in the median wage, what would the national minimum wage need to be to meet the government's target of it equalling 60% of the median wage? (2 marks)

- 5 Define the term 'national minimum wage'. (3 marks)

- 6 Explain two possible advantages of the national minimum wage. (6 marks)

- 7 Explain two possible disadvantages of the national minimum wage. (6 marks)

- For Question 8, choose to answer the 'Either' or the 'Or' question.

- 8 Refer to Figure 27.1. Explain one possible reason why the minimum wage as a percentage of the median wage has:

EITHER increased so dramatically in South Korea;	
OR decreased so dramatically in Eire.	

(5 marks)

Discrimination in the labour market

Key concepts from Year 1

No prior knowledge from Year 1 is needed for this chapter.

This chapter examines discrimination in the labour market. It considers the conditions necessary for wage discrimination and the impact of gender, ethnicity and other forms of discrimination on wages, levels and types of employment. The chapter concludes with an assessment of the advantages and disadvantages of wage discrimination for workers, employers and the economy as a whole.

Introduction

'Equal work' does not require the same job to be undertaken. Whether two jobs are roughly equivalent in worth is based on four main factors: skill, effort, level of responsibility and working conditions.

Direct discrimination exists when an employer pays someone more than a person from another group for doing the same job.

Wage discrimination is not just concerned with equal pay for the same jobs. If two different jobs require comparable levels of skill, even if the types of skill differ, then the two jobs should have equal rates of pay. For example, case law in the UK has led to warehouse jobs being declared comparable to till workers in retail outlets. However, such comparisons require value judgements, and so most of the actual analysis of discrimination in the labour market refers to direct discrimination.

Most frequently wage discrimination is applied to imbalances in pay between genders, as male employees tend to earn more than female employees. However, it can apply to other groups of workers.

The main reasons workers suffer discrimination is because of their gender, race, disability, age or sexual orientation.

The conditions necessary for wage discrimination

In Chapter 17 we examined the conditions necessary for price discrimination. Price discrimination requires that:

- firms have a certain degree of monopoly power;
- firms are able to identify and separate different groups of customers;
- the product cannot be sold to another consumer, through what is known as secondary markets;
- different groups of customers have different price elasticities of demand.

Key term

Wage discrimination

exists when an employer pays a woman or a person from a minority group less than the wage paid to a man for 'equal work' (i.e. a job of comparable worth).

Wage discrimination is very different because it is not related to income – a payment received by a firm; instead, it is a cost – a payment to a factor of production. Nevertheless, similar conditions are needed for wage discrimination. They are:

- Firms must have a degree of **monopsony** power. In a perfectly competitive labour market, perfect information would allow workers who experience discrimination to offer their services to other firm. A firm practising wage discrimination would thus experience a lack of workers and incur higher wage costs than those that did not discriminate. It would also recruit lower quality workers. These higher costs would threaten its existence in the marketplace.
- Employers must be able to identify and separate different groups of workers. For most types of discrimination, these factors are easily recognised and so employers can discriminate on the basis of factors such as gender and race. Discrimination on the basis of sexual orientation may be more difficult for an employer.
- A person selling their labour cannot usually operate through a secondary market. However, where companies subcontract work, they do not directly control the labour force and so would be unable to discriminate. The subcontractor, on the other hand, can still discriminate.
- Workers can be discriminated against if their supply of labour differs from other groups. Employers can use factors such as working hours to identify certain categories of workers who might be less flexible in the hours that they can offer. This can be used as the basis of discrimination.

Research by the European Commission cites the following factors as the main causes of wage discrimination based on gender. Similar logic can be applied to some of the other types of discrimination.

The main causes of wage discrimination based on gender

- **The undervaluing of women’s work:** Jobs dominated by women tend to be lower paid than those dominated by men. Those skill areas where female employees dominate, such as cashier work and responsibility for people, are valued less highly than other skill areas where male employees dominate, such as those requiring physical strength and responsibility for capital.
- **Segregation in the labour market:** Women and men tend to undertake different jobs. Women often work in service sectors such as health, education and public administration. Firms value these occupations less than jobs in which males predominate, such as engineering and some manufacturing.
- **Underrepresentation at senior management level:** Women tend to be less likely to earn promotion, often because employers do not believe that they will stay in their jobs as long as men.
- **Traditions and stereotypes:** Female educational paths and consequently their eventual career choices tend to be biased towards lower-paid jobs. Most significantly, women are expected to reduce their working hours or leave the labour market to carry out childcare and elderly care, thus blocking their career progression.
- **Work–life balance:** In general, family care and domestic responsibilities are not shared equally. In the EU, 89% of men with children are working; for women this figure is 66%. This factor means that women are more likely to be involved in part-time work than men. Differences in wages arise partly because women work

fewer hours than men, although the hourly rate of pay for women is lower than the rate for men.

An ONS survey into gender discrimination concluded that there are five key factors that explain the pay gap between males and females:

- 22% of the gap is due to the different industries and occupations in which women work.
- 21% of the gap is due to differences in years of full-time work.
- 16% of the gap is due to the negative effects on wages of having previously worked part-time or having taken time out to look after the family.
- 5% of the gap is due to formal education levels.
- The remaining 36% of the pay gap could not be explained by these factors. This suggests that direct discrimination, where employers are prepared to pay men more than women regardless of their backgrounds, is likely to be the most important cause of the pay gap.

The impact of gender, ethnicity and other forms of discrimination on wages, levels and types of employment

Discrimination is based on prejudice and imperfect information and is thus a cause of market failure in labour markets. In effect, most discrimination does not consider the marginal revenue productivity of labour accurately. In practice, it is very difficult for an employer to know the MRP potential of a worker and so estimates of the worth of a worker can be influenced by personal opinion, allowing employers to introduce prejudice and discrimination into their decision making.

By underestimating the worth of women, people of certain ethnicity, disabled workers, people of certain ages and people of certain sexual orientation, employers who discriminate are reducing the demand for labour amongst certain groups of workers. A fall in the demand for labour will lead to a fall in wages and a decline in the level of employment for those workers that are being undervalued. However, the corollary of this situation is that the value of other workers is overestimated in comparison to their true worth.

The effective supply of labour to an occupation by minority groups will decrease because the employer may discard their applications. In the long run, this may discourage applications for jobs by minorities, further reducing the supply.

The overall effect of these factors is that employers who discriminate will face restricted supply and will overvalue the employees that are likely to be chosen. This will lead to an increase in the wage rate for those employees and an increase in their level of employment. For women and minority groups, it will lead to lower pay and fewer jobs.

The type of employment will also be affected. Workers experiencing discrimination will be forced into less favourable working terms, such as part-time or temporary work. In Europe, 32% of women work part-time compared to only 8% of men.

Gender pay gaps

A survey by the TUC in 2013 revealed that, on average, women in full-time employment are paid 15% less than men. For part-time work, women are paid 35% less per hour than men.

The biggest differences in pay occurred amongst health professionals, people working in culture, media and sport, and those in manufacturing occupations. On average, women were paid about 25–30% less in these occupations. Women earned less than men in 32 of the 35 occupations that were examined in the survey.

In the public sector, females earned 13.6% less than males; in the private sector the difference was 19.9%. Perhaps the most striking conclusion related to the top-level gender pay gap. The survey examined the top 10% of earners in the UK. Amongst these elite workers, men were paid 54.9% more than women.

The report also found that women were more likely to lose their jobs in an economic downturn than men.

The gender pay gap is not just a UK phenomenon. Table 28.1 shows the pay gap in selected European countries in 2006 and 2013.

Table 28.1 *Percentage differences in male pay and female pay in Europe, 2006 and 2013*

Country	Pay gap in 2006 (%)	Pay gap in 2013 (%)
European Union	17.7	16.4
United Kingdom	24.3	19.7
France	15.4	15.1
Germany	22.7	21.6
Ireland	17.2	14.4
Italy	4.4	7.3
Netherlands	23.6	16.0
Poland	7.5	6.4
Spain	17.9	19.3

In 2006 the UK had the fourth highest pay gap in the European Union and in 2013 it had the sixth highest.

The advantages and disadvantages of wage discrimination for workers, employers and the economy

For workers, wage discrimination will lead to lower wages and lower levels of employment for groups that are discriminated against. Data on the UK labour market suggest that this applies to females and all of the minority groups, with disabled and ethnic workers experiencing fewest opportunities. For workers favoured by discrimination, such as male, white workers, both wages and the level of employment are likely to increase.

For employers, discrimination is likely to lead to higher costs if they are recruiting from a smaller labour force. It can also lead to lower productivity because discrimination prevents some of the most productive workers from being employed. The only possible advantage is that if there is widespread discrimination, people who experience discrimination may lower their personal perception of their value to the labour market. This may lead to employees offering their labour at a wage rate that is below its true value. This may enable firms to increase their profits.

For the economy, discrimination is an example of irrational decision making. In order to achieve productive efficiency and allocative efficiency and thus a perfect allocation of resources, all economic agents must make decisions that are rational

and based on their desire to maximise profit/economic welfare. Irrational decisions will thus cause market failure and lead to a suboptimal allocation of resources in the economy as a whole.

REALWORLD ECONOMICS 28.1

Types of discrimination

In 2009 and 2012 the European Commission undertook a survey of perceptions of discrimination on different grounds in order to try to ascertain whether non-gender factors were as significant a cause of discrimination as gender and to see if perceptions of discrimination were improving or worsening. Respondents were asked if a certain type of discrimination was 'very widespread', 'fairly widespread', 'rare' or 'very rare'. Table 28.2 indicates the number who believed discrimination was very or fairly widespread.

The report suggested that discrimination based on gender is much lower than other types of discrimination. Ethnic origin is perceived to be a much more significant type of discrimination, with disability, sexual orientation and age (over 55) being widespread at similar levels. Other than age discrimination against younger



Discrimination can have a significant impact on the labour market

Type of discrimination	Percentage who believed discrimination was 'widespread'	
	2009	2012
Ethnic origin	61	56
Disability	53	46
Sexual orientation	47	46
Age (over 55)	n/a	45
Religion or beliefs	39	39
Gender	40	31
Age (under 30)	n/a	18

Table 28.2 *Perceptions of discrimination in the European Union*

workers, gender discrimination is seen to be the least problematic in Europe. Between 2009 and 2012, Europeans perceived a distinct fall in discrimination in most categories.

In the UK there is evidence that discrimination unrelated to gender also influences the labour market. The Equality and Human Rights Commission issued a report in 2015, entitled 'Is Britain Fairer?' Some of its conclusions indicate continuing discrimination:

- people from ethnic minorities remain less likely to be directors of companies;
- 16–24 year-olds have the lowest employment rates;
- unemployment rates have increased for disabled people;
- unemployment rates are significantly higher for people from ethnic minorities – employment amongst Pakistani/Bangladeshi women is less than half the average female

employment rate;

- Muslims experience the highest unemployment rates and the lowest hourly pay rates;
- the pay gap between younger workers and older workers has increased;
- there has been no improvement in the existing gender pay gaps of graduates;
- disabled people have experienced a greater decline in their average pay than other workers;
- poverty rates are higher for children in households headed by someone from an ethnic minority.

Discussion points

Have the UK and the EU focused too much on gender discrimination at the expense of other types of discrimination? What additional information would you require to make such a judgement?

Review questions

Total: 30 marks

- 1 For purposes of wage discrimination, the definition of 'equal work' is based on certain factors. Which *one* of the following factors is *not* used to assess 'equal work'?
 - A Level of responsibility
 - B Pay
 - C Skill
 - D Working conditions

(1 mark)

- 2 For purposes of wage discrimination, the definition of 'equal work' is based on certain factors. Which *one* of the following factors *is* used to assess 'equal work'?
 - A Effort required
 - B Individual worker's length of service
 - C Job title
 - D Location of job

(1 mark)

- 3 Which *one* of the following factors is *not* a possible cause of gender discrimination?
 - A Overvaluation of women's work
 - B Poorer promotion prospects for women
 - C Stereotyping
 - D Work-life balance

(1 mark)

- 4 A recent survey in the European Union concluded that people perceive the highest level of discrimination to be experienced by:
 - A Disabled people
 - B Ethnic minorities
 - C Older people
 - D Women

(1 mark)

- 5 Define the term 'wage discrimination'. **(3 marks)**

- 6 Define the term 'direct discrimination'. **(3 marks)**

- 7 What is meant by the term 'pay gap'? **(3 marks)**

- 8 Explain *two* conditions necessary for wage discrimination. **(6 marks)**

- 9 Explain how wage discrimination is likely to affect those workers who are not discriminated against. **(5 marks)**

- 10 Explain the impact of wage discrimination on the economic welfare of the country. **(6 marks)**

Topic 6 Exam-style questions

A-LEVEL PAPER 1

SECTION A Context – Trade unions and wage differentials

Extract A Trade union membership and density

Year	Trade union membership	Density (%)
2000	7.1 million	29.8
2005	7.1 million	28.6
2010	6.5 million	26.6
2015	6.4 million	25.0

Extract B Wage differentials between occupations

Surveys conducted by the Office for National Statistics indicate that significant wage differentials remain between different groups of occupations. Based on money wages, the average weekly wage of workers in the financial and insurance sector of the UK economy was £1018 in 2015. However, this was surpassed by weekly wages in mining and quarrying, which averaged £1150 in 2015. Wages in these high-paid sectors of the economy are considerably more than mid-range occupations such as the manufacturing of engineering products, where the average worker received £652 per week. Towards the lower end of the scale were health and care workers on £414 per week and retail workers on £316 per week.

These figures do not reveal the considerable variations often received within a sector. Doctors and nurses, for example, earn significantly more than care assistants. In part, the variations may also be a reflection of differing rates of part-time and temporary work within sectors.

Year	Occupational groups										CPI
	Mining and quarrying		Manufacturing–engineering		Retail		Finance and insurance		Health and care		
Trade union density	21%		20%		12%		17%		24%		
	AWW (£)	Index	AWW (£)	Index	AWW (£)	Index	AWW (£)	Index	AWW (£)	Index	
2000	567	100	408	100	186	100	534	100	254	100	100
2005	741	131	494	121	219	118	629	118	333	131	107
2010	966	170	572	140	261	140	887	166	398	157	123
2015	1150	203	652	160	316	170	1018	191	414	163	138

Table A Wage differentials of selected occupational groups, 2000–15

Note: AWW = average weekly wage; index = index of AWW in that occupation Sources: ONS and other sources

Table A shows how average weekly wages in these five sectors have changed since the beginning of the millennium. The increasing differentials in pay are particularly noticeable. Wages in the most highly paid sector – mining and quarrying – have more than doubled. The index of average weekly wages shows an increase from 100 to 203 in this industry. Similarly,

the finance and insurance sector recorded a 91% increase in wages over this period. The lower paid sectors received wage increases varying from 60% to 70% over the 15-year period. Prices, measured by the CPI, rose by 38% in this period.

Extract C The impact of trade unions

The growing pay gap between occupations is leading to some people reconsidering the role of trade unions. Since 1979, membership of trade unions has halved and trade union density is now 25% in the economy as a whole.

The trade union wage gap is the percentage difference in average hourly wages of trade union members in comparison with non-trade union members. In 1995, the trade union wage gap was 30.3% in the public sector and 15.3% in the private sector. By 2014, this had fallen to 21.6% in the public sector and 8.1% in the private sector. Although there are many other factors impacting upon these differentials, the trade union wage gap is some indication of the impact that trade unions have on wage levels.

The impact of trade unions on wages has prompted a campaign to increase restrictions on their power. However, although trade unions distort perfectly competitive labour markets, some economists believe that in monopsony labour markets they can reduce the degree of market failure. While organisations such as the CMA can restrict the monopoly power of oligopoly firms in product markets, some economists believe that trade unions are a necessary tool to prevent such firms exploiting labour markets through their monopsony buying power.

Questions

Total: 40 marks

- 1 Using the information in Extract A, calculate to one decimal place the percentage decline in trade union membership from 2000 to 2015. (2 marks)
- 2 Explain one reason why trade union membership in the retail industry is so low in comparison to other industries. (4 marks)
- 3 Select two occupational areas from Table A and analyse two possible reasons for the wage differentials between them. (9 marks)
- 4 Assess the view that trade unions are a negative influence on perfectly competitive labour markets but may help to reduce market failure in monopsony labour markets. (25 marks)

SECTION B Essays

Total: 40 marks

'Although the wage differential has increased between top executives and other employees, the government cannot tell people what they should be paid. However, it should act where there has been market failure because labour markets are not perfect.'

'In Europe, countries do not intervene in a way that affects workers on high levels of pay. However, only a handful of countries have not implemented a national minimum wage.'

- 1 Explain *three* reasons why labour markets may be imperfectly competitive. (15 marks)
- 2 Assess the view that countries should not intervene in a way that influences workers on high levels of pay, but that they should implement a national minimum wage. (25 marks)



Topic 7

**The distribution of
income & wealth:
poverty & inequality**

The distribution of income & wealth

Key concepts from Year 1

With the exception of income and wealth, which were introduced in Chapter 6 of the Year 1 companion textbook, no prior knowledge from Year 1 is needed for this chapter.

This chapter examines the difference between income and wealth and studies the various factors that influence their distribution. The difference between equality and equity in relation to the distribution of income and wealth is explained. The use of the Lorenz curve and Gini coefficient as measures of inequality and the interpretation of these measures are also explained. The chapter compares the likely benefits and costs of more equal and more unequal distributions. It concludes by examining the view that excessive inequality is both a cause and a consequence of market failure and that people's views of what constitutes an equitable distribution of income and wealth are based on value judgements, which thus influence policy decisions.

The difference between income and wealth

Income is a flow of earnings paid to labour over a period of time. A person's income is usually based on the number of hours worked or in the form of an annual salary. The actual payment is usually made monthly or weekly.

Economists often use 'disposable income' as the most appropriate measure of income. **Disposable income** is the level of income after the deduction of tax and national insurance payments. This is used because it represents the amount of money available for spending more accurately than **gross income** – the individual's income before tax and national insurance payments. This means that consumers' disposable incomes are influenced by government. If the government increases income tax and/or national insurance payments, there will be a decrease in consumers' disposable incomes.

Wealth is a stock of assets owned by an individual or organisation. It may take the form of cash, bank deposits and ownership of company shares, household goods (such as furniture and appliances), cars and property.

Wealth may be obtained through inheritance or through the accumulation of assets over a lifetime. It can be converted into cash (or may already be in the form of cash), but for assets such as property it may take time to convert into cash. An asset that cannot be turned into cash without loss or delay is known as an **illiquid asset**. The impact of wealth will vary according to the individual's circumstances. An individual with a low income but considerable wealth in the form of property may struggle because the upkeep and maintenance of the property may involve high expenditure.

The various factors that influence the distribution of income and wealth

An unequal distribution of income leads to inequality and occurs when there are large differences in incomes between inhabitants of a country. In the UK, as with most countries, the distribution of wealth is more unequal than the distribution of income. Inequality in income and inequality of wealth tend to be interconnected. People on high incomes can devote more of their resources to the accumulation of assets, such as property and stocks and shares. These assets often generate further income and are frequently key influences in causing an increase in the level of inequality of income. However, assets such as property can also cause a drain on an owner's resources and so, although high income tends to lead to high wealth, it is not such a certainty that high wealth leads to high income.

Since families with high wealth have tended to accumulate that wealth over many generations, there is a tendency for the distribution of wealth to be more unequal than the distribution of income.

Although the factors that influence the distribution of income and the distribution of wealth tend to be very similar, it is possible to distinguish between those that tend to affect the distribution of income and those that tend to affect the distribution of wealth.

Factors influencing the distribution of income

- **Skills and qualifications:** Skills and qualifications have a significant impact on wage levels. Workers with high levels of skill are likely to be able to secure higher incomes for two reasons: first, they are able to generate much greater output for their employers, and so their high marginal revenue productivity (MRP) causes the demand for their labour to increase; and, secondly, fewer workers are able to develop high levels of skill in most occupations and so workers with high skills are in short supply. A decrease in (or low level of) supply will lead to higher wages.
- **High levels of education:** Workers with degrees typically earn significantly more than workers who left school without formal qualifications. To some extent, educational opportunities can be influenced by a person's family background and wealth, and so education is frequently a factor that causes more inequality. Conversely, it is often the key factor enabling people to change their social class and move into a different level of income.
- **Inequality in levels of wealth:** Individuals or families with high levels of wealth can earn profit if the wealth is in the form of a business, interest if the wealth is in the form of capital, and rent if the wealth is in the form of land. In addition, these individuals have their labour to offer in return for income. Consequently, owners of wealth have the potential to generate much higher levels of income, in various forms, than people who do not possess wealth.
- **Household composition and age:** The distribution of income can be measured on the basis of individuals or households. Whichever measure is used, the composition of the household in terms of the number of dependent children and the number of working adults can have a big influence on the average or total income of a household. A household with two working parents and two grown-up children in employment will tend to have a high total income. However, a household with one

Key terms

The **distribution of income** is a comparison of the current income levels of inhabitants of a country or community.

The **distribution of wealth** is a comparison of the value of the stock of assets held by inhabitants of a country or community.

working parent and five children still in education will tend to have a low income. The age of individuals can be seen to be a key factor in household composition. The income of an individual will tend to rise as the person gains experience in the labour market and decline in the period around or after retirement, and so in itself this is a key factor influencing the distribution of income.

- **Government policy:** Although it is government policy to seek a redistribution of income in order to create greater equality in society, the extent to which different governments aim to achieve this goal varies. The methods used will also have a significant impact. Current government policy is to increase people's incomes by encouraging workers to seek employment rather than rely on benefits. However, primarily this policy is intended to increase the amount of income available to the workforce as a whole – if everyone is in employment, GDP should increase. However, where this policy does not work, it is likely to cause greater inequality in the distribution of income.
- **Market structures:** In a perfectly competitive market, it can be seen that incomes of workers become more equal and owners of other factors of production only earn normal profit. Although this is a theoretical model, the closer a market is to a perfectly competitive structure, the more equal will be the distribution of income. In contrast, markets in which there is monopoly power for sellers, or monopsony power for buyers, tend to lead to high levels of inequality in the distribution of income.
- **The functional distribution of income:** In the global economy, labour's share of national income remained at about 67% over a long period of time, although in the UK it tended to exceed 70%. In 1975 this functional distribution for labour peaked at 78% in the UK, but since then it has fallen to 70%. This has led to a significant increase in the percentage of national income going to entrepreneurs as profit and so it has created a more unequal distribution of income.

The factors noted above are just some of the key influences on the distribution of income. This distribution can also be affected by factors such as discrimination, the number of hours worked, the level of entrepreneurship in an economy, levels of unemployment and the rate of economic growth.

Factors influencing the distribution of wealth

- **Inequality of income:** Since the distribution of wealth arises from the accumulation of assets, individuals or families with high levels of income are more likely to be able to increase their holdings of wealth. It could be argued that income is the key factor influencing the distribution of wealth. This factor is closely interlinked with time. A family enjoying high levels of income over many generations will be able to maintain an increase in its wealth. In the UK, some wealthy families can trace their wealth back over many hundreds of years.
- **Inequality of wealth:** Since the ownership of wealth creates many opportunities to generate further income, a high level of wealth is likely to lead to even more wealth.
- **Inheritance traditions:** In countries such as the UK, it has been a tradition for the eldest son to inherit an estate. This tends to maintain the inequality in the distribution of wealth in the UK, as high-value estates remain in the hands of single individuals. In many other countries, wealth is shared more evenly between

the children of a family, and this is also the tradition amongst lower wealth owners in the UK. This approach leads to a more even distribution of wealth over time. This factor can be further enhanced by a tendency for wealthy individuals to marry other wealthy individuals. However, despite this factor, the distribution of wealth in the UK is more equal than the distribution of income, *in comparison to other countries*.

- **Government policy:** Governments have faced problems in their attempts to tax wealth rather than income because a tax on wealth (such as a country estate) might force the owner to sell the estate in order to pay the tax, since selling small parcels of land is unlikely to be a practical solution. Consequently, wealth tax in Britain is almost entirely based on inheritance tax. This is aimed at restricting the level of wealth that can be passed from one generation to another, and thus serves to redistribute wealth from the rich to the poor. However, there are a number of ways in which the impact of inheritance tax can be limited and so it does not redistribute wealth to the extent that was originally intended.
- **Propensity to save:** People with high incomes have a higher propensity to save because they can satisfy their wants without needing to spend most of their income. Increased savings will represent an increase in holdings of wealth, thus further increasing inequalities in the distribution of wealth.

Commentary on the distribution of income and wealth

Income

According to the *Equality Trust*, the UK has a very high level of income inequality compared to other developed countries. Figure 29.1 shows the distribution of annual incomes by household in the UK in 2015.

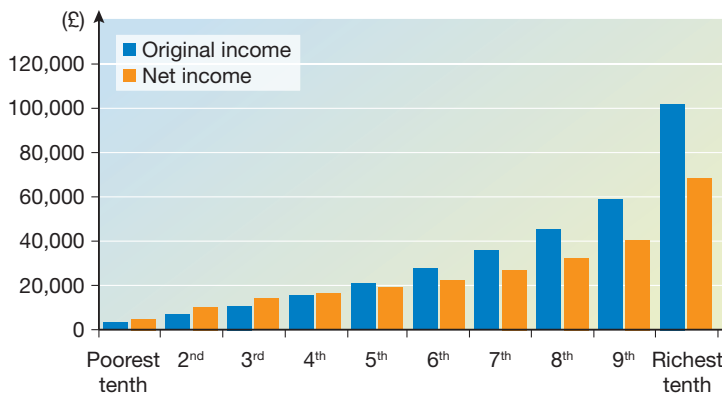


Figure 29.1 The spread of annual incomes by household in the UK, 2015 (by decile*)

Note: *A decile examines groups that represent 10% of a population – see Key note overleaf.
Source: ONS

Original income is gross income, before tax and benefits. Net income shows income after taxes and benefits have been included. The redistribution effects of taxation and government spending can be seen in Figure 29.1, with the lowest four deciles showing higher net income than original income and the highest six deciles paying more in tax than they receive in benefits. In effect, the blue columns show the extent of inequality as a result of the workings of the economy, while the orange columns indicate the level of inequality after government intervention.

Key note

In statistical analysis, groups of a population are often classified together. The four groupings below are the ones that are used most commonly in economics:

- **Percentiles:** A percentile represents 1% of the population. The first percentile is usually described as the one with the lowest value (income or wealth in this chapter), whilst the hundredth percentile is the one with the highest value.
- **Deciles:** A decile represents 10% or one-tenth of the population. The first decile represents the 10% of the population with the lowest value; the tenth decile represents the 10% of the population with the highest value. *(This pattern applies to each type of grouping.)*
- **Quintiles:** A quintile represents 20% or one-fifth of the population.
- **Quartiles:** A quartile represents 25% or one-quarter of the population.

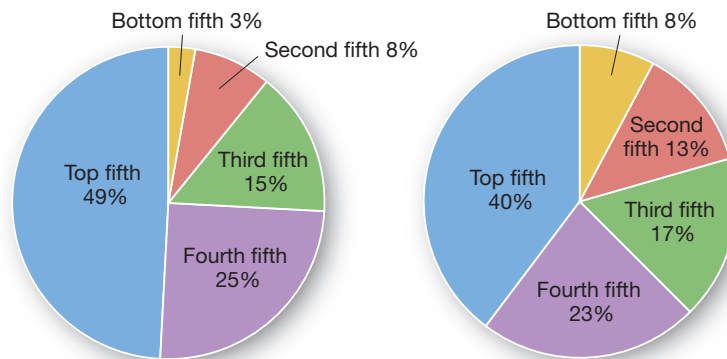
The mean household gross annual income in the UK in 2015 was £33,155; after taxes and benefits, this became a disposable (net) income of £31,786.

It can be seen from Figure 29.1 that the average original income of households in the first decile (10% of the population) was only £3738. After taxes and benefits, this rose to a disposable (net) income of £8468. The average gross income in the tenth decile was £102,366, falling to £79,042 after taxes and benefits. Before government intervention, the average original income of the tenth decile was more than 27 times higher than that of the first decile; after the effects of government redistribution, it was nine times higher.

The graph does not show the full extent of the difference between the rich and poor, because the tenth decile represents 2.67 million households. The top 1% of the UK population had an average annual income in excess of £250,000 and the top 0.1% (about 30,000 households) averaged over £1 million.

Figure 29.2 shows the distribution of income in the UK by quintile (a quintile is 20% of the population).

Figure 29.2 The share of income in the UK by quintile



(a) The share of original income (b) The share of net (disposable) income

Based on original incomes, the top quintile earns 49% of all income in the UK, with the lowest quintile earning 3%. This differential is reduced to 40% for the top quintile and 8% for the bottom quintile after redistribution.

Wealth

As with most countries, wealth in the UK is more unequally distributed than income. Figure 29.3 shows the distribution of wealth.

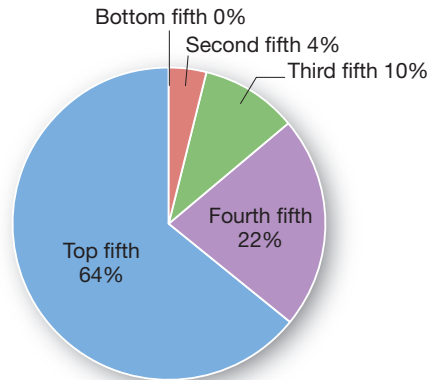


Figure 29.3 The distribution of wealth in the UK by quintile

Figure 29.3 shows that wealth is more unevenly distributed than income in the UK. The richest quintile owns 64% of the wealth in the UK (whereas the richest quintile based on income has 40% of the income). The bottom quintile owns 0% of the wealth but has 8% of the income. 50% of the UK population owns just 8% of the UK's wealth. This is less than the percentage of UK wealth owned by the richest 1% of its inhabitants.

Although the UK has a more unequal distribution of wealth than income, this is a global pattern in developed countries. In fact, the UK's distribution of wealth is more even than in most other European countries, whereas its distribution of income is more uneven than in most other European countries.

The difference between equality and equity in relation to the distribution of income and wealth

Equality is a positive measure of variations in income and wealth because it is measured by using factual data on the income and wealth of a country or region. The extent to which a country has equality or inequalities of income and wealth can be shown by a Lorenz curve or Gini coefficient (see next section).

Equity is a normative concept because it is based on people's value judgements. People take different views on whether the distribution of income and wealth is considered fair. Some focus on the unfairness of the vast differences in spending power that different levels of income can provide for different individuals and argue that unequal distributions of income are unfair. Others focus on the contributions that individuals make to the creation of income and wealth in a country through their work, entrepreneurship and ownership of factors of production. They tend to believe that inequality is fair if it reflects the greater value of different individuals to the GDP of the economy.

Equity is viewed in two different ways:

- **Horizontal equity** is when identical individuals are treated the same. If this situation applies, then it would be seen as 'just' or 'fair'. For example, individual employees have their pay deducted through the PAYE (Pay-As-You-Earn) system.

Key terms

Equality measures the extent to which different individuals have the same level of income and wealth.

Equity means 'fairness' and examines the justness of the extent to which different individuals have the same or different levels of income and wealth.

Firms are given details of how to deduct tax in a way that ensures that two individuals with the same circumstances and on the same level of income will pay the same level of tax. In recent years, there have been questions about whether there is a horizontal equity in corporation tax. The declared profits of some multinational organisations, such as Google, Amazon and Starbucks, appear to have underestimated their profits in comparison to firms that are based wholly in the UK. Critics in the UK have argued that UK-based companies, such as Costa, are competing in an unfair market because they are paying more corporation tax than their main competitors.

- **Vertical equity** relates to the treatment of individuals with different circumstances. It recognises that people are different and should therefore be treated differently, but that overall there should be fairness (justness) in the differences in the way they are treated. For example, it is accepted that richer people should pay more tax than poorer people. Furthermore, it is also widely accepted that the percentage of income taken in taxation should be greater for the person on a higher income than that taken from a person on a lower income. This is known as progressive taxation. However, there is not widespread agreement on exactly how progressive taxation should be, as this is very much a matter of opinion. Vertical equity therefore depends on value judgements. In the UK, the system of taxation and government spending generally became more progressive after the Second World War, until the end of the 1970s. Since then, it has tended to become less progressive, but Figure 29.1 shows that it is still progressive because it redistributes income from higher earners to lower earners.

The Lorenz curve and Gini coefficient

The Lorenz curve and Gini coefficient are universally agreed ways in which inequality can be measured. Both can be used for measuring the distribution of income and/or wealth; they are also used in other aspects of economics, such as measuring monopoly power.

The Lorenz curve

The Lorenz curve is a diagrammatic way of presenting inequality. The 'x' axis measures the percentage of the population while the 'y' axis measures the percentage

Key term

The **Lorenz curve** is a graphical device used to show distributional inequality. Primarily it is used to show inequalities in income and wealth.

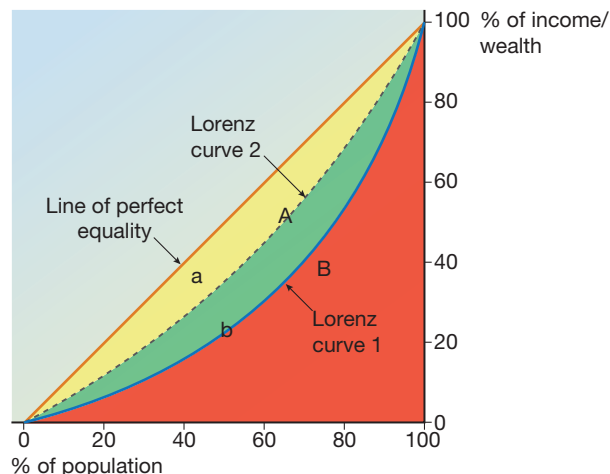


Figure 29.4 The Lorenz curve

of income. Each point on the Lorenz curve indicates the percentage of income held by a particular percentage of the population. In Figure 29.4 Lorenz curve 1 shows that the poorest 20% of the population have only about 5% of the income/wealth, whereas the richest 20% own almost 50% of the income/wealth.

In a Lorenz curve, complete equality would occur if each quintile/decile etc. had exactly the same percentage of income/wealth. This is shown in Figure 29.4 by the 'line of perfect equality'. If the Lorenz curve shows a high degree of curvature away from this line of perfect equality, then there is a high level of inequality of income/wealth. The blue line (Lorenz curve 1) shows a relatively high degree of inequality. If the Lorenz curve shows only a slight degree of curvature away from this line of perfect equality, then there is a low level of inequality of income/wealth. The dotted line (Lorenz curve 2) shows a relatively low degree of inequality.

The Gini coefficient

The Gini coefficient uses the same data to provide a mathematical calculation of inequality. It is calculated by comparing the area of the diagram between the 'line of perfect equality' and the 'Lorenz curve', with the total area to the right of the 'line of perfect equality'.

Key term

The Gini coefficient is an arithmetic measure of inequalities in income and wealth. Typically, this number ranges between zero (perfect equality) and one (perfect inequality).

Author tip

The AQA A-level specification does not require students to calculate the Gini coefficient. However, you will be expected to interpret values expressed as a Gini coefficient and explain their significance. Primarily, the Gini coefficient is expressed as a decimal between the values of 0 and 1. For example, the Gini coefficient for the distribution of income in the UK in 2013–14 is about 0.35. However, many organisations express it as a percentage and so Gini coefficients may vary between 0 and 100. In this case the current UK Gini coefficient would be expressed as '35'. You need to be prepared to come across either measure. The decimal value is the most common and is used by organisations such as the United Nations. However, the ONS – the UK government's primary provider of economic data – is one example of an organisation that uses the percentage method.

There are two Lorenz curves in Figure 29.4. For Lorenz curve 1, Area 'A' represents the area between the line of perfect equality and the Lorenz curve. Area 'A' is the area of the diagram shaded in yellow and the area shown in green. Area 'B' is the area to the right of Lorenz curve 1. This is the area shaded in red.

The formula for the Gini coefficient is as follows:

$$\text{Gini coefficient} = \frac{\text{Area A}}{\text{Area A} + \text{Area B}}$$

For Lorenz curve 2, Area 'a' represents the area between the line of perfect equality and the Lorenz curve. Area 'a' is the area of the diagram shown in yellow. Area 'b' is the area to the right of Lorenz curve 2. This is the area shaded in green and the area shown in red.

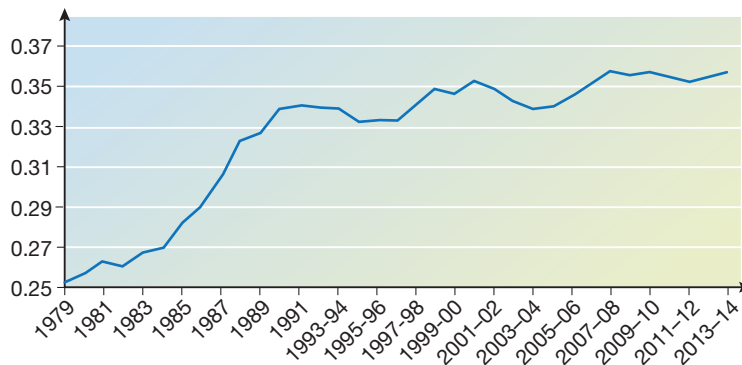
The formula for the Gini coefficient is as follows:

$$\text{Gini coefficient} = \frac{\text{Area a}}{\text{Area a} + \text{Area b}}$$

If the Lorenz curve is identical to the line of perfect equality, then there will be no Area A(a), and so the Gini coefficient for perfect equality is zero. If one person had all of the income/wealth, the Lorenz curve will trace around the bottom line and right-hand vertical line. In this case there would be no Area B (b), and so the Gini coefficient for complete inequality is one.

The Gini coefficient is expressed as a decimal between zero and one. Government measures of the Gini coefficient in the UK date back to 1960, when it was 0.25. From 1960 to 1978 it fluctuated slightly to either side of this figure, reaching its lowest ever value of 0.23 in 1978–79.

Figure 29.5 *The Gini coefficient for income distribution in the UK, 1979 to 2013–14*



It can be seen from Figure 29.5 that the Gini coefficient for UK disposable income showed a high level of equality of 0.25 at the end of the 1970s. It then rose quite steeply to 0.34 by the end of the 1980s. Since then it has fluctuated slightly but remained close to its 1989 level.

The Gini coefficient is used to show how levels of inequality have changed over time. It is also used to compare different levels of inequality, such as the differences in equality between original income and disposable income. However, it is more usually used to compare levels of inequality between different countries. This is shown later in this chapter (Real World Economics 29.1).

The likely benefits and costs of more equal and more unequal distributions

Inequality is both a cause and consequence of market failure, and so to some extent its benefits and costs depend on the exact nature of the causes of inequality. The analysis in this section will be based on an understanding of inequality in the context of the UK and other developed economies. In many instances the extent to which a factor is considered to be a significant benefit or cost will depend on value judgements.

This section will examine the benefits and costs of an unequal distribution of income. If appropriate, the opposite arguments can be applied to an analysis of the benefits and costs of a more equal distribution of income.

Benefits of an unequal distribution of income

- Inequality can be beneficial to society if it is the result of some individuals working harder than others. The prospect of achieving a higher income than other people provides an incentive for workers to stay in education, develop skills, be entrepreneurial and inventive, and work hard. Society benefits from the wealth

created by this motivation and so inequality can be a driving force behind the creation of a higher GDP in a country. A higher GDP has the potential to bring benefits to all inhabitants of a country.

- Labour is not homogeneous. Some individuals have particular skills that enhance society and therefore it is only fair that they should benefit from those skills. If consumers are prepared to pay more for a good or service, such as a visit to the theatre, then the person satisfying those needs should reap the rewards.
- Similarly, attempts to redistribute income may be deemed to be unfair. If two individuals are identical in circumstances, is it equitable to take money from a person who has contributed a lot to the economy and give it to someone who has been given the same opportunities but chosen not to make the same level of effort?
- People have different values. Many economists criticise income and GDP as measures of economic welfare, citing the need to take into consideration factors such as lifestyle and happiness. In an economy, a person who values lifestyle and happiness may consciously choose to prioritise other activities, such as time with family, at the expense of economic activity, such as work. Consequently, they would not expect to benefit from the redistribution of national income if this lifestyle choice led to their having a lower level of income for that individual.
- In the UK, inequality is most notably linked with self-employment. People at the very highest levels of income are invariably those that own businesses, rather than those who are employees of a business. These high incomes are often the reward for the level of risk taken by entrepreneurs. Thus inequality is a reflection of the importance of risk taking in society, as this enables the economy to develop and new products to be provided.

Costs of an unequal distribution of income

- Diminishing marginal utility also applies to income. The law of diminishing marginal utility applies to all goods and services. As more of a good is consumed, the additional utility from the marginal good continues to diminish. Logically, people with high incomes are likely to get less satisfaction from their marginal spending than people on low incomes who are still spending money on necessities that provide them with very high value. As economic welfare is a measure of how many wants are satisfied by an economy, it is logical to take money away from people who have a low marginal utility from their income and give it to people who have a high marginal utility. In this way, the economic welfare of the population as a whole will be increased (as long as the rich person does not lose the incentive to work).
- Inequality can be a reflection of monopoly power. In the UK, market failure is often linked to the power of monopolists and oligopolists to use their monopoly power to restrict output and charge high prices. In this way, the inequality is a reflection of a poorly performing economy that is misallocating scarce resources. Redistribution of income is therefore a way of overcoming this market failure.
- Inequality can also be a reflection of monopsony buying power. The functional distribution of income has moved in favour of the entrepreneur in recent decades, as a result of a weakening of trade union power and a decline in the use of collective bargaining whereby groups of workers meet with large employers. Currently, many individual workers agree their wages with employers, a situation that reflects considerable imbalance between the strength of the buyer of labour

and the weakness of the seller.

- Reduced spending can result from excessive inequality. People on high incomes have a low propensity to consume, as they can meet their basic needs with a small fraction of their income. Thus, inequality can reduce GDP and lead to high levels of unemployment, further lowering GDP. The redistribution of income from rich to poor will tend to have a favourable effect on GDP, unless resources are already fully employed.
- Social problems can arise from inequality. These types of costs will be examined in Chapter 30.
- Where inequality arises from inherited wealth, many people deem it to be unfair. The argument that it provides an incentive towards effort does not apply if a person has achieved a high income without any personal effort.

The weighing up of these benefits and costs requires value judgements and will also be dependent on the extent of the inequality of income. However, there is widespread agreement that the existence of inequality can be a major incentive to individuals to make great efforts to accumulate income and will therefore often have a favourable impact on a country's output and thus its productive efficiency. There is also widespread agreement that such inequality is likely to lead to a misallocation of resources and thus allocative inefficiency.

REALWORLD ECONOMICS 29.1

The distribution of income and wealth in OECD countries

A recent study by the *Equality Trust* examined the distribution of income and wealth amongst selected countries that belong to the OECD (the Organisation of Economic Cooperation and Development).

It should be noted that different organisations use different sources

of data and so these figures might not correspond exactly to those from other sources. It should also be noted that the two different ways of measuring the Gini coefficient are being employed. The Gini coefficients for the distribution of income are shown as decimals

(Figure 29.6), whereas the Gini coefficients for the distribution of wealth are shown as percentages (Figure 29.7).

The figures show that income distribution is more even than wealth distribution in all 30 countries that were surveyed. The

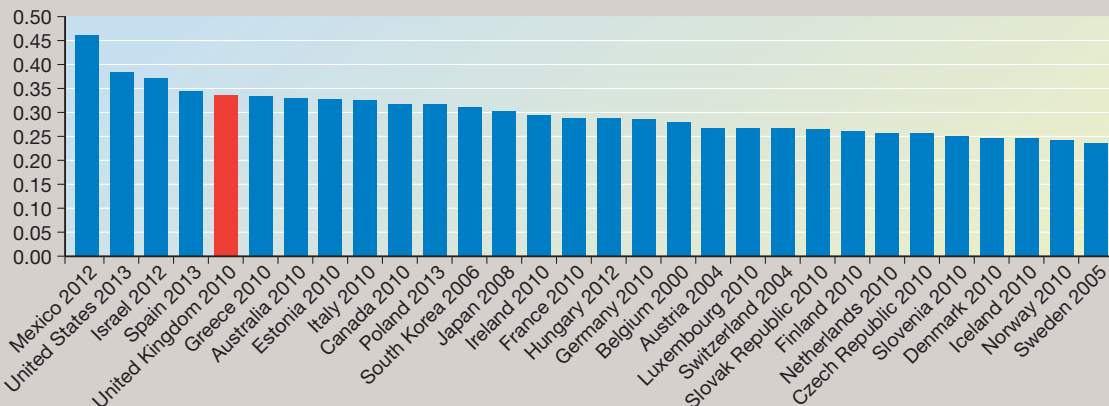


Figure 29.6 Income Gini coefficients of OECD countries

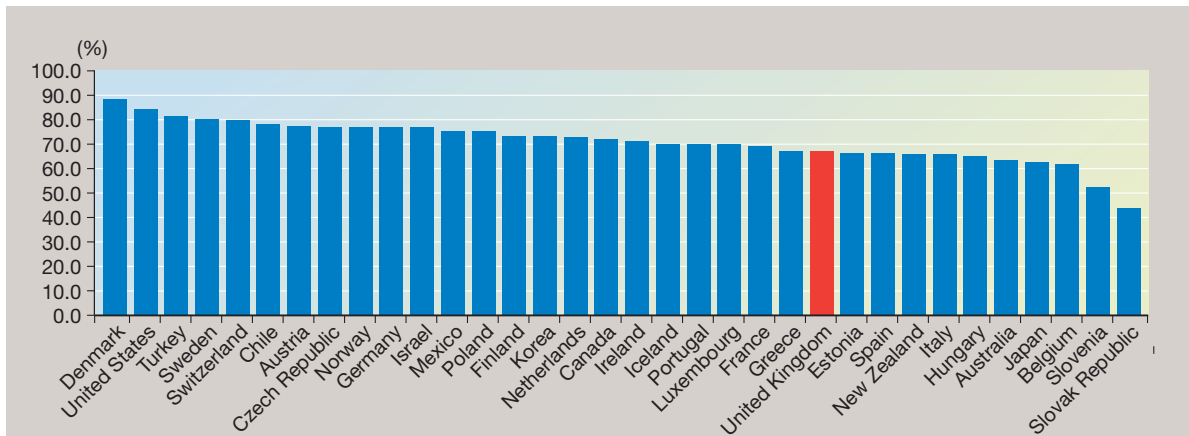


Figure 29.7 Wealth Gini coefficients of OECD countries

highest Gini coefficient for income is that of Mexico, at 0.45 (45%), whereas Sweden has the lowest Gini coefficient for income, which is below 0.25. For wealth, the highest coefficient is Denmark, with a Gini coefficient that is just under 0.9 (90%). The lowest wealth coefficient is that of the Slovak Republic, the only country with a wealth coefficient below 0.5 (50%).

In general, the Scandinavian countries tend to be those with the most even distribution of income,

with Mexico and the USA being the two countries with the most unequal distribution of income.

The pattern of wealth distribution is less clear-cut, although two former Communist states show the most equal distribution of income.

The survey showed that the UK was the country with the fifth highest inequality in the distribution of income, of the 30 countries surveyed. However, its distribution of wealth was the eleventh most equal of the 34

countries that were surveyed.

Exercise

Explain why the UK's distribution of income is so different from its distribution of wealth. (10 marks)

Discussion

You may wish to discuss the reasons for the distribution of income of another country or why there are such significant differences between the two sets of figures shown in Figures 29.6 and 29.7.

Review questions

Total: 40 marks

- 'A flow of earnings paid to labour over a period of time.' This phrase describes the term:
 - Consumption
 - Illiquid assets
 - Income
 - Wealth
- In comparison to the average distribution of other OECD countries, the UK has:
 - Greater equality of income and wealth
 - Lower equality of income and wealth
 - Greater equality of income but lower equality of wealth
 - Lower equality of income but greater equality of wealth

(1 mark)

(1 mark)

- 3** If identical individuals are treated the same, it is known as:
A Horizontal equality
B Vertical equality
C Horizontal equity
D Vertical equity *(1 mark)*
- 4** If individuals with different circumstances are treated differently, but with overall fairness, this is known as:
A Horizontal equality
B Vertical equality
C Horizontal equity
D Vertical equity *(1 mark)*
- 5** Define the term 'distribution of wealth'. *(3 marks)*
- 6** Explain why it is less practical for government to tax wealth rather than income. *(4 marks)*
- 7** What is the difference between 'equality' and 'equity'? *(4 marks)*
- 8** Explain how greater wealth can lead to a higher level of income. *(5 marks)*
- 9** Explain *two* possible benefits to an economy arising from an unequal distribution of income. *(6 marks)*
- 10** Explain *two* possible costs/problems to an economy arising from an unequal distribution of income. *(6 marks)*
- 11** Explain *two* factors that influence the distribution of income in the United Kingdom. *(8 marks)*

The problem of poverty

Key concepts from Year 1

No prior knowledge from Year 1 is needed for this chapter.

This chapter follows on from the discussion of inequalities of income and wealth by looking at the problem of poverty. It explains the difference between absolute poverty and relative poverty and examines the causes and effects of poverty.

The difference between relative and absolute poverty

The United Nations and other international organisations, such as UNESCO and the World Bank, define absolute poverty as being unable to afford the basic needs of life – food, clothing, shelter and so on. It is an absolute measure because it is compared to a fixed standard of living.

The UK has attempted to create a definition of absolute poverty. This definition is that a person is in absolute poverty if their income is below 60% of the median income in the UK in 2010–11. If living standards continue to improve, levels of poverty should decrease, based on this measure. Relative poverty in the UK tends to be measured by a similar formula. It is defined as a level of income below 60% of the current median income in the UK. Thus, a person who is just above the poverty line may fall into poverty if their income increase is below that of the median income increase. The Department for Work and Pensions (DWP) produces figures for poverty using two measures: income before housing costs; and income after housing costs.

‘Income before housing costs’ means that spending on items such as rent and mortgages are treated the same as spending on items such as food and leisure. Increased income means that you are better off, and you may choose to spend this on housing.

Since housing costs vary so much between different regions of the UK, additional spending on housing costs in places such as London are probably not a true reflection of a better standard of living. The use of ‘income after housing costs’ eliminates this effect and allows comparison to be made with the amount of disposable income

Key terms

Poverty is a state or condition in which a person or community lacks the financial resources and essentials needed to enjoy the basic needs of life, such as food, clothing and shelter.

Absolute poverty means that a person’s income is considered to be insufficient for them to satisfy basic human needs.

Relative poverty is when a person’s income is a certain level below a society’s average (median) income.

available after expenditure on housing costs. This approach also overcomes the problem of people receiving housing benefits. If rents increase, housing benefits often rise too, but this extra income is not a reflection of a better living standard.

These two measures are included because housing can account for a large proportion of spending of people in poverty.

While it is possible to overcome absolute poverty by creating a society in which all individuals are provided with their basic needs, overcoming relative poverty depends on how it is defined. There will always be people with lower levels of income than others and so, as living standards improve, there is a tendency to raise the threshold of relative poverty.

In the UK, measures of poverty are based on household income. Therefore, people sleeping rough, those living in bed and breakfast accommodation and those in institutions such as nursing homes are all excluded from government figures. These omissions mean that measures of poverty will underestimate the true level of poverty in the UK. However, measures based on household income may not accurately describe the person's situation at the time. If a person is experiencing short-term unemployment, they will have low income but may be spending their savings, and therefore their consumption of goods and services will not be in keeping with their current level of income.

Key note

The distribution of income and wealth is uneven. Since a few individuals earn much more than most of the population, the mean (total income/total population) is much higher than the median (wages earned by the person who is on the fiftieth percentile in the distribution of income). In the UK, the mean income is more than 20% higher than the median income and is close to the level of income earned by the top end of the sixth decile rather than the median (the top of the fifth decile).

Author tip

Different organisations use different ways of measuring poverty, and existing measures change over time. You should be prepared to be flexible when reading articles on poverty because definitions used may vary from those given in this chapter, as these are based on specific sources. The key point is to remember what these definitions of poverty are trying to show.

The causes and effects of poverty

The causes of poverty

Many of the factors that influence inequality in the distribution of income are linked closely to the factors that cause poverty. However, as poverty focuses on those individuals or households with the lowest levels of income, the causes of poverty are those factors that lead to lower levels of income. Some of the main causes of poverty are noted below.

- **Low levels of income from employment:** Although unemployment creates more poverty, widening differentials in pay have led to far more families experiencing poverty (relative poverty), even though the main income earner of the family is

employed. The government has announced a change in policy from the national minimum wage to a national living wage. The target for the national living wage is for it to reach 60% of the median income in the UK. As this is the line that defines relative poverty, the national living wage should lead to a significant reduction in relative poverty, provided this strategy succeeds.

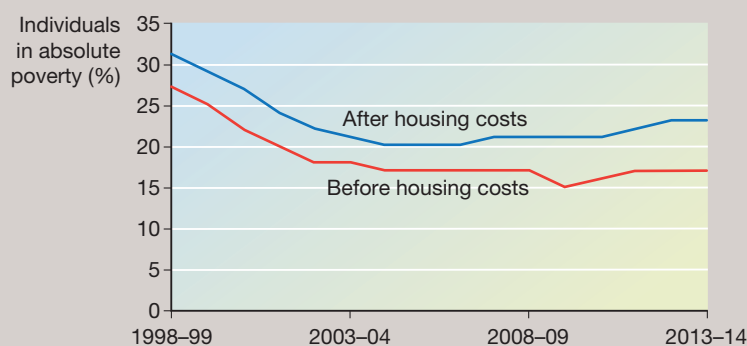
- **Unemployment:** This is a major cause of poverty, particularly for individuals who are experiencing long-term unemployment. Without a regular income, unemployed individuals are reliant on welfare payments from the state. In recent years, there has been a particular emphasis on the need for individuals to be self-supporting and so cuts in these benefits have tended to accentuate levels of poverty amongst the unemployed.
- **Dependency on others and age:** Unemployed workers are dependent on the state. Other groups that suffer relatively high levels of poverty are those dependent on other people's incomes, such as housewives and children. Levels of child poverty have also increased noticeably in the UK. People of retirement age are dependent on state support, and early analysis of poverty tended to focus on the elderly as the group that suffered most heavily. However, the level of effort made to prevent poverty amongst the elderly has led to significant improvements in their situation, in comparison to some other groups in society. Age is a key factor because people of working age are more able to take measures to reduce or overcome poverty.
- **Government policy:** Government approaches to the redistribution of income will be considered in the next chapter. The extent to which the government prioritises this aim can have a major impact on levels of poverty. It can also have a different impact on absolute and relative poverty. Current government policy focuses on wealth creation and the trickle-down effect this has on everybody's level of income. Consequently, absolute poverty should improve if this strategy is successful. However, since it depends on wealth creation through the market mechanism, it is likely to widen income differentials and therefore is less likely to reduce levels of relative poverty.

REALWORLD ECONOMICS 30.1

Poverty in the UK

An investigation of poverty in the UK by Full Fact, a pressure group that provides independent information on topical issues, reveals that absolute poverty has declined in the UK since 1998–99. However, it also shows that absolute poverty after housing costs actually increased between 2003–04 and 2013–14.

Figure 30.1 shows that, 'after housing costs', 31% of the UK population were in absolute poverty in 1998–99. This figure fell to 20%



Source: Households below average income (HBA): 1994–95 to 2013–14

Figure 30.1 Absolute poverty in the UK, 1998–99 to 2013–14

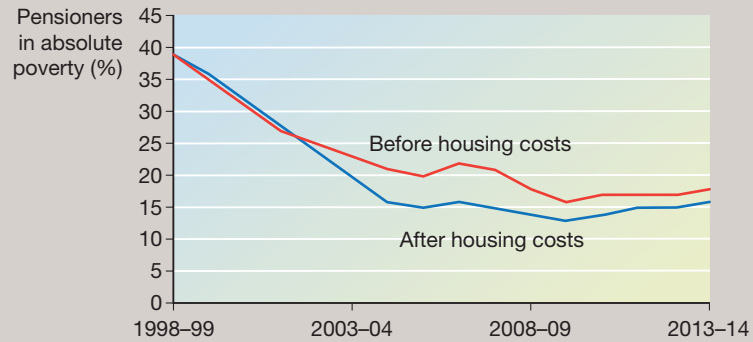
in 2004–05, after which it stabilised but then increased slightly to 23%. 'Before housing costs', absolute poverty was 27% in 1998–99 and then fell to 18% in 2002–03. It then stabilised but fell slightly to 15% in 2009–10, since when it has risen slightly to 17%.

Figures 30.2 and 30.3 provide data on absolute poverty for pensioners and children respectively. It can be seen that pensioner poverty fell dramatically from 1998 to 2004, since when it has stabilised at a level below the average UK household. Although child poverty also fell in this period, it has remained above the UK average in every year.

Source: Full Fact (www.fullfact.org.uk)

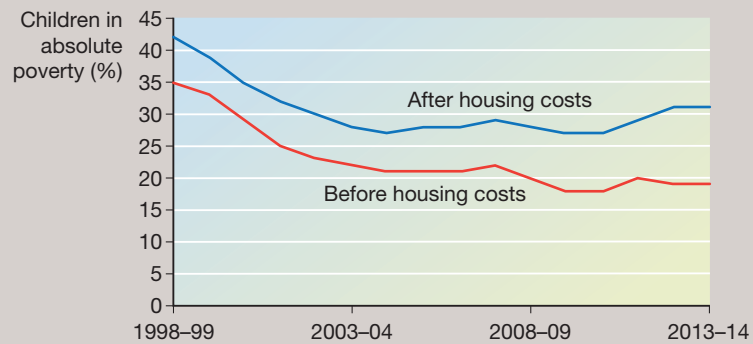


The level of child poverty remains above the UK average



Source: Households below average income (HBA): 1994-95 to 2013-14

Figure 30.2 Pensioner poverty, 1998-99 to 2013-14



Source: Households below average income (HBA): 1994-95 to 2013-14

Figure 30.3 Child poverty, 1998-99 to 2013-14

Exercises

Total: 20 marks

- 1 Absolute poverty 'before housing costs' is lower than absolute poverty 'after housing costs' for most households. However, absolute poverty 'before housing costs' is higher than absolute poverty 'after housing costs' for pensioners. Explain why this is the case. (6 marks)
- 2 Explain why poverty for children has not fallen in the same way as poverty for pensioners. (6 marks)
- 3 Analyse why poverty fell significantly in the early part of the millennium but has subsequently stabilised. (8 marks)

- **Low levels of education and training:** Since education and training are key influences in enabling people to improve their living standards, individuals who have low levels of education and training are unlikely to have high earning power and are more likely to be unemployed.
- **Country and/or region of residence:** For absolute poverty, country of residence is invariably the key factor that influences an individual's chance of living in poverty. In many African countries, the infrastructure of the country is not adequate to

support its population in their attempts to achieve living standards that would be considered to be relative poverty in most European countries.

- **Family background:** People born into wealth are more likely to avoid poverty than people that have little or no wealth. Social mobility has declined in the UK and so children in poverty are more likely to experience adult poverty than children from more privileged backgrounds.
- **Health:** Ill-health or illness can be major causes of poverty, as they can prevent individuals from taking steps to overcome poverty.

The effects of poverty

- **Ill-health and death:** Poverty can prevent individuals from taking measures to ensure good health and so ill-health and a lower life expectancy are more likely to be experienced by people living in poverty. Poverty is also a major cause of stress and can affect mental health too.
- **Limited education and training and thus poor job prospects:** Limited education and training is both a cause and an effect of poverty. Children in households that experience poverty are less likely to stay in education for as long as other children and also less likely to receive job training. Consequently, their job prospects are poor and the cycle of poverty is likely to be repeated for them as adults.
- **Lower productivity for the individual and economy:** People living in poverty tend to have fewer labour skills and poorer health and so they are less likely to be as productive in their jobs. Workers with low labour productivity tend to have a low MRP. Consequently, the demand for their services will be low and so they are likely to experience low wages and/or unemployment.
- **Lifestyle:** People in poverty tend to experience a poorer lifestyle because they are unable to afford as many goods and services as others and they are less likely to have domestic luxuries or take holidays.

Review questions

Total: 25 marks

- 1 In the UK, poverty is defined as a level of income which is below:
 - A 70% of the median income
 - B 60% of the median income
 - C 50% of the median income
 - D 40% of the median income

(1 mark)
- 2 What is the difference between absolute poverty and relative poverty?

(4 marks)
- 3 Explain why the UK's median income is used when measuring poverty, rather than the UK's mean income.

(4 marks)
- 4 Explain two possible causes of poverty for pensioners.

(8 marks)
- 5 Explain two possible effects of poverty on children.

(8 marks)

Government policies to influence the distribution of income & wealth & to alleviate poverty

Key concepts from Year 1

This chapter builds on Chapter 28 of the Year 1 companion textbook. No other prior knowledge from Year 1 is needed.

This chapter examines the government policies that are available to influence the distribution of income and wealth and to alleviate poverty. It concludes by considering the economic consequences of such policies.

Policies available to influence the distribution of income and wealth and to alleviate poverty

(The redistribution of income and wealth was discussed in Chapter 28 of the Year 1 book. You are advised to study Tables 28.1 and 28.2 in that chapter because they provide useful data to show how government policies influence the distribution of income and wealth and alleviate poverty.)

Some of the main ways in which government can influence the distribution of income and wealth and thus alleviate poverty are through:

- **welfare benefits**, such as unemployment benefit, housing benefits and pensions. By providing cash benefits to those without income or on low income, the government can improve the standard of living of those most likely to experience poverty. Many welfare benefits take the form of **tax credits**. Where a person does not use all of their tax allowances because their income is low, a form of *negative* income tax is paid – payments from the government to support that person's income;
- **subsidised government services**, such as health. By subsidising these services, government is giving 'benefits in kind' to members of society. These benefits are often not related to income. However, where these services are 'means-tested', people on lower incomes are more likely to receive 'benefits in kind'. People on high incomes are more likely to pay for services such as education and health and so, overall, these benefits can help to alleviate poverty;
- **progressive taxation** to fund spending. A progressive tax system is one that takes a larger percentage of income from those on higher incomes.

Other policies to influence the distribution of income and wealth and alleviate poverty are indicated below:

- **Legislation such as the national living wage:** From April 2016, the national living wage replaced the national minimum wage. The rate rose from £6.70 per hour to £7.20 per hour for workers over the age of 25. By guaranteeing a minimum wage for most workers, the government is focusing on ensuring that unscrupulous employers do not force workers into poverty through exploitation. However, this wage is below 60% of the median wage in the UK and so the million workers who are currently paid the national minimum wage are likely to be classified as living in poverty, if they are living independently.
- **Reducing unemployment:** This is one of the government's macroeconomic targets. As long-term unemployment is a major cause of poverty, measures to reduce unemployment are likely to have a favourable impact on poverty. Since macroeconomic targets were introduced in 1944, low unemployment has been one of the key focuses of successive governments.
- **Spending on social funds:** Government spending on services such as subsidised housing, child benefit, free education and free health all contribute towards easing the burden on people on low incomes and thus alleviating poverty. The greater the proportion of GDP being spent in these areas, the greater is the impact on the alleviation of poverty. A recent government decision to limit child benefit to the first two children in a family from April 2017 has been criticised by many charities, which believe that this will create higher levels of poverty amongst vulnerable families.
- **Expenditure to encourage equality of opportunities in education and jobs:** Since education and training are instrumental in helping people to escape poverty, government policies that increase equality of opportunities in education will help people to escape poverty. Furthermore, equality of opportunities in employment, through legislation such as disability discrimination and race relations, protects minority groups from unfair treatment by employers in the labour market.

Key notes

Two important factors that influence the redistribution of income and wealth are 'fiscal drag' and 'the earnings trap'. These are described below.

Fiscal drag: Since tax allowances do not always keep pace with inflation, taxpayers on low incomes can find themselves moving from the situation of being below the income tax threshold to one in which they are liable to pay income tax. This can therefore make it hard for a person on very low income to escape poverty. Many people who are classified as being in poverty in the UK still earn sufficient money to pay income tax. Since 2010, the basic tax allowance threshold for income tax has increased significantly above the rate of inflation and so fiscal drag as a cause of poverty has become less of an issue in recent years. (The main impact of fiscal drag currently applies to people paying the 40% tax rate.)

The earnings trap: Since many benefits are 'means tested' (based on income and personal circumstances), people can find it hard to escape poverty because any additional income they earn leads to a reduction in the benefits to which they are entitled. This 'earnings trap' means that it is very difficult to improve living standards if one is on a low income. This situation provides a disincentive to work. Recent government policy has been to reduce the impact of the earnings by reducing welfare payments in order to encourage people to earn income rather than receive benefits.

The economic consequences of policies to redistribute income and wealth

The consequences of policies to redistribute income and wealth extend far beyond the realms of economics. Commentary will often reflect the political perspectives of commentators. There is also a moral perspective to such decisions, focusing on whether it is right or wrong to allow people to experience poverty. *The argument below focuses on economic consequences only.*

Equity or efficiency?

The policies described in the previous section should achieve one of the government's central aims: the redistribution of income and wealth and the alleviation of poverty. However, it has other objectives too, notably its macroeconomic aims of full employment, economic growth, low inflation and a balanced balance of payments.

The use of the government's budget to overcome poverty conflicts with the use of the government's budget to achieve its macroeconomic aims. Supply-side policies have become the key feature of fiscal policy in recent decades. These measures to increase the productive capacity of an economy (its production possibility boundary) conflict with the use of the government's budget to alleviate poverty.

A central feature of supply-side policy is the creation of incentives to work. This requires fiscal policy to encourage the working of the market mechanism in order to achieve the most efficient use of scarce resources. By encouraging entrepreneurship and a willingness to work, the economy will be both dynamically efficient (inventive) and productively efficient, as workers strive to improve their skills and earnings. Taxation policy to encourage this aim requires the lowering of taxes on income, so that workers keep most of their earnings and are therefore more motivated to increase their incomes. The implication for government spending is that there should be a reduction in benefits paid to people who are not working, in order to encourage them to work. If these policies succeed, there will be **economic efficiency**.

The redistribution of income and wealth is based on the philosophy of equity (being 'fair' to people). For vertical equity, this requires help for those on lower incomes, so that they can enjoy a reasonable living standard. In terms of the government's budget, it requires a progressive system of taxation and spending.

Efficiency and equity are in direct conflict because they require opposite approaches. Equity takes away incentives to work and may therefore reduce GDP. As a consequence, there will be less income to redistribute and so there is a debate as to whether the reduction of wage differentials is desirable for the economy as a whole.

Redistribution of income can increase overall levels of economic welfare because people on low incomes gain greater satisfaction from their marginal expenditure than those on higher incomes, and thus equity can lead to greater **allocative efficiency**.

An alternative view is that the redistribution of income can help productive efficiency of an economy because people on low incomes have a higher propensity to consume. Thus the redistribution of income from rich to poor leads to greater consumption, providing entrepreneurs with a greater incentive to produce goods and offering more opportunities for employment in the labour market.

***Conclusion:** Although there is conflict between redistribution of income and the desire to create incentives to support the growth of the economy, economists do not all agree on the outcomes of these two approaches.*

REALWORLD ECONOMICS 31.1

Comparison between the distributional effects of the 2010/11 Budget and the proposed 2017/18 Budget

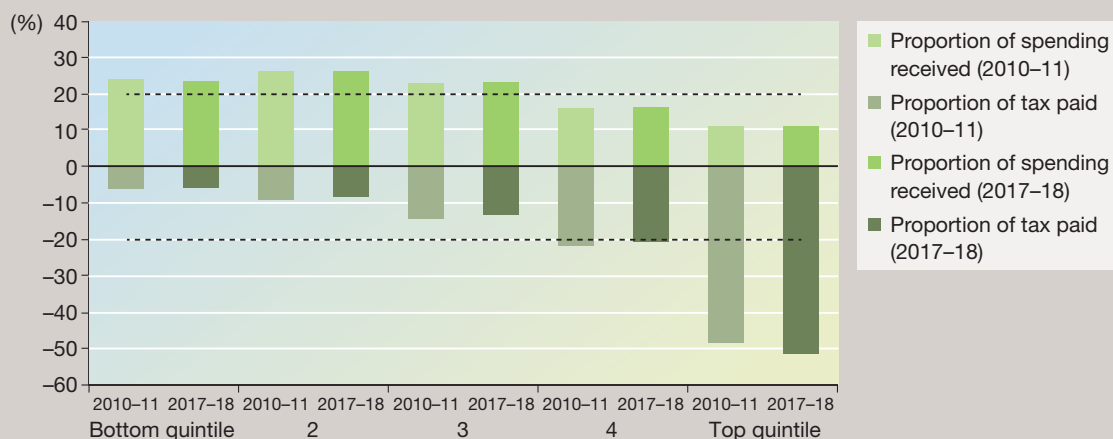


Figure 31.1 Impacts of policy decisions on the distribution of public spending, 2010–11 and 2017–18

In July 2015, The Treasury published documents showing how policy decisions on taxes and public spending would affect household income and expenditure by quintile.

In Figure 31.1, the two dotted lines show a perfectly even distribution of income (20% each) between each quintile. Any figure above the top dotted line shows a quintile that received an above

average percentage of government spending. Any figure below the bottom dotted line shows a quintile that paid an above average percentage of taxation.

It can be seen from Figure 31.1 that the second quintile continues to benefit more than the first quintile.

Figure 31.1 also shows how taxation is expected to become slightly more progressive, with only

the top quintile paying a higher percentage of tax in 2017–18 than in 2010–11.

Source: HM Treasury, July 2015

Exercises Total: 10 marks

- Why does government spending benefit the second quintile more than the first quintile? **(4 marks)**
- Why does the bottom quintile still pay as much as 6% of its income in tax? **(6 marks)**

Impact on individual markets

The redistribution of income affects different markets in different ways. Goods with high income elasticity of demand are usually luxuries, such as yachts and expensive houses, which appeal to those on high incomes. Goods with negative income elasticity of demand are known as inferior goods, such as cheap cuts of meat, and are usually purchased because an individual's income is insufficient to enable them to buy a superior version of this good. A redistribution of income from rich to poor would lead to fewer people being able to afford luxury products and will also lead to a reduction in sales of inferior goods, as people on lower incomes can now afford better quality products.

The decline in both of these individual markets will be beneficial for mid-range products, such as family cars and shoes.

Overall, this will lead to a decline in the price and consumption of both luxuries and inferior goods, and an increase in the price and consumption of mid-range goods.

Review questions

Total: 30 marks

- 1 Which *one* of these taxes is progressive?
A Council tax
B Fuel duty
C National Insurance
D VAT (1 mark)

- 2 Which *one* of these taxes is the *most* regressive?
A Fuel duty
B Income tax
C National Insurance
D Tobacco duty (1 mark)

- 3 Define the term 'fiscal drag'. (3 marks)

- 4 Explain the term 'earnings trap'. (3 marks)

- 5 Explain how government taxation and spending can be used to redistribute income in order to alleviate poverty. (8 marks)

- 6 Explain *two other* policies that can be used to redistribute income in order to alleviate poverty. (6 marks)

- 7 Analyse the conflict between the government's desire for efficiency and its desire for equity. (8 marks)

Topic 7 Exam-style questions

A-LEVEL PAPER 1

SECTION A Context – The distribution of income and wealth

EXTRACT A Taxation and redistribution

	Quintile groups of ALL households					All households
	Bottom	2 nd	3 rd	4 th	Top	
(a) Percentages of gross income						
Direct taxes	9.7	11.0	15.6	19.2	23.5	18.9
Indirect taxes	28.1	19.1	17.8	15.3	11.3	15.3
All taxes	37.8	30.2	33.3	34.5	34.8	34.2
(b) Percentages of disposable income						
Indirect taxes	31.2	21.5	21.0	18.9	14.8	18.8
(c) Percentages of expenditure						
Indirect taxes	20.7	20.5	20.4	19.5	17.0	19.0

Source: ONS

Table A Taxes as a percentage of gross income, disposable income and expenditure for all households by quintile groups, financial year ending 2014, UK

EXTRACT B Poverty or inequality?

In the year 2000, the UN declared a target of halving the global percentage of people in absolute poverty by 2015. Absolute poverty was defined as anyone living on an income of less than \$1 a day. As a result of price changes, the target was subsequently increased to \$1.25. This target was achieved in 2010, five years ahead of schedule.

In 1980, more than half the world's population was living in absolute poverty. It is now just over 20%. However, in areas such as sub-Saharan Africa, the number of people in poverty has grown.

Changes in the definition of poverty make it difficult to assess progress. In 1944, the Beveridge Report defined it as the lack of 'what is necessary for subsistence'. The current definition (60% of median income) gives an income of approximately £14,000. With such a wide difference between the definitions of UK poverty and world poverty, UK citizens are less persuaded by arguments relating to the need to overcome poverty in the UK. However, discussions concerning inequality bring about more of a reaction relating to unfairness.

According to YouGov, support for austerity has grown since the recession. Between 2011 and 2014, the number of people who believe that cuts are good for the economy grew from 34% (with 51% opposed) to 46% (with 35% opposed).

Source: Article in the *Guardian*, by John Lanchester, 5.9.14

EXTRACT C Foodbanks

The growth of food banks in the UK during and after the recession suggests that poverty is not a problem that has been solved. The Trussell Trust is a charity that provides help for people in

need. In 2015, it provided over 1 million three-day emergency food supplies to people in crisis – an increase of 3% on 2014. The main reasons for food bank referrals are benefit delays (29%) and benefit changes (14%). Low income accounts for 22% of referrals to its food banks. On average, people use foodbanks three or four times a year.

Poverty and benefit changes have also been associated with changes in spending patterns. In recent years there has been a relative increase in ‘non-discretionary’ expenditure, such as rent, and an increase in focus on low-priced goods, such as inferior goods.

Sources: Trussell Trust; ONS

Questions

Total: 40 marks

- 1 The average amount of tax paid by the first quintile of income earners is £4886. Using the information in Extract A, calculate the average income of an individual in the first quintile of income earners. (2 marks)
- 2 Explain why there is no worldwide agreement on how countries should measure poverty. (4 marks)
- 3 With the help of a diagram, explain how overcoming poverty might lead to a change in the pattern of expenditure on inferior goods in the UK. (9 marks)
- 4 Assess the view that government’s use of government spending and taxation should focus on equity rather than efficiency. (25 marks)

SECTION B Essays

Total: 40 marks

Inequality of income and wealth are permanent features of the UK economy. The effective working of the market economy can be one of the factors leading to inequality, but it is also a key factor in creating higher GDP, and thus enabling the UK to reduce levels of poverty.

- 1 Analyse three reasons for the inequality of income and wealth in the UK. (15 marks)
- 2 Assess the view that taxation is the most important government policy for the redistribution of income in order to alleviate poverty. (25 marks)



Topic 8

The market mechanism, market failure & government intervention in markets

How markets & prices allocate resources

Key concepts from Year 1

Chapter 22 of the Year 1 companion textbook provided full coverage of the 'rationing, incentive and signalling functions of prices in allocating resources and coordinating the decisions of buyers and sellers in a market economy'. It is vital that this content is revisited in order to understand this chapter.

This chapter builds on the functions of prices by examining the advantages and disadvantages of the price mechanism and of extending its use into new areas of activity.

The advantages and disadvantages of the price mechanism

Advantages

The rationing, incentive and signalling functions of prices provide a template for the advantages of the price mechanism as they show how it achieves an efficient allocation of resources. These three advantages are:

- **Rationing:** Price adjusts in a way that ensures that scarce resources are rationed so that they are allocated to the production of the most highly desired goods. Consumers have unlimited wants. They prioritise those wants in the form of offering prices for goods: offering a high price indicates a high level of want, whereas a low price would be offered if the good provides relatively little satisfaction to the consumer. The price also impacts on suppliers. A high price encourages more resources to be allocated to producing a good, whereas a low price will discourage supply. Combining supply and demand leads to an equilibrium price at which the willingness to supply matches the effective demand for a good. Thus the price mechanism is an effective way of balancing the cost of providing a good, which influences supply, with the value that consumers place on a good, which influences demand. Due to scarce resources, virtually all goods are rationed to some extent – many people prefer to use the word 'allocating' rather than 'rationing'.
- **Incentive:** Price acts as an incentive for consumers to decide on the best pattern of expenditure. A low price means that more of a good can be consumed and more wants satisfied, but this must be matched to the economic welfare or happiness (utility) received from the good. Price also acts as an incentive to suppliers – a high price encourages firms to supply more goods. Equally important, price encourages efficient use of resources because if a firm can produce a good at a cost that is below the market price, then it can increase its profit.
- **Signalling:** Changes in price provide a signal to buyers and sellers to show them

whether to buy or sell a different quantity of goods in a particular market. A high price signals to firms to supply more goods but also acts as a signal to buyers to contract their demand. As a high price is usually an indication of scarcity, this signalling helps to overcome scarcity because it encourages firms to produce more while encouraging consumers to buy less, until scarcity is overcome.

There are other advantages of the price mechanism.

- It is impersonal and therefore unlikely to be influenced by personal bias. The ‘invisible hand’ acts objectively in decisions to allocate resources.
- It is cost effective because there are no administration costs, unlike in markets such as a command economy. Changes in the allocation of resources do not therefore involve any additional costs beyond the prices paid by buyers.
- It is very quick to identify scarcity and shortages. Prices can be highly volatile if there are sudden changes in supply or demand, in order to shift resources quickly in the direction that suits the needs of consumers and sellers.
- In contrast to other types of allocation, the price mechanism encourages effort and inventiveness, and so it should lead to both productive and dynamic efficiency. Furthermore, if there are no externalities, the resources will be allocated according to the wishes of consumers and so allocative efficiency will be achieved.

Disadvantages

The advantages described above work on the premise that the price mechanism works perfectly in order to allocate resources. However, where market failure occurs, the price mechanism does not lead to a perfect allocation of resources. Consequently, the disadvantages of the price mechanism can be described by the various causes of market failure.

These arguments have been presented in Chapters 23–28 of the Year 1 book and so they will not be developed here. Table 32.1 provides a summary of the nature of the disadvantages arising from each of these types of market failure.

Cause of market failure	Nature of resulting disadvantage
Public goods	Underproduction of public goods
Positive externalities	Underproduction of goods with positive externalities
Negative externalities	Overproduction of goods with negative externalities
Merit goods	Underproduction of merit goods
Demerit goods	Overproduction of demerit goods
Imperfect information	Non-rational decisions, leading to an imperfect allocation of resources
Asymmetric information	Non-rational decisions, to the detriment of the person with least information
Monopoly power	Underproduction and excessive price
Monopsony power	Underuse of resources and lower rewards to sellers (usually of factors)
Immobility of factors	Underproduction of goods
Inequality of income/wealth	Unfair allocation of resources

Table 32.1 *Potential disadvantages of the price mechanism*

The price mechanism provides an impersonal but objective way of allocating

resources. It reacts quickly to changes in supply and demand and is cost effective because it requires no administration. The advantages are likely to outweigh the disadvantages where the market is perfectly competitive, or close to being perfectly competitive. However, where there is market failure, the price mechanism will lead to an imperfect allocation of resources.

Extending use of the price mechanism into new areas of activity

In the last 35 years, the use of the price mechanism has increased, with fewer goods being provided through direct government involvement. This change has applied both to the UK and to most other developed nations. Furthermore, where government allocates goods and services, their production is often left to the private sector. Decisions on production often rely on price or market mechanism.

There are three main new areas of activity into which the use of the price mechanism has been extended. These are:

- charitable goods or those provided by voluntary labour;
- internalisation of markets within firms in order to use the price mechanism;
- use of private sector organisations to provide public sector goods and services.

Charitable goods or those provided by voluntary labour

Most charities now use the market and price mechanism in their activities. For example, charity shops will only accept donations if they can be sold at a price that provides them with a profit. If a gift is deemed unlikely to sell, then it will be refused by the charity.

Monetary gifts to charities are often organised by external marketing organisations. Employees of these firms use marketing data to identify potential charitable donors and then use telephone or internet marketing to persuade people to donate or increase the level of their existing donations. These employees are paid a commission, based on the level of donation they achieve.

Many charities continue to use this approach, though it has led to criticism, because it can be very effective in achieving high levels of donation.

The market for body parts for medical needs has also changed in recent years. The DVLA has been instrumental in achieving much higher levels of organ donation, as a result of their communications with people requiring driving and car licences. Effective 'marketing' seems to attract supply.

Internalisation of markets within firms in order to use the price mechanism

Traditionally, firms produced their own goods to sell. The firm's structure was also based on the essential services that a firm would need, such as a marketing department, IT services, payroll, HRM and so on.

Over time, firms have outsourced many of these services. External organisations bid to run a service, such as payroll, for a particular price. If this price is lower than the costs of running the firm's own payroll, the service is outsourced. In this way, the price mechanism operates to improve the efficiency with which resources are used – if the firm cannot provide a service more cheaply than another organisation, it leaves the provision of that service to the external organisation.

This logic also led to outsourcing of production to other countries, particularly China.

REALWORLD ECONOMICS 32.1

The market for blood

Research by the World Health Organisation (WHO) shows that in wealthy countries, blood is mostly supplied through voluntary donations. In poorer countries, blood supplies come from paid donors or family and friends of the patient in need. In 2011, the WHO estimated that 36% of countries rely solely on voluntary donations, 28% receive the majority of their blood through voluntary donations, while the remaining 36% cover less than half of their needs through voluntary donations.



Donating blood – a voluntary activity in the UK

In the UK, donating blood is a voluntary activity. In many other countries, donors are paid for their blood. Although the latter process tends to encourage higher supply, it is also associated with lower quality blood supply, particularly since AIDS and mad cow disease. Since 1975, the WHO recommends that 100% of blood supply should be covered by donations. Quality is not the only reason for this recommendation; many believe that it is unethical to buy and sell products such as blood.

However, there is a market for plasma – one of the components of blood. In the USA, paid donors provide 81% of plasma products, and USA firms collect almost 70% of the world's plasma for use across the world. The market for plasma operates according to economic theory – the price fluctuates according to supply and demand. In times when there is a greater need for plasma (excess demand),

the price rises and attracts more donations.

Matching supply and demand of blood is more difficult because there is no price mechanism. Supply tends to be lower in the winter and holiday periods. Interestingly, sometimes when demand is high, such as a natural disaster, donors respond by increasing supply. Moreover, in order to balance supply with demand, organisations such as the NHS time their marketing effort to coincide with occasions when blood supply is low. Although research indicates that using the price mechanism to influence supply is relatively unsuccessful in the UK, it also indicates that non-monetary benefits, such as gift cards, pens and lottery tickets, can have a positive effect.

Discussion point

Should the price mechanism be used in the market for blood, as it is in the market for plasma?

However, although this process generally provides a lower cost service or good, the quality can suffer and so many firms are reconsidering this strategy. Bringing the production of goods and services back to the firm is known as insourcing.

Use of private sector organisations to provide public sector goods and services

In order to overcome some types of market failure, the government uses the price mechanism. Two examples of this process are 'property rights' (see Chapter 33) and 'pollution permits' (see Chapter 36).

In the UK, government provision of public goods and merit goods is increasingly based on the price mechanism rather than a command economy approach. The use of the price mechanism by the public sector tends to take three main forms:

- **Outsourcing:** This occurs where a private organisation bids to offer a particular public service. For example, roadbuilding has always been funded by the public sector but built by construction companies, based on the price charged (and other

aspects of their bids). Most government offices outsource services such as cleaning and catering.

- **Public-private partnerships (PPPs):** The private sector company builds, maintains and operates public sector infrastructure, such as a hospital or school. In return, it is paid an agreed price by government for its services.
- **Vouchers:** State provision of some services, such as nursery care, is often based on vouchers. Parents whose children attend a nursery can use these vouchers to pay a private sector nursery for their services.

Review questions

Total: 25 marks

- 1 Which *one* of the following is *not* a function of the price mechanism?
A Incentive function
B Matching function
C Rationing function
D Signalling function (1 mark)

- 2 Which *one* of these statements describes a feature of the price mechanism?
A It encourages effort
B It is costly to operate
C It is slow to react to market changes
D It requires a high level of bureaucracy (1 mark)

- 3 Which *one* of these examples of market failure leads to an *overproduction* of goods?
A Goods with negative externalities of consumption
B Immobility of factors
C Monopsony power
D Public goods (1 mark)

- 4 Which *one* of these examples of market failure leads to an *underproduction* of goods?
A Demerit goods
B Goods with negative externalities of consumption
C Goods with negative externalities of production
D Monopoly power (1 mark)

- 5 Explain *two* advantages of the price mechanism as a means of allocating resources. (8 marks)

- 6 Explain how there may be problems with the price mechanism as a means of allocating resources if there is asymmetric information. (4 marks)

- 7 Explain how the price mechanism is used to allocate resources through the use of nursery education vouchers. (4 marks)

- 8 Explain how outsourcing of government projects uses the price mechanism. (5 marks)

Aspects of market failure

Key concepts from Year 1

This chapter combines three elements of market failure. These are:

- public goods, private goods and quasi-public goods;
- positive and negative externalities in consumption and production;
- merit and demerit goods.

The vast majority of the AQA A-level content on these three areas is covered in Chapters 24–26 of the Year 1 companion textbook, so it is essential that you review them.

This chapter focuses on three main elements of market failure, which are included in the AQA A-level specification but *not* included in the AQA AS specification and were therefore not covered in the Year 1 (AS) book. These three elements are listed below:

- ‘The tragedy of the commons’
- ‘Property rights’
- ‘Using diagrams showing MSB and MSC to illustrate market failure’

Combining your understanding of ‘the tragedy of the commons’ with your knowledge from Chapter 24 of the Year 1 book will provide you with a good understanding of ‘Public goods, private goods and quasi-public goods’.

Combining your understanding of ‘property rights’ with your knowledge from Chapter 25 of the Year 1 book will provide you with a good understanding of ‘Positive and negative externalities in consumption and production’.

In the Year 1 book, the use of supply and demand diagrams to illustrate market failure was shown in Chapter 25 (to illustrate positive and negative externalities) and Chapter 26 (to illustrate merit and demerit goods). This chapter will show you how to use MSB and MSC diagrams to illustrate market failure. Although supply and demand diagrams can be used to illustrate market failure, MSB and MSC diagrams are a more sophisticated and precise way of showing market failure and you should be able to show levels of market failure using this technique.

The tragedy of the commons

The economist W.F. Lloyd introduced this concept in 1833. In the UK, it was customary at that time for communities to have common land, which was available for use by all residents. Lloyd noted that if an individual owner of cattle grazed more cows on the common land, then that individual would reap greater benefits. Thus it could be seen to be a rational decision. However, if these additional cows led to an overgrazing of the land, the common could be depleted or even become incapable of supporting cattle – an outcome that would be detrimental to other users of the common and the community as a whole.

Key term

The tragedy of the commons occurs when individuals pursue rational self-interest in the use of a common resource, but in so doing they create adverse effects for the group or society.

In 1968, an article in the journal *Science* by ecologist Garrett Hardin reawakened interest in this situation. From an economic perspective, it is an example of market failure relating to public goods. If all members of the community have access to common land, it can be described as ‘non-excludable’. Since public goods do not operate within the market mechanism, their provision by government (usually local government) is an ideal solution. However, public ownership of such land becomes more of a problem if there is overgrazing by certain individuals.

Overfishing is a very topical example of ‘the tragedy of the commons’. If one fisherman restricts the size of his catch, in order to preserve stocks of fish, it is likely to be detrimental to that individual if other fishermen do not take the same action. In places, voluntary agreements are often the best solution to the ‘tragedy’. However, due to the international nature of fishing, it is more difficult to achieve voluntary restrictions on fishing. As a consequence, organisations such as the governments of certain countries and the European Union have implemented fishing policies, with fishermen being allocated a restricted level of fish that they are able to catch.

The tragedy of the commons applies to most environmental problems. While damage to the environment reduces the welfare of society as a whole, the people or organisations that cause the damage are likely to benefit financially from the actions that they are taking. Ultimately, intervention is needed to ensure that the needs of the community are protected against problems created by the self-interest of individual economic decision makers.

Why the absence of property rights leads to externalities in both production and consumption and hence market failure

Key term

Property rights describe the legal ownership of a resource by an individual or organisation.

Where an individual has a property right, such as ownership of an asset like a building, they can choose how to use the asset, when or if to sell it to someone else, and how to seek compensation if someone damages the value of their asset. In economics, the entrepreneur’s ownership of resources such as capital equipment is another example of property rights.

Market failure can occur when there is an asset to which no one has property rights. For example, no one owns the environment as such. Consequently, when firms make economic decisions about production, and when consumers make economic decisions about consumption, the environment is unlikely to figure in their thinking. A firm will receive no direct negative consequences from pollution of the environment; likewise, a consumer, such as a smoker, will not directly receive negative consequences from the smoke that they emit into the atmosphere.

Overall, a lack of property rights leads to three main problems:

- overuse of scarce resources;
- overproduction of goods with negative externalities;
- free-rider problems, such as using intellectual property without paying.

Typically, government can intervene to prevent these negative externalities through legislation, such as limits to pollution or the banning of certain substances.

However, **extending property rights** can lead to fewer negative externalities. The advantage of this process is that it uses the market mechanism to determine the value of the externality, rather than relying on a subjective assessment by a government official or judge. Globally, there has been agreement on the concept of ‘territorial

water'. Governments of countries can claim ownership of the sea within a certain distance of their coastline and license/sell the rights for fishing in that area of the sea. These licences are set by government to ensure a desirable level of fishing in order to protect stocks.

Property rights have also been extended into assets such as music, books, design, intellectual property and pollution.

Using marginal social benefit (MSB) and marginal social cost (MSC) diagrams to illustrate market failure

Introducing MSB and MSC

In Year 1 we saw that an **externality** is the effect of an economic decision, such as the manufacture or purchase of a good, on other people or organisations, whose interests were not taken into account. A **positive externality** exists when the effect is beneficial to the outsiders. A **negative externality** occurs when the action has an adverse effect on the outsiders. The four types of externalities are:

- positive externalities in consumption;
- negative externalities in consumption;
- positive externalities in production;
- negative externalities in production.

In Chapter 25 of the Year 1 textbook, a number of concepts were introduced. The definitions of these concepts are reproduced below.

Private costs – the financial cost to an individual or firm of an economic transaction undertaken by that individual or firm. We have seen earlier that the supply of goods in a perfectly competitive market is the marginal cost line. Therefore, when examining supply, private costs = marginal cost = supply.

Private benefits – the financial benefit to an individual or firm of an economic transaction undertaken by that individual or firm. This is measured by the amount that consumers are prepared to buy at any given price. Therefore, private benefits = marginal benefit = demand.

External costs – the value of negative externalities arising from the production and/or consumption of a particular good.

External benefits – the value of positive externalities arising from the production and/or consumption of a particular good.

Social costs – the full cost to society of an economic activity, taking into consideration both private costs and external costs.

Social benefits – the full benefit to society of an economic activity, taking into consideration both private benefits and external benefits.

Showing the optimum allocation of resources using MSB and MSC

Equilibrium can be shown through marginal analysis. Where the additional benefit (marginal benefit) of the consumption (demand) of a good is equal to the additional cost (marginal cost) of the production (supply) of that good, the market is in equilibrium. This is where supply equals demand.

In a perfectly competitive market, the best allocation of resources is achieved where

supply (marginal private costs) equals demand (marginal private benefits). Where there is market failure, the market equilibrium quantity is not the best output because externalities are not considered by the market mechanism. To assess whether market failure has occurred, social costs and social benefits need to be considered:

Marginal social costs (MSC) – the cost to society of the last/marginal unit of output, taking into consideration both private costs and external costs.

Marginal social benefits (MSB) – the benefit to society from the last/marginal unit of output, taking into consideration both private benefits and external benefits.

Thus:

Marginal social costs (MSC) = Marginal private costs (MPC) + External costs (the level of negative externalities)

Marginal social benefits (MSB) = Marginal private benefits (MPB) + External benefits (the level of positive externalities)

Since marginal social costs show the true cost to society of the use of resources and marginal social benefits show the true benefits to society from consuming those resources, a perfect allocation of resources occurs where the two are equal.

Conclusion: A perfect allocation of resources occurs where $MSC = MSB$.

In any situation in which externalities exist, there is likely to be market failure because there is a discrepancy between social costs/benefits (which should be used to ensure that a perfect allocation of resources exists in a market) and private costs/benefits (which decide the allocation of resources in a perfect market).

Demand = supply where $MPB = MPC$. Up to the equilibrium quantity, every good sold gives the consumer more benefit than its cost to society, as the MPB line is above the MPC line. Therefore, $MPB = MPC$ where welfare is at its highest *if there are no externalities*. However, the market mechanism does not take externalities into account and so market failure occurs if there are externalities.

Externalities are a value judgement and so economists will place a value on them.

Showing externalities in consumption

If a product provides **positive externalities in consumption**, such as a talented songwriter whose songs are enjoyed by people listening to the radio, then the MPB underestimates its true economic value and so economists add the positive externality to the MPB line to get marginal social benefit (MSB). The MSB measures the benefit to both consumers and those affected externally by the product. (If there are negative consumption externalities, then these are subtracted from the MPB to show MSB. In effect, $MPB + \text{value of externalities} = MSB$.)

In Figure 33.1, Point X shows the market equilibrium where MPB (demand) and MPC (supply) are in equilibrium. Q_1 is the equilibrium quantity. As the product gives positive externalities of consumption, MSB is greater than MPB, as shown by the red arrows (the vertical distance between MPB and MSB). In this diagram, it is assumed that there are no production externalities and so MPC and MSC are the same. Society's welfare is greatest where $MSB = MSC$. This is at Point Y. Q_2 is consequently the level of output that maximises economic welfare. The market equilibrium quantity (Q_1) is too low and so the difference between Q_1 and Q_2 is the level of underproduction of this good. Market failure has occurred.

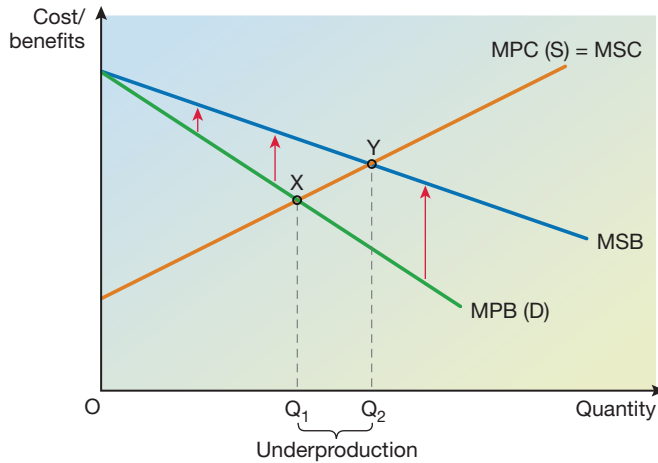


Figure 33.1 Positive externalities in consumption

If a product provides **negative externalities in consumption**, such as an alcoholic acting aggressively towards strangers, then the MPB overestimates its true economic value and so economists deduct the negative externality from the MPB line to get marginal social benefit (MSB). The MSB measures the net benefit to both consumers and those affected externally by the product.

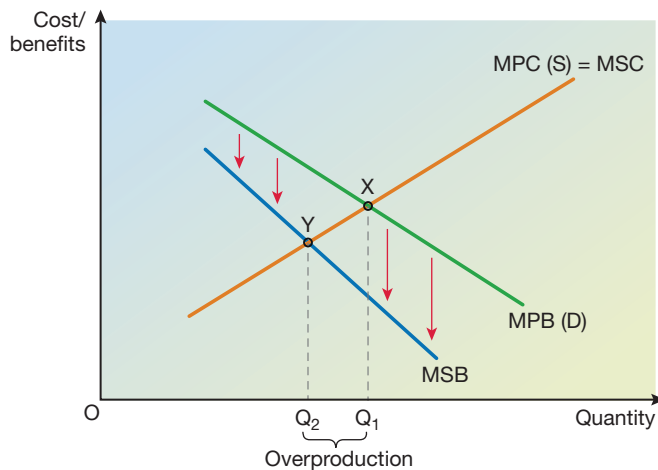


Figure 33.2 Negative externalities in consumption

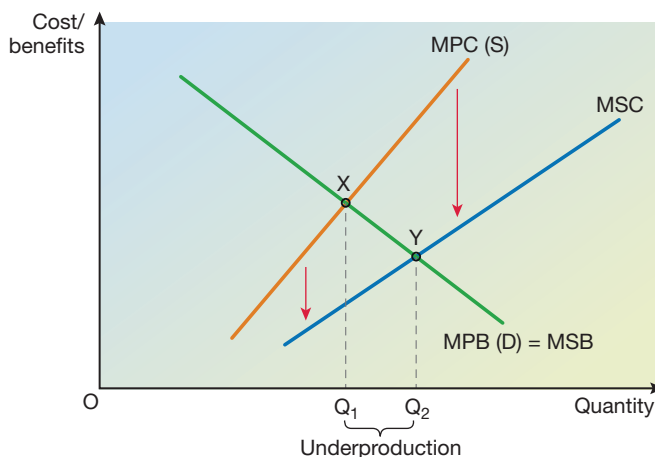
In Figure 33.2, Point X shows the market equilibrium where MPB (demand) and MPC (supply) are in equilibrium. Q_1 is the equilibrium quantity. As the product gives negative externalities of consumption, MSB is lower than MPB, as shown by the red arrows (the vertical distance between MPB and MSB). In this diagram, it is assumed that there are no production externalities and so MPC and MSC are the same. Society's welfare is greatest where MSB = MSC. This is at Point Y. Q_2 is consequently the level of output that maximises economic welfare. The market equilibrium quantity (Q_1) is too high and so the difference between Q_1 and Q_2 is the level of overproduction.

Showing externalities in production

Where there are externalities in production, MPC is adjusted to take into account production externalities.

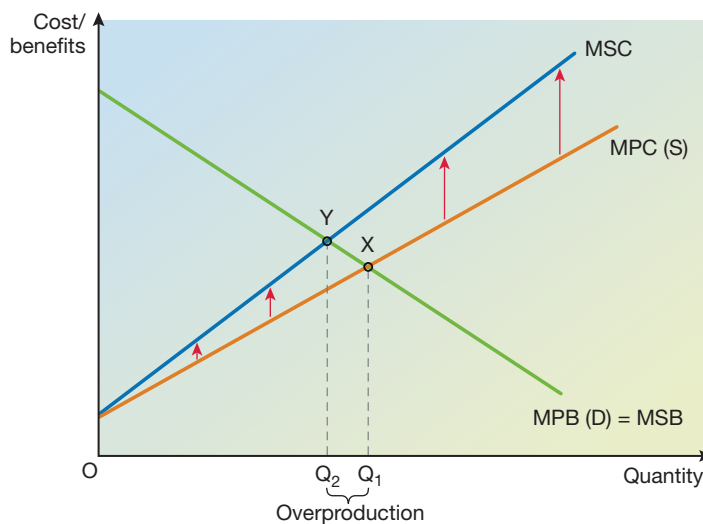
A product may provide **positive externalities in production**, such as a firm changing its materials in order to use renewable resources and thus helping to conserve scarce resources. In this case, the MPC underestimates its true economic value and so economists include the positive externality by amending the MPC line to show a lower level of marginal social cost (MSC). The MSC measures the benefit to both producers and those affected externally by the product.

Figure 33.3 Positive externalities in production



In Figure 33.3, Point X shows the market equilibrium where MPB (demand) and MPC (supply) are in equilibrium. Q_1 is the equilibrium quantity. If there are positive externalities in production, then the MSC (MPC + externalities) moves to the right, like an increase in supply. As the product gives positive externalities in production, MSC is less than MPC, as shown by the red arrows (the vertical distance between MPC and MSC). In this diagram, it is assumed that there are no consumption externalities and so MPB (demand) and MSB are the same. Society's welfare is greatest where MSB = MSC. This is at Point Y. Q_2 is consequently the level of output that maximises economic welfare. The market equilibrium quantity (Q_1) is too low and so the difference between Q_1 and Q_2 is the level of underproduction of this good.

Figure 33.4 Negative externalities in production



A product may provide **negative externalities in production**, such as a firm's factory polluting the local environment. In this case, the MPC overestimates its true economic value and so economists allow for the negative externality by amending the MPC line to show a higher level of marginal social cost (MSC).

In Figure 33.4, $MPB = MSB$ and so there are no consumption externalities. However, costs to society shown by MSC are greater than MPC (as shown by the vertical red lines). This difference represents negative production externalities. The market equilibrium is shown by Point X, and so market equilibrium (where $MPB = MPC$) is Q_1 . However, economic welfare is maximised where $MSB = MSC$. This is at Point Y and so Q_2 is the best level of output for society. The difference between Q_1 and Q_2 is overproduction in this case.

Showing the optimum output of merit and demerit goods, using MSB and MSC

A **merit good** is a good that is underprovided by the market mechanism. An example is health provision. In a free market it has been found that people do not tend to set aside sufficient money to meet their health needs in later life.

A **demerit good** is a good that is overprovided by the market mechanism. An example is alcohol, as people do not always fully recognise its adverse effects.

The diagrams used in the previous section on externalities can also be used to show the optimum output of merit and demerit goods. For merit goods, the MSB will be greater than the MPB (and so more should be produced). This situation is illustrated in Figure 33.1. For demerit goods, the MSB will be less than the MPB (and so less should be produced). This situation is illustrated in Figure 33.2.

REALWORLD ECONOMICS 33.1

Noise pollution

Owners of property in the UK have clear property rights – they are able to sell their property or earn income from it. However, UK law has, in effect, extended the rights of property owners by providing compensation in instances where externalities can threaten the value of the property (and thus the scope for financial gain for the owner).



Should local residents be compensated for noise pollution?

Residents living close to RAF Wittering in Lincolnshire complained about the constant level of noise arising from the training of pilots – on average, 70 flights a day. RAF Wittering argued that the training was for the public benefit, and positive externalities, such as search and rescue missions, meant that on balance the positive externalities to the wider society outweighed the negative externalities suffered by local residents. Furthermore, RAF Wittering argued that residents who bought property after the RAF base had been established could be expected to know that they would experience negative externalities. The price they paid for their property

probably reflected these negative externalities.

The court accepted that there was a public benefit from the training of pilots but concluded that the cost should not be borne by local residents. RAF Wittering was required to pay approximately £1 million in compensation.

Source: UK Environmental Law Association

Discussion points

- 1 Should residents who bought property, knowing that the RAF base existed, be entitled to compensation?
- 2 This article shows how property rights can be extended into noise pollution. In what other ways can property rights be extended?

Review questions

Total: 25 marks

For each situation described in Questions 1–5, identify which diagram shows the cause and extent of market failure.

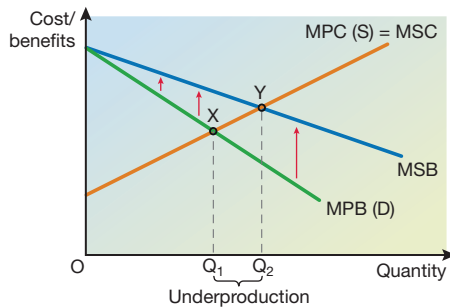


Figure 33.5

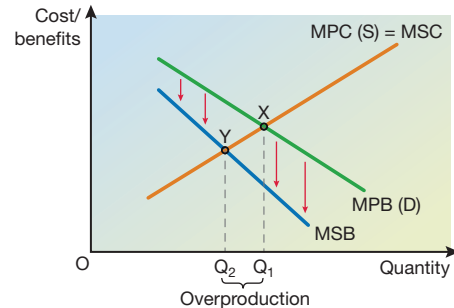


Figure 33.6

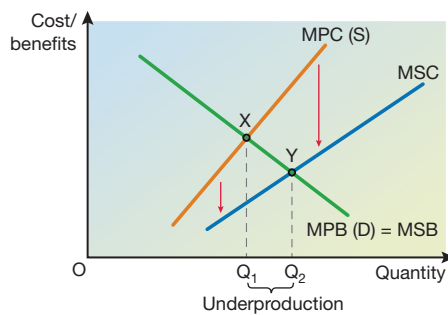


Figure 33.7

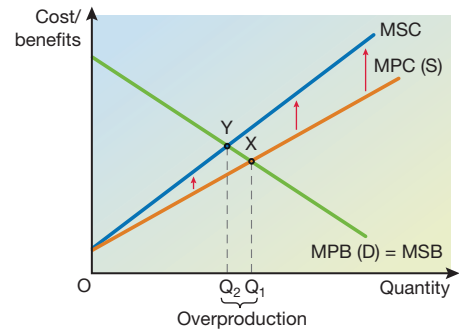


Figure 33.8

- 1 Customers of a takeaway throw away empty cartons in the high street. (1 mark)
- 2 Health and safety training for workers at a local factory leads to employees saving the lives of some local residents involved in a car accident. (1 mark)
- 3 The noise level from the nightshift of a local factory leads to a lack of sleep for local residents. (1 mark)
- 4 A large firm offers local residents free access to its sports facilities. (1 mark)
- 5 A high take-up of flu vaccination reduces the risk of catching flu for people who were not vaccinated. (1 mark)
- 6 Which one of the following is *not* an example of ‘the tragedy of the commons’?
 - A Overcrowding of beaches
 - B Overfishing of coastal waters
 - C Overgrazing of common land
 - D Overuse of fields at the Glastonbury Festival (1 mark)
- 7 Complete the formula: $MPB + ? = MSB$ (1 mark)
- 8 Explain the meaning of ‘the tragedy of the commons’. (6 marks)
- 9 Why might an absence of property rights lead to market failure? (5 marks)
- 10 Explain, with a diagram, how a merit good can lead to market failure. (7 marks)

Competition policy

Key concepts from Year 1

No prior knowledge from Year 1 is needed for this chapter.

This chapter examines competition policy. It studies the general principles of both UK and EU competition policy and assesses the benefits and costs (problems) of competition policy.

The general principles of UK competition policy

In April 2011, the Department for Business, Innovation and Skills published a paper entitled 'Principles for Economic Regulation'. It reiterated earlier approaches to competition policy by reaffirming that the key principles of such a policy are:

- to protect the interests of consumers by ensuring that markets are competitive;
- to liberalise markets and encourage competition in order to improve the efficiency of markets and thus benefit both consumers and UK firms;
- to regulate markets where there are restrictions on competition in order to restrict the prices charged by dominant companies, while ensuring that those prices promote efficiency and provide a reasonable return on the firm's assets;
- to restrict anti-competitive agreements, such as members of a cartel dividing up a market;
- to ensure that regulatory measures do not inhibit dynamic efficiency and the development of vital infrastructure for the UK economy, particularly in the regulated sectors such as telecoms, water, energy and rail.

Overall, competition policy is about applying rules to make sure businesses compete fairly with each other. Competition policy helps consumers by:

- encouraging enterprise and efficiency;
- creating a wider choice of variations of goods and services for consumers;
- trying to reduce prices;
- helping to improve quality;
- encouraging innovation and invention.

The government department responsible for the promotion of competition is the Department for Business, Innovation and Skills (BIS). Its objectives are:

- to promote free and fair markets, for the benefit of consumers;
- to increase productivity, innovation and invention, and labour skills;
- to encourage the commercial exploitation of innovation and invention;
- to improve the performance of UK industry, particularly in regions that suffer from low productivity.

Key term

Competition policy is an attempt by government to ensure that markets operate more efficiently, competitively and in the interest of consumers. Usually the interest of consumers is the paramount objective.

Where markets do not operate freely and competitively, regulatory bodies are responsible for ensuring that restrictions on competition do not disadvantage consumers and the potential for growth in the economy.

The main regulatory body in the UK is the Competition and Markets Authority (CMA). This regulatory body came into operation on 1 April 2014. It took over duties from the Office of Fair Trading (which supervised consumer protection) and the Competition Commission (which investigated potential mergers and made recommendations to government on whether those mergers should be allowed, and any conditions for allowing mergers that were accepted).

In the UK, there are separate regulatory bodies for privatised industries. These regulatory bodies, such as Ofwat (water), have certain powers of control over the industry – primarily the power to set prices or advise government on the need to restrict prices.

Customers are expected to be alert to possible exploitation by firms; the original principle was ‘let the buyer beware’. If a fault with a product is obvious on purchase, the consumer may not claim compensation if they bought it knowingly. However, with many goods being packaged, and the quality being impossible to assess at the time of purchase, legislation has been introduced to protect consumers. Examples include:

- weights and measures;
- misrepresentation of goods and services;
- false advertising;
- complex contracts and conditions;
- misleading financial statements.

In order to ensure that markets operate effectively, ground rules have been established on where monopoly power is being abused. Firms can be investigated if:

- a cartel appears to exist, with firms agreeing not to compete;
- there is an abuse of monopoly power (defined in the UK as a market share in excess of 25%);
- there is evidence of unfair competition, such as misuse of asymmetric information;
- supply is restricted in order to boost price;
- a firm benefits from state aid to the detriment of its competitors;
- a merger is likely to lead to a substantial lessening of competition (SLC).

European Union competition policy

The European Commission enforces the European Union’s competition policy. Understandably, it has similar objectives to the UK, seeking to encourage competitive markets through:

- examining cartels that restrict competition;
- monitoring situations where market leaders abuse their dominance;
- preventing mergers, where they might not act in consumers’ interests;
- opening up markets to more competition.

Additionally, the EU monitors state support for industries to ensure it is not creating unfair competition. Recently the commission ruled that the Netherlands and

Luxembourg had agreed unlawful tax deals with Starbucks and Fiat Chrysler, giving the two companies an unfair competitive advantage as a result of ‘state aid’.

European Union priority

European Union (EU) law takes precedence over UK law in the field of competition policy.

- **EU regulations** are established by the European Union and act as laws in member countries. Member countries must comply with them.
- **EU directives** are instructions from the European Union to member countries. Each member country must pass laws to follow these instructions. Directives are intended to ensure that each country has the same legislation.

The benefits and costs of competition policies

Benefits

The benefits of competition policy are described in the principles of competition policy. If competition policy is successful in achieving its aims, then the following benefits will accrue:

- **Lower prices for consumers:** If competition policy succeeds in liberalising markets, then competitive forces will drive down prices because firms will only be making normal profits. Where markets are uncompetitive, firms can increase their prices and make supernormal profits. Regulatory bodies can investigate such activities and take action to prevent excessive prices and profits, so that even where markets are not liberalised, prices are kept low.
- **Greater choice:** The model of perfectly competitive markets is theoretical rather than realistic. In practice, many firms in competitive markets operate through non-price competition. This leads to product differentiation and thus a much wider choice of products for consumers. Since different consumers have different preferences, this wider choice enables consumers to satisfy their wants more precisely.
- **Better quality:** In imperfectly competitive markets, quality is one of the most significant ways of differentiating products. In order to remain competitive, firms must prioritise this aspect of the products and so consumers benefit from better quality products. In addition, liberalising imperfect markets creates a culture of dynamic efficiency. Thus, firms will strive to introduce new ideas and use new technology to improve their products. Ultimately the consumer benefits from the resulting higher quality products.
- **Innovation and invention:** An effective competition policy is one that balances the needs of consumers, primarily through low prices, with the needs of firms to achieve an appropriate return on their capital employed. By ensuring that regulations allow firms to make reasonable financial returns, the firms have sufficient profit to invest in new technology and research and development in order to improve their processes and invent new products to satisfy (and possibly create) new customer wants.
- **Greater competitiveness in global markets:** Successful competition policy will lead to more efficient firms. This enables UK firms to compete with competitors outside the EU. Likewise, EU competition policy is designed to help European firms compete with the rest of the world.

REALWORLD ECONOMICS 34.1

Is West Ham United's use of the Olympic Stadium anti-competitive?

After the London Olympics, the London Legacy Development Corporation (LLDC) took over responsibility for the Olympic Stadium. The Olympic bid had promised a legacy in the form of facilities that would continue to be used for sport. Many Olympic stadiums in other countries have fallen into disuse.

Unfortunately, athletics is unable to generate high year-round levels of use for the stadium. After various disputes, the decision was taken to convert the stadium so that it could be used for football, but



The former 2012 Olympic Stadium is being restructured to accommodate West Ham United

with retractable seats that would enable it to be used for other sports – primarily (but not exclusively) athletics. This decision was taken following an agreement that West Ham United would become tenants of the stadium, which will continue to be owned by the LLDC.

The original stadium cost £429 million to build. However, the restructuring of the stadium to suit its new purposes is estimated to cost £272 million. West Ham has paid a £15 million contribution to this restructuring. In return, it will pay an annual rent of £2.5 million for a 99-year lease. Since the stadium is owned by the LLDC, the agreement is that the LLDC (funded by taxpayers) will be responsible for employing ticket checkers, medical staff, security and stewards, and cleaning staff on match days. These costs are estimated to be close to the £2.5 million a year rent, although profits from catering will belong to the LLDC.

Football clubs such as Leyton Orient and Charlton Athletic have objected to the project on the grounds that it is providing an unfair subsidy for West Ham and will threaten their fan base. Other clubs are more concerned about the financial advantages this will give West Ham in activities such as the transfer market. The European Commission has already set up investigations into similar arrangements for football clubs in Spain and Holland, but it has indicated that it does 'not at this stage intend to pursue the matter further' in respect to West Ham's use of the Olympic Stadium.

Discussion point

The central principle of competition policy is that it should protect the interests of consumers. Is the agreement between West Ham and the LLDC against the interests of consumers? Are there other grounds for arguing that it is anti-competitive?

Costs

Some criticisms of competition policy are that:

- **It creates an expensive bureaucracy:** The work involved in selecting suitable situations to investigate, obtaining and interpreting the necessary data, and then drawing a well-reasoned evaluation of the situation requires considerable skill. As a consequence, specialist labour must be employed. The CMA has 700 staff and a budget of over £60 million. However, it could be argued that this aspect of government spending is an efficient use of taxpayers' money. The CMA has a target to achieve £10 of direct consumer benefits for every £1 of cost to the taxpayer. In its 2015 report, it recorded benefits of £11 for every £1 spent.
- **Relatively few mergers are fully investigated:** In the light of the data above, it could be argued that this is a more serious problem. In 2015, the CMA reviewed 83 mergers. Of these, only 10 were deemed likely to cause a substantial lessening of competitiveness (SLC).

- **Collusion between firms is difficult to prove:** Since a cartel is an illegal organisation, its existence is unlikely to lead to formal records by the firms involved. Consequently, it can be very difficult to prove collusion, even where it does exist. Furthermore, the initial outcome of collusion is often the fact that all firms charge the same price – this is exactly the outcome that would be expected in a perfectly competitive market. The possibility of collusion can often be detected after its existence because firms charging the same price may all reveal supernormal profits in their accounts at the end of the year, but even then it is difficult to be certain that these profits were the result of collusion between firms.
- **Firms often find loopholes in legislation:** Consumer protection depends on legislation, but its central premise is ‘unfair to consumers’. Unfairness is a normative concept and it therefore relies on a personal judgement. Legislation relating to consumer protection is hence often couched in normative language. Regulators cannot be certain that judges will find in their favour if a dispute goes to court.
- **Regulatory capture reduces the effectiveness of policies.**

Key term

Regulatory capture exists when a regulated industry is able to gain influence over the regulator, so that regulators support the interests of the industry rather than meeting the needs of the public interest that they (the regulators) are meant to serve. (Regulatory capture is explained in detail in the next chapter.)

Review questions

Total: 25 marks

- 1 Which *one* of the following statements is *not* a general principle of competition policy? Competition policy should:
 - A Encourage competition
 - B Encourage dynamic efficiency
 - C Prevent mergers
 - D Protect the interests of consumers

(1 mark)
- 2 A monopoly in the UK is defined as a firm with a market share of at least:
 - A 20%
 - B 25%
 - C 50%
 - D 100%

(1 mark)
- 3 Define the term ‘competition policy’.

(3 marks)
- 4 Explain *one* way in which consumers may benefit from competition policy.

(4 marks)
- 5 Explain *two* ways in which consumers can be protected from exploitation.

(8 marks)
- 6 Explain *two* possible problems arising from the implementation of competition policy.

(8 marks)

Public ownership, privatisation, regulation & deregulation of markets

Key concepts from Year 1

No prior knowledge from Year 1 is needed for this chapter.

This chapter examines the relative merits of public ownership and privatisation and the extent to which the regulation of markets is required. It outlines the arguments for and against: the public ownership of firms and industries; the privatisation of state-owned enterprises; the regulation of markets; and the deregulation of markets. The chapter concludes with a brief study of the problem of regulatory capture.

The arguments for and against the public ownership of firms and industries

Where public ownership involves a firm or an industry, it is often described as 'nationalisation'. In the UK, many industries were nationalised in the postwar period of 1945 to 1950. By the end of that period, the Bank of England, coal, railways, steel, electricity, gas and water supply were in public ownership. Since the beginning of the 1980s, all of these industries have been privatised, except the Bank of England.

The arguments in favour

Some of the reasons for nationalisation are as follows:

- **To reduce costs by producing on a very large scale:** Since nationalised industries produce on such a large scale, they are able to benefit from internal economies of scale to a much greater extent.
- **To prevent existing monopolies from exploiting consumers:** Most of the companies that were nationalised were either natural monopolies or operated as local monopolies. Given the nature of many of the goods and services that they provided, and the fact that essential services such as utilities were natural monopolies, it was argued that these necessary services should serve the needs of the consumers rather than be a source of profit for owners.
- **To make decisions that consider the needs of all stakeholder groups:** Firms tend to focus on the needs of shareholders and profit is therefore the main objective. This means that the needs of other stakeholders, such as the workers, consumers and local community, tend to be ignored. Since government needs to remain popular, it was felt that public ownership would meet stakeholder needs better.
- **To ensure that externalities are considered in decision making:** Private sector

Key term

Public ownership is a situation in which the government owns property, a firm or an industry.

organisations consider only private benefits and costs when making decisions, which influences the allocation of resources. Direct ownership of a business gives the government a much greater scope to ensure that externalities are considered in order to achieve an optimal allocation of resources.

- **To enable the country to achieve greater equality:** Most governments aim to achieve greater equality and, in the wake of the Beveridge Report published in 1942, this has become an official government macroeconomic aim. Direct ownership of enterprises allows the government to ensure greater equality of pay and, because publicly owned enterprises aim to break even rather than make a profit, public ownership should lead to a reduction in the price paid by consumers.

The arguments against

In practice, nationalised industries do not always enjoy the advantages outlined above. Some of the arguments against public ownership are as follows:

- **Diseconomies of scale often dominate economies of scale:** The scale of many nationalised industries was huge. For example, the National Coal Board employed over 700,000 people when it was first nationalised. This led to difficulties of communication and coordination that reduced productive efficiency in the industry.
- **Low levels of dynamic efficiency:** Nationalised industries were expected to provide dynamic efficiency because they would have the financial backing of the government for new investment and could afford to employ the best quality managers. However, in practice managers were not always given the freedom to manage. For political reasons, governments were reluctant to reduce employment in the nationalised industries. Dynamic efficiency was also lacking because there was no profit motive. Since the financial target of the nationalised industries was to achieve breakeven, there was little incentive to take risks in order to achieve higher revenue.
- **Political aims and interference:** The priority given to the wider interests of stakeholders, rather than shareholders, was seen to be an advantage of nationalised industries. However, in some cases the government exerted political influence to make sure that their wishes became the main priorities of the industry. This often manifested itself in protecting workers from unemployment. For example, nationalised industries, such as electricity, were required to purchase coal from the National Coal Board in order to protect the jobs of coal miners. Consequently, electricity was produced at much higher cost than necessary and so consumers of electricity were, in effect, subsidising the coal industry.
- **Financial losses:** Many of the nationalised industries were making losses before nationalisation and continued to do so after nationalisation. These losses were partly because they were declining industries and partly because of attempts to maintain employment levels in these industries. This meant that the taxpayer was subsidising these industries. This could be deemed to be both market failure, because the industry was overproducing, and government failure, because government action was compounding this problem.
- **Limited choice and low quality:** Since the nationalised industries were monopolies, they did not tend to provide choice for customers. Quality was also inhibited because of the absence of competition. This lack of choice and quality meant that allocative efficiency was not achieved.

The arguments for and against the privatisation of state-owned enterprises

Key term

Privatisation is the transfer of ownership of a business from government (the public sector) to the private sector.

Since 1979, most of the former nationalised industries have been privatised. The arguments for and against privatisation are summarised below. Since privatisation is the opposite of nationalisation, in most cases the advantages and disadvantages of privatisation are the reverse of the arguments for and against nationalisation.

The arguments in favour

- **There is greater incentive to improve efficiency and cut costs:** Where industries are privatised, competition is introduced wherever possible. Whereas the nationalised industries had little incentive to cut costs, because losses were subsidised by the government, the privatised firms need to make profit for their shareholders. Consequently, there is a drive to improve productive efficiency by cutting costs.
- **There is more innovation, leading to greater choice and quality:** In a competitive market there is more incentive for dynamic efficiency, as profits are made through product differentiation and new inventions and innovation.
- **Private firms find it easier to access finance:** Under nationalisation, governments are reluctant to borrow money to provide investment because the initial investment requires an increase in taxation. In order to avoid political unpopularity, it is easier to neglect investment in these industries.
- **There is less chance of politically motivated business decisions:** The extent to which some nationalised industries were driven by political decisions can be indicated by their demise after privatisation. Since privatisation, the coal and steel industries in the UK have virtually disappeared, with each having fewer than 10,000 employees. Although this has caused considerable difficulty to the individuals affected by this decline, it does demonstrate the extent of market failure and the high level of misallocation of resources that existed in these industries.
- **A boost to government revenue:** In the short run, privatisation provides extra revenue for government. This enables the government either to finance additional government expenditure or to cut taxation. However, the opportunity cost of the sale of these industries is the loss of an asset and a lessening of the government's direct influence on the economy.
- **Wider share ownership and entrepreneurship are encouraged:** By selling shares to individual members of the public, there is a significant increase in the number of UK residents directly owning shares in businesses. Since choice of suppliers is no longer restricted as an element of privatisation, there are more opportunities for small businesses to supply their goods to privatised firms.

The arguments against

- **Most privatised firms have monopoly power and consumers can suffer:** Despite efforts to create competitive environments for privatised firms, in many cases consumers have no real choice. Consequently, privatised firms are able to use this monopoly power to earn supernormal profits, and the imperfections in the market allow these supernormal profits to be maintained in the long term.
- **Externalities are less likely to be considered:** Since privatised firms are likely to seek profit, their decisions will be based on private costs and private benefits. Consequently, market failure will occur if there are externalities.

- **Nationalised industries are often sold too cheaply in order to attract buyers:** A common criticism of the process of privatisation has been the selling price set by government for shares in privatised firms. Recently the National Audit Office has questioned the price obtained for shares in Royal Mail.
- **Government has less direct control of the economy:** The decline of the public sector has made it more difficult for government to influence the macroeconomy directly in order to achieve aims such as low inflation (energy prices, for example, were sometimes held down in order to keep inflation low). Currently, the government is expecting the former nationalised industries to be the main drivers of investment in infrastructure over the next few years. However, because these firms are privately owned, the government has no direct impact on most of the UK's infrastructure, such as railways, gas and electricity.
- **Short-termism:** A frequent criticism of the UK economy is the short-term approach taken by many firms and organisations. This involves making decisions that are likely to lead to a short-term profit, possibly at the expense of the long-term efficiency of the business. This short-termism tends to lead to a lack of dynamic efficiency because it discourages activity, such as research and development and invention, which tend to take a long time to provide financial returns. Under nationalisation, government was more likely to take a long-term view and therefore invest in projects that would take a long time to pay for themselves. The process of privatisation was also seen as short-termism. The proceeds of privatisation were used to cut taxes. A central principle of finance is that capital should not be used to fund everyday expenses (for example, an individual should not sell their house in order to enjoy a holiday or buy food). However, this is exactly what happened with privatisation and so the long-term effect has been that the state now has fewer assets, since all of the financial gains from the sale of those assets were consumed in the year in which the privatisation took place.

The arguments for and against the regulation of markets

Regulations can apply to the public sector, such as a requirement for a bidding process when offering building contracts to private sector firms, and to the private sector, such as the licensing of taxi drivers and financial advisers. There are a variety of purposes for regulation, which are summarised in the section below.

The arguments in favour

- **To safeguard public interests:** Views on 'public interest' require a value judgement. Consequently, the perception of a need for regulation can vary over time. For example, up until the 1970s, regulations required banks to hold certain levels of cash and liquid assets in order to restrict the chances of them suffering liquidation. The removal of these regulations enabled banks to take more risks. This led to increased profits but was instrumental in creating the banking crisis of 2008. Regulations have subsequently been reintroduced in order to ensure that banks have sufficient assets in the form of cash or liquid assets.
- **For the safety of consumers:** Regulation of production processes, such as food processing, and the labelling of products are intended to enable consumers to assess whether a product is safe or desirable for them to consume.
- **For workers' safety:** Many regulations apply to the workplace and are intended to safeguard people in their place of employment. Without such regulations there is

Key term

Regulation is the imposition of rules by government in order to modify or influence the economic behaviour of organisations and individuals.

a danger that people may be injured as a consequence of unsafe working practices, as unscrupulous employers may try to cut costs.

- **To enable consumers to make rational decisions:** It can be difficult for consumers to compare different products and their value for money. Regulations, such as the need to display weights and measures and unit pricing, enable consumers to make more rational decisions when purchasing goods.
- **To protect quality:** In industries such as television and radio broadcasting, regulations are used to protect the position of the BBC. This is designed to ensure that the quality of programming is not diluted by competition based on low cost. It is also aimed to ensure that programmes are provided for minority interests.
- **To regulate the level of supply:** The transport industry, including taxi services, is regulated to ensure that supply meets demand. The aim is to prevent excess supply. That would drive down prices and force suppliers out of the market, leaving customers with a limited choice.
- **To protect certain interests:** In the past, some regulations were intended to ensure the wellbeing of workers or other government enterprises. Thus, council employees undertook local government projects, such as the maintenance of parks and flowerbeds.

The arguments against

- **Regulations can reduce productive efficiency:** Regulations, such as the guarantee of work for council employees, encourage inefficiency. In the case of council employees, the lack of competition led to inefficiency, as they knew that they would be employed for the job even if it were not conducted efficiently. In general, the target was to provide a satisfactory job.
- **Increased costs:** Costs will rise because of inefficiency but also because regulations need a bureaucracy to enforce them. For example, regulations of taxi services require licensing in order to create the ideal supply. However, a freely operating market mechanism should ensure that the number of taxis in an area matches the demand. It could be argued that the market mechanism will adjust to the required supply much more effectively than a council employee trying to assess the number of taxis needed in an area.
- **Regulations provide barriers to entry:** Regulations, such as licensing of premises, provide a barrier to entry and therefore are likely to restrict supply and thus increase the price of a good or service that is regulated.

The arguments for and against the deregulation of markets

Deregulation allows more goods and services to be provided using the market mechanism. This process is often known as the **liberalisation** of markets. It involves the government removing controls that restrict competition, in order to improve efficiency. Deregulation can apply to industries that were previously nationalised. However, it occurs most frequently in the provision of services to certain activities that remain in the public sector. For example, although there was widespread support for the privatisation of nationalised industries, public support for government ownership of services such as health and education has remained strong. In order to provide elements of free market behaviour in services such as health and education, the government has introduced deregulation through removing the provision of some

Key term

Deregulation is the process of removing restrictions and/or regulations that hinder free competition in the supply of goods and services.

services from the state. Thus, in both schools and hospitals, support services, such as catering and cleaning, are subcontracted to private firms. Healthcare services are also frequently provided by the private sector.

In order to encourage efficiency through deregulation, the government has introduced schemes, such as:

- **Competitive tendering:** Services such as cleaning and catering are contracted out to private enterprise. The private firm that presents the best (often cheapest) bid will provide the service.
- **Internal markets:** In this instance, different providers within the public sector compete to provide a good or service. In the NHS, certain specialist provision is provided across a wider area by a particular NHS trust. The trust that provides a specialist service has earned the right through a tendering process. This process is designed to award the contract to the most efficient bidder, based on criteria such as low cost and high quality.
- **Public–private partnerships (PPPs):** PPPs usually take the form of a private sector firm building and maintaining infrastructure – for example, a new hospital, which is then leased to the NHS trust. Many schools are also funded by public–private partnerships. In this case a private firm builds and maintains the school. In return, an annual payment is made to the private firm.

The arguments in favour

- **Encouraging enterprise:** By opening up many small services to the private sector, deregulation has encouraged the growth of entrepreneurship in the UK, which has more than doubled since 1980.
- **Lower costs/prices:** By using the price mechanism to decide on which firms provide a particular service, greater competition is introduced and the price of deregulated services is reduced. Where these services are provided for public sector organisations, such as the NHS, it leads to lower costs and a more efficient use of taxpayers' money.
- **Removal of bureaucracy:** Some regulations require considerable expenditure on enforcement. Deregulation reduces bureaucracy.
- **Removal of barriers to entry:** Where regulations act as a barrier to entry, such as licensing requirements, they restrict the contestability of markets by limiting the scope for new firms to enter. Thus, deregulation provides more contestable markets.

The arguments against

- **Increased risk to consumers:** Where regulations are intended to protect consumers from unsafe products and services or from exploitation, deregulation can lead to greater risk for consumers. For example, taxi licensing helps to ensure that passengers are safeguarded. Deregulation of such activities can lead to less safety for consumers. This also applies to the safety of employees in their workplaces.
- **Less control of quality:** Where services are provided by outside organisations, it can be difficult for the service provider to guarantee the quality. This issue is compounded by the fact that there is an incentive for the outside provider to cut costs in order to make profit.

REALWORLD ECONOMICS 35.1

Should there be more regulation of banking?

Until 1971, commercial banks in the United Kingdom were required to hold 8% of their assets as cash and a total of 28% as liquid assets (cash and assets that could be turned into cash quickly, such as overnight loans and Treasury bills).

Deregulation in the 1970s and 1980s reduced or eliminated requirements to hold liquid assets, so that at the time of the banking crash in 2008 banks held only 0.2% of their assets as cash.

These changes were introduced to provide greater competition in the banking industry, which had a reputation for a lack of competition in respect to competitive features such as interest rates. The changes were also intended to overcome restrictions that made it difficult for UK banks to compete internationally. As payment by credit card became more widely accepted, fewer customers withdrew cash and the 8% cash requirement became outdated.

Over the next 40 years, the nature of banking changed. From being a risk-averse industry with steady predictable profits, it became

much more dynamic and high risk. This changed the shareholder base from people seeking security to shareholders wishing to make high short-term profits.

High-risk assets, such as subprime mortgages, paid high interest rates and were therefore very profitable for the banks. However, the recession led to a fall in house prices and many of these high-risk mortgage holders defaulted on their payments. As the security for these loans was the property, the fall in property prices meant that the security could not cover the value of the loan.

In 2008, the banking crisis occurred because the banks were unable to pay their debts due to a lack of cash and liquid assets. This led to the government providing loans and then buying shares in order to inject capital into the banks. As a result, the government ended up injecting £20 billion into Lloyds Bank in return for a 43% shareholding, and £46 billion into Royal Bank of Scotland (RBS) in return for a 73% shareholding. Two building societies were also rescued

– Bradford & Bingley, whose ownership was eventually passed to Santander, and Northern Rock, which was nationalised. Northern Rock was eventually returned to the private sector when it was sold to Virgin Money, for £747 million. The government had injected £3.4 billion to save it.

The negative publicity about this loss has led to a government commitment to ensure that the taxpayer does not subsidise the purchase of Lloyds Bank and RBS shares. As RBS share prices have remained low, the government has maintained its RBS holding, except for 5.4% of the shares, which were sold in August 2015. Its shareholding in Lloyds Bank is now just over 10%.

Meetings of the G-20 group of countries have led to the creation of a Financial Stability Board. Its main aim is to end the problem of banks that are 'too big to fail'. Stricter regulations are being introduced to prevent excessive growth by individual banks, and existing large banks must possess capital buffer requirements according to their



The Royal Bank of Scotland and Lloyds – 'too big to fail'

global systematic importance. All banks must hold a 7% capital buffer to ensure that they have reserves available during a crisis, and large banks must possess an additional capital buffer according to their global systematic importance. Class 4 banks, such as HSBC, must hold an extra 2.5% (9.5% in all).

Critics in the UK argue that these restrictions are preventing bank lending, particularly to

small firms, which are seen to be a higher risk. A campaign to establish a publicly owned bank, geared particularly towards the needs of small firms, was successful. This led to the formation of the British Business Bank in 2014, but its primary function is to advise small firms where to obtain finance – it cannot directly provide the finance needed on its own.

Exercises

Total: 25 marks

- 1 Explain one advantage of regulating banking and one advantage of deregulating banking. (8 marks)
- 2 Should government regulate to prevent any firm from being 'too big to fail'? (5 marks)
- 3 Assess the case for public ownership of a commercial bank. (12 marks)

The problem of regulatory capture

The process of privatisation is intended to reduce government control and bureaucracy. Ironically, since most privatised firms have significant monopoly power, government has needed to set up a number of regulatory bodies to monitor their behaviour. In some ways this has led to an increase in the level of regulation in the UK.

Regulatory bodies in the UK include the following regulators for privatised industries:

- Office of Rail and Road (ORR);
- Office of Gas and Electricity Markets (Ofgem);
- Water Services Regulation Authority (Ofwat);
- Office of Communications (Ofcom).

Regulatory capture tends to occur for two main reasons:

- In order to carry out the role, the regulator needs to have a clear understanding of the working of the industry and must be in regular contact with the leaders of firms in that industry, resulting in close personal relationships. This often means that the regulator is perceived to be an 'insider', with sympathies for the industry that they are meant to be regulating. Consequently, they may find it difficult to prioritise the views of the consumers rather than the supplier.
- Regulators need data on which to assess the activities of their industry. This requires the cooperation of members of the industry. Since consumers will only have limited information – often an individual experience of one event – the majority of information on which an assessment will be made is provided by the industry. This example of asymmetric information is likely to lead to decisions in favour of the industry.

Key term

Regulatory capture

occurs when agencies that have been formed by government to monitor and regulate an industry are 'captured' by the industry and start to operate in the interests of the industry rather than the consumers they are meant to protect.

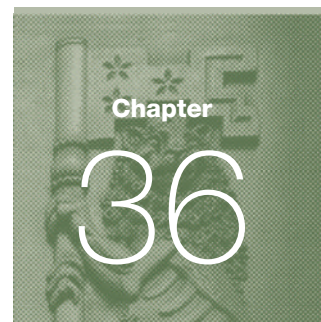
Review questions

Total: 40 marks

- 1 Which *one* of the following organisations is in the public sector?
 - A Bank of England
 - B British Gas
 - C BT
 - D Royal Mail

(1 mark)

- 2** Which *one* of the following is not an example of 'deregulation'?
- A Competitive tendering
 - B Nationalisation
 - C PPPs
 - D Using internal markets
- (1 mark)**
- 3** What is meant by the term 'privatisation'?
- (3 marks)**
- 4** Define the term 'deregulation'.
- (3 marks)**
- 5** Explain *one* way in which 'regulatory capture' might occur.
- (4 marks)**
- 6** Explain *two* disadvantages of privatisation.
- (6 marks)**
- 7** Explain *two* reasons for the regulation of markets.
- (6 marks)**
- 8** Analyse *two* advantages of public ownership.
- (8 marks)**
- 9** Analyse *two* advantages of privatisation.
- (8 marks)**



Government intervention in markets

Key concepts from Year 1

Most of the content of this subject area was covered in Chapter 29 of the Year 1 companion textbook. In order to place the two new elements – government intervention and government failure – in context, it is vital that you revise that chapter before reading this one.

The final two elements of the AQA A-level economics specification on microeconomic issues are government intervention and government failure. These two elements were covered almost in their entirety in the Year 1 book, in Chapter 29 (Government intervention in markets) and Chapter 30 (Government failure).

Table 36.1 shows the A-level content for these two elements. (All of these items were covered in the Year 1 book except ‘the extension of property rights and pollution permits to correct market failure’.)

Government intervention	Government failure
The existence of market failure, in its various forms, provides an argument for government intervention in markets.	Government failure occurs when government intervention in the economy leads to a misallocation of resources.
Governments influence the allocation of resources in a variety of ways, including through public expenditure, taxation and regulation.	Inadequate information, conflicting objectives and administrative costs are possible sources of government failure.
Governments have a range of objectives and these affect how they intervene in a mixed economy to influence the allocation of resources.	Governments may create, rather than remove, market distortions.
The use of indirect taxation, subsidies, price controls, state provision and regulation, the extension of property rights and pollution permits to correct market failure.	Government intervention can lead to unintended consequences.

Table 36.1 *Government intervention and government failure in the AQA A-level economics specification*

Consequently this chapter completes the study of government intervention by studying property rights and pollution permits.

The extension of property rights to correct market failure

In Chapter 33 we saw how the absence of property rights leads to externalities in both production and consumption and hence market failure. In this chapter, the extension of property rights in order to correct market failure will be considered.

Where property rights are clear-cut, such as ownership of a building, there is a strong

Key term

Property rights describe the legal ownership of a resource by an individual or organisation.

financial incentive for the owner to protect their property so that it can be used effectively in economic activity in order to generate income for the owner. However, where there is no clear-cut ownership, such as ownership of rivers, the sea and air, there is no incentive to protect this natural resource. Historically, these properties or assets have suffered considerable neglect and overexploitation. The extension of property rights has been particularly significant in terms of ownership of the sea. With no restriction on access to the sea, resources such as fish were overexploited over a long period of time. This led to a lack of sustainability of many species of fish.

In the European Union, the EU agrees fish quotas. Over a long period of time these quotas have been cut in order to preserve stocks of fish in places such as the North Sea. When the UK joined the European Union in 1973, there were 400 trawlers based in Grimsby. That figure has declined to just five trawlers. Grimsby's main role is processing fish caught elsewhere, particularly by Icelandic ships. However, it could be argued that this has just been the result of market forces – overfishing has led to the almost total elimination of many species of fish in the North Sea. By extending property rights to fishermen so that they catch a certain quota, government has been able to protect this property (the sea) from overexploitation.

In recent years, there have been signs of this extension of property rights reaping benefits. Fish numbers are being replenished and, in December 2015, the EU increased quotas for certain fish. For example, the quota for haddock has been increased by 47% and cod has received a 15% increase in its quota. In specified areas, quotas on plaice, hake and sole have all seen significant increases.

The use of pollution permits to correct market failure

Pollution permits can be given to firms or bought by firms, usually from government. These permits usually measure pollution in weight, such as a tonne of CO₂.

Pollution is a common negative externality that leads to market failure because the market mechanism makes no allowance for this externality. Two approaches that have been used to correct environmental market failure are:

- **regulations**, whereby polluters are fined if they break laws relating to pollution. Although this approach can generate income for government, it also creates bureaucracy. Furthermore, there is no reward for firms that pollute well below the target level of pollution and so it is likely to have limited impact on many businesses;
- **taxation**, where taxes are imposed on pollution (e.g. the landfill tax). This approach has been more effective and acts as a source of revenue for government. It also tends to reward those firms that perform very effectively in the elimination of pollution.

A more recent approach, however, has been to use market forces through the provision of pollution permits and the creation of a market in which pollution permits can be traded.

In order to limit negative externalities and thus improve the allocation of resources, governments often license an industry by giving permits to pollute to individual firms such that the total level of pollution is acceptable. Some firms may find it easier/cheaper to limit their pollution below this level and so they will reduce their pollution levels below their target and sell their pollution permits to other firms. In

Key term

Pollution permits give firms a legal right to create a certain level of pollution over a particular period of time.

**REALWORLD
ECONOMICS** 36.1

The EU Emissions Trading System (EU ETS)

In 2002, the **Kyoto Protocol** led to EU countries such as the UK agreeing to reduce their emissions, between 2008 and 2012, to 87.5% of their 1990 levels. In order to encourage its members to meet this target, the EU established the European Union Emissions Trading System (EU ETS). This system is the world's biggest emissions trading market and accounts for over 75% of international carbon trading.

This system targets high emitting industry sectors, such as power stations and manufacturing plants. The system has been severely criticised by manufacturing companies because households and transport create a large percentage of emissions in the UK and EU, yet their emissions were not covered by the system (government was worried that this would prove to be politically unpopular for households). Aviation is now part of the process, with a requirement to cut emissions by 5% per annum. Overall, around 45% of total EU emissions are limited by the EU ETS.

Firms are given (or pay for) emission allowances. For every tonne of CO₂ emitted, the firm

must give up one of its allowances. Firms that pollute more than their allowance can buy permits from other firms; otherwise they must pay a fine. This provides an incentive for firms to reduce pollution so that they do not have to pay for additional permits or can earn revenue by selling their unused permits to other firms.

The first two stages of the EU ETS, from 2005 to 2012, had limited success. This was because permits were initially given free of charge and too many were allocated. This meant that the price of permits was so low that there was very little incentive for firms to improve their pollution control. The second stage coincided with the recession and, with a fall in economic activity, production levels fell, which meant that although there had been a reduction in allocations, the supply of permits was still too high.

From 2013, a system of payments for permits was introduced, with a view to phasing out free allocations by 2027. Manufacturing industry received 80% of its allowance free of charge in 2013, but this decreases annually until it reaches 30% in

2020. It will then be phased out over the following seven years.

In the third stage of EU ETS, permits are sold through auctions, so that they reflect market value. In order to meet its Kyoto Protocol targets, the supply of permits is reducing by 1.74% every year from 2013. Between 2005 and 2013, the system had reduced emissions by about 1% per annum. This more stringent target is likely to increase the value of pollution permits and therefore encourage firms to find ways of reducing pollution.

Exercises

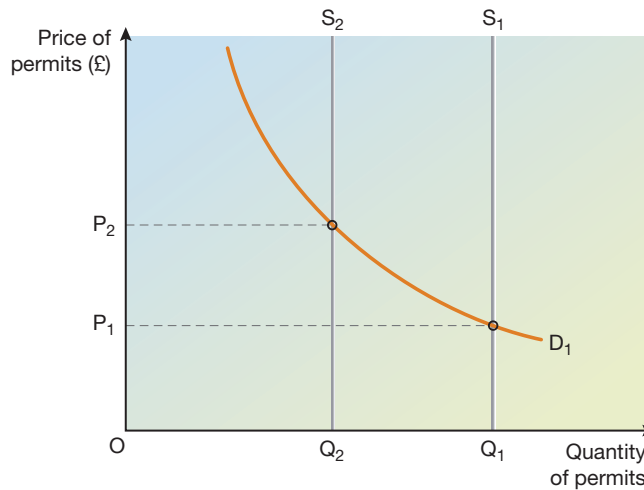
Total: 20 marks

- 1 Explain why the more stringent target for pollution is expected to lead to an increase in the value of pollution permits. **(4 marks)**
- 2 In the UK, some negative environmental externalities are controlled through taxes, such as the landfill tax. Explain one reason for using taxation to control emissions and one reason against using taxation to control emissions. **(8 marks)**
- 3 Explain one strength and one weakness of the EU ETS system of pollution permits. **(8 marks)**

this way, firms that can reduce pollution most cost effectively earn money from those that cannot and so pollution is reduced as cost effectively as possible. This system ensures that every firm has an incentive to reduce pollution.

Pollution permits, once issued, can be more effective in reducing overall costs of pollution control because, unlike regulations that require close monitoring and enforcement, the market mechanism is employed. However, pollution permits tend to be less effective at reducing the overall level of pollution because the number of permits is often quite high. In order to be effective, they should be severely restricted.

Figure 36.1 Pollution permits



In Figure 36.1, D_1 shows the initial demand for pollution permits. S_1 shows the supply of permits from the government, which, as it is fixed, is vertical at Q_1 . Q_1 is the level of permitted pollution. If the government wishes to reduce the level of pollution, it can cut back the supply of permits. For example, if it wishes to restrict the level of pollution to Q_2 , then the line S_2 shows the new supply of pollution permits. This cut in supply forces the price up from P_1 to P_2 . Any firms that cannot or will not pay this price are unable to produce and so the level of pollution falls. In practice, some firms may still pollute above the permitted level because governments will have a system of fines to punish excessive polluters.

Pollution is difficult to control because it is difficult to measure pollution levels from a particular organisation. The use of pollution permits is undermined by the fact that less developed countries tend to have surplus permits that can be bought cheaply by more heavily polluting industrialised nations.

Review questions

Total: 20 marks

- 1 Define the term 'property rights'. (3 marks)
- 2 Define the term 'pollution permits'. (3 marks)
- 3 Explain how the extension of property rights relating to fishing has helped to overcome market failure. (6 marks)
- 4 Analyse, with the help of a diagram, how pollution permits can help to reduce pollution. (8 marks)

Topic 8 Exam-style questions

A-LEVEL PAPER 1

SECTION A Context – Inflation and the energy market

EXTRACT A Inflation and energy prices

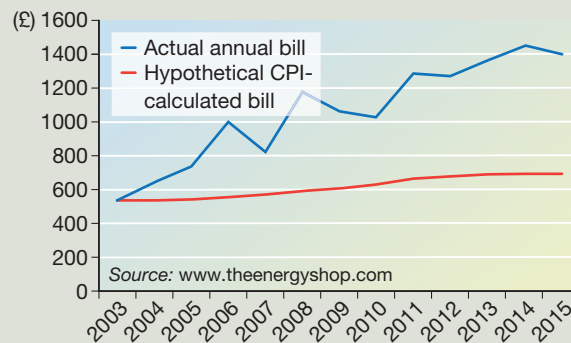


Figure A Actual average annual household bills and hypothetical bills based on the Consumer Prices Index, 2003–15

Year	Retail Prices Index	Index of energy prices
1947	7.3	4.5
1960	12.5	9.4
1970	18.5	15.3
1980	66.8	61.9
1987	100	100

Source: ONS

Note: Base year for indices – 1987

Table A Inflation and energy prices during period of nationalisation

EXTRACT B CMA investigation into the energy market in 2015

In July 2015, the CMA published provisional findings of its investigation into the level of competition in energy markets in the UK. The CMA identified adverse effects on competition arising from the following features of the market:

- inadequate and asymmetric information, particularly amongst domestic customers and small businesses, as a result of barriers to accessing information. This led to these customers being unlikely to switch suppliers, even where there were clear differentials in prices between suppliers;
- a tendency for customers to be on a tariff that was not the cheapest tariff available from that supplier;
- the existence of regulations that tended to inhibit innovation amongst suppliers;
- vertical integration within some suppliers, leading to monopoly power when supplying competitors in retailing gas and electricity to customers;
- a lack of robustness and transparency in actions in the decisions made by the regulator.

Source: Ofgem – Retail Energy Markets in 2015

EXTRACT C **Should the energy suppliers be nationalised?**

According to city analysts, a proposal by Jeremy Corbyn to nationalise energy suppliers would cost the government £185 billion.

The head of Ofgem, Andrew Wright, stated that 'people had a deep mistrust of anything the energy companies do or say', following a consistent trend for energy prices to exceed the rate of inflation (see Extract A, Figure A). Since 2009, the average profit per household made by an energy company has risen from £8 to over £100.

However, Ann Robinson of the uSwitch price comparison website believes that 'nationalisation would stifle innovation and remove incentives to improve customer service and efficiency'. Energy price rises exceeded inflation during the period of nationalisation (see Extract A, Table A).

The level of competition in the gas market indicates that the existing energy companies have monopoly power, particularly when dealing with households and small businesses. For both electricity and gas supply, the four-firm concentration ratio for the small business sector is the same – 66%. However, new entrants are taking a bigger share of the household market.

Energy UK, the body that represents the energy suppliers, argues that there is a massive amount of investment in infrastructure required – around £110 billion in the next seven years. The energy companies are also devoting finances to improving the impact of their activities on the environment. A spokesman stated that 'the only way that it is going to be funded is if the industry is able to make a profit and attract investors'.

However, the government has a stronger credit rating than the big six energy firms and so would be able to borrow money at a 3–4% rate of interest, in comparison to 6% charged to the energy companies. Furthermore, any surpluses made by a nationalised energy company could be put back into making infrastructure and environmental improvements.

It should be noted that the decision would need to be ratified by the European Union. A government-owned monopoly supplier would be seen as a lessening of competition and therefore nationalisation of the whole industry is unlikely to be accepted by the EU.

Sources: www.independent.co.uk/environment and other sources

Questions

Total: 40 marks

- 1 Calculate the RPI rate of inflation over the period 1970–80. (2 marks)
- 2 Explain *one* possible reason why energy price rises have consistently exceeded the rate of inflation between 1947 and the current day. (4 marks)
- 3 With the help of a diagram, explain how the monopoly power of the energy suppliers is likely to lead to market failure. (9 marks)
- 4 Assess the view that energy suppliers in the UK should be nationalised. (25 marks)

SECTION B Essays

Total: 40 marks

The operation of free markets and the price mechanism are the best way of achieving a perfect allocation of resources in an economy. Where market imperfections exist, competition policy can be used in order to improve the competitiveness of markets.

- 1 Analyse the problems of using competition policy to improve markets in order to benefit consumers. (15 marks)
- 2 Assess the view that the price mechanism leads to a perfect allocation of resources. (25 marks)

Section 2

The national & international economy



The first part of the document discusses the importance of maintaining accurate records of all transactions. It emphasizes that every entry, no matter how small, should be recorded to ensure the integrity of the financial data. This includes not only sales and purchases but also expenses and income. The text suggests that a systematic approach to record-keeping is essential for identifying trends and making informed decisions.

Next, the document addresses the issue of budgeting. It explains that a well-defined budget is a critical tool for managing resources and controlling costs. By setting clear financial goals and limits, individuals and organizations can avoid overspending and ensure that their financial objectives are met. The text provides practical advice on how to create a realistic budget that accounts for all potential expenses and income sources.

The third section focuses on the importance of regular financial reviews. It argues that periodic assessments of financial performance are necessary to catch any discrepancies or errors early on. This process involves comparing actual results against the budget and identifying areas where adjustments may be needed. The document stresses that consistent monitoring is key to staying on track and achieving long-term financial success.

Finally, the document concludes by highlighting the value of seeking professional advice when needed. It notes that while many financial tasks can be handled internally, complex situations may require the expertise of accountants or financial planners. Consulting with professionals can provide valuable insights and help navigate challenging financial decisions. The overall message is one of proactive financial management and continuous learning.



Topic 9

The measurement of macroeconomic performance

Uses of national income data

Key concepts from Year 1

This chapter builds on some important economic topics that were developed in Chapters 31–33 of the Year 1 companion textbook.

- Chapter 31 looked at what governments in the UK (and elsewhere) aim to achieve by implementing their macroeconomic policies. These so-called macroeconomic objectives include achieving high (but not unsustainable) rates of economic growth, minimising unemployment, price stability (which in the UK means a rate of inflation of around 2% annually) and a stable balance of payments on current account. We also considered the extent to which conflict might arise when governments attempt to achieve these objectives simultaneously.
- Chapter 32 discussed the methods used by governments to measure the performance of the economy and thus judge the extent of their success in achieving macroeconomic objectives. Relevant methods of measurement include rates of growth of real GDP and real GDP per capita, the Consumer Prices Index (CPI), the rate of unemployment and the balance of the UK current account.
- Economists make extensive use of index numbers to measure macroeconomic and microeconomic performance because the data they handle are frequently large and complex. Chapter 33 covered how index numbers are calculated (including base years and the use of weights), how they are interpreted and their use in measuring economic variables.

This chapter builds on the material covered in Year 1 to consider how national income data can be used to assess changes in living standards over time as well as their limitations when used for this purpose. It also considers the value and limitations of national income data in making comparisons between living standards in different countries and the importance of purchasing power parity (PPP) exchange rates in making such comparisons.

Key terms

National income is the monetary value of the total output of an economy over a specific time period.

Economic growth

occurs when, over time, an economy expands its capacity to produce goods and services.

Measuring national income

National income measures the monetary value of the total production of goods and services produced within an economy over a time period. It is a fundamental measure of the level of activity in an economy. The calculation of national income in the UK provides much important economic data including gross domestic product (GDP) and gross national product (GNP) as well as information on, for example, household saving and disposable income.

There are many millions of transactions taking place within the UK and other economies each day that add to the country's national income. The size and complexity of an economy, such as that of the UK, means that measuring national income accurately is a challenge. We shall consider the difficulties later in this chapter.

How is national income measured?

1 Different methods of measurement

The ONS can use three methods to measure the national income of the UK.

- **The output measure:** This approach measures the total of the value added through the production of goods and services within the economy. The cost of resources used in production and the value of goods and services produced are compared. It also provides the first estimate of GDP as well as data on the contribution of different industries to national income.
- **The income approach:** This measures the total income generated by the production of goods and services within the UK economy. It gives information on different categories of income: for example, income earned by companies and the self-employed.
- **The expenditure approach:** This measures the total expenditure on all finished goods and services produced within the economy by the government, households and firms as well as those based overseas who buy UK goods and services.

Each of these three measures should, with suitable adjustments, give the same figure for the UK's national income.

2 Different measures of the size of an economy

There are a number of measures of the size of an economy. In the Year 1 companion textbook we used the concept of gross domestic product or GDP. This is probably the most commonly used measure of size and, when indicating changes in an economy's size, of economic growth. However, GDP is not the same as national income, even though it is frequently used as a proxy for it. The three different measures of an economy's size are shown in Figure 37.1.

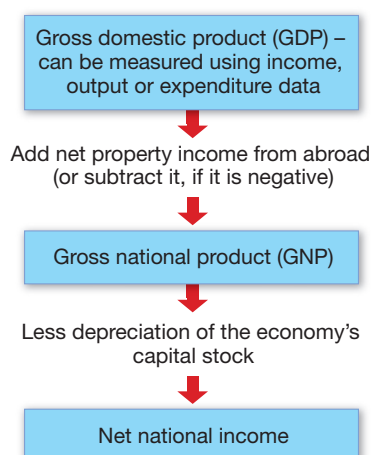


Figure 37.1 *Different measures of an economy's size*

- **Gross domestic product (GDP):** This measures the value of output produced by the resources available within the UK.
- **Gross national product (GNP):** Some of the productive resources in the UK are owned by foreigners (Toyota's factory in Derby, for example) and these result in a flow of income to the owners overseas. Equally, UK businesses and individuals own productive assets based abroad which generate flows of income into the UK. When these two flows are added together, the outcome (which could be a negative

or positive figure) is called ‘net property income from abroad’. When this is added to gross domestic product it gives a figure that is termed gross national product or GNP.

- **National income or net national income:** Over a period of time, an economy’s capital stock (i.e. assets such as its factories, offices, bridges, roads and airports) will decline in value. This decline, which is difficult to measure accurately, is called depreciation. If depreciation is subtracted from gross national product (GNP), the result is net national income or, as it is more commonly called, national income.

Why do governments measure national income?

The level of national income in an economy is measured for a number of reasons.

1 It helps governments to make decisions

Most governments in the world pursue fairly similar macroeconomic objectives, such as price stability and sustainable rates of economic growth. National income data provide governments with important information relating to the output of the economy, such as the rate of growth of GDP. This allows some assessment of the effectiveness of the government’s existing monetary, fiscal and supply-side policies.

National income data are also instrumental in helping the government and monetary authorities (for example, the Bank of England in the UK) to forecast future levels of national income and to design and implement suitable macroeconomic policies for the future. The UK authorities were anxious to see whether their macroeconomic policies (such as reducing the bank rate of interest to 0.5%) were effective and the data for national income would have provided important evidence of this.

2 It helps governments to make judgements about living standards

Governments and other interested parties can make judgements about the living standards of an economy’s citizens using national income data. An increase in a country’s national income per capita in real terms shows average incomes adjusted for the effects of inflation. Increases in this figure have the potential to increase the standard of living of everyone. Data relating to real GDP per capita (or per head) can also be used to illustrate the trend in living standards in an economy over time. Figure 37.2 illustrates the change in the UK’s real GDP per head (based on quarterly data) from Quarter 1 in 1997 to Quarter 2 in 2015. The effect of the 2008–09 recession and its aftermath on average living standards is evident.

Key terms

Gross domestic product (GDP) is a measure of the value of all goods and services produced within an economy over a given time period.

Gross national product (GNP) is the value of all goods and services produced within an economy plus net property income from abroad.

Net property income from abroad is the income earned from overseas assets owned by UK businesses and individuals minus the income from UK assets owned by foreigners. This figure can be negative or positive.

The **standard of living** is the degree of comfort through the consumption of goods and services enjoyed by individuals, households or other groups in society.

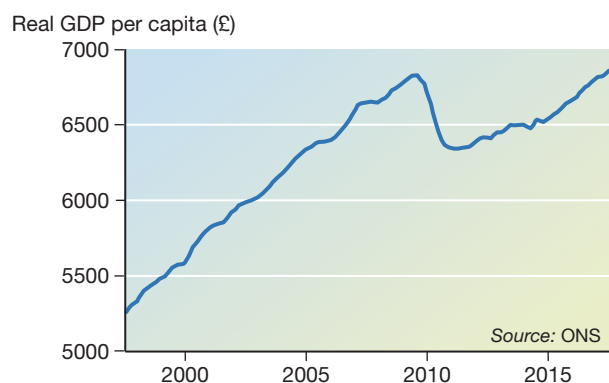


Figure 37.2 *Real GDP per capita for the UK, 1997–2015*

As we shall see later in this chapter, national income data can also be used to make comparisons between living standards in different countries.

The limitations of using national income data to measure living standards

National income data are used by most governments to measure living standards in their countries over periods of time – a process known as trend analysis. However, there are a number of limitations in using national income data in this way.

1 The data ignore the distribution of income

Figure 37.2 illustrates that the average GDP per head in the UK in real terms has risen from about £6400 in 2010 to approximately £6900 in 2015. At first glance this appears to suggest that the living standards of all the UK's inhabitants have risen as the country's income has increased. However, this is an average figure. In reality some groups in society might have enjoyed rapid rises in their incomes, while the real incomes of others may have fallen over this period.

If the distribution of income becomes less equal over a period of time, this means that the country's income is shared less equally between its inhabitants. In such circumstances, a rise in income results in some groups benefiting much more than others.

Figure 37.3 shows the changes in the distribution of income in the UK between 1961 and 2013–14. Since 1977, income distribution in the UK has generally become less equal, although there have been periods when the opposite has been true, such as 2001–04. Therefore it cannot simply be assumed that a rise in average national income per head results in higher living standards for all in an economy.

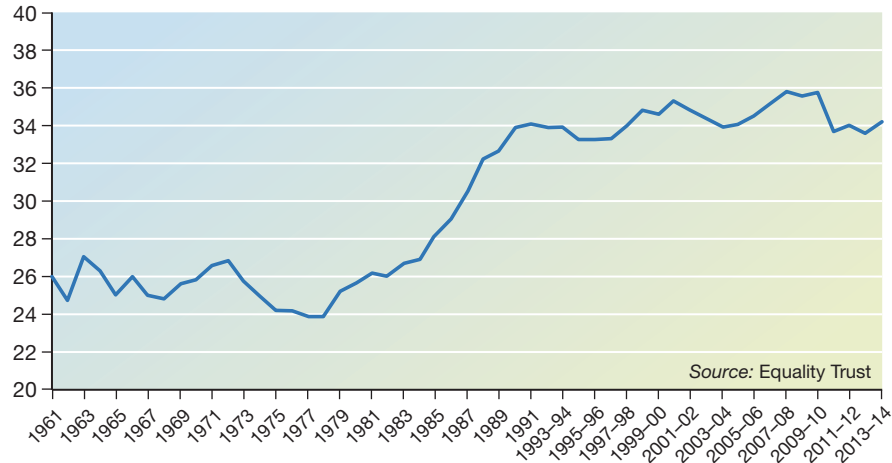
Key terms

The **distribution of income** measures how a nation's total income is spread between that country's inhabitants.

The **shadow (or underground) economy** comprises activities that are not regulated or recorded by the government and can be illegal.

Negative externalities describe the problems experienced by third parties due to an economic activity. These problems can be passed on as a result of the consumption or production of a product.

Figure 37.3 The UK's distribution of income as measured by the Gini coefficient, 1961–2013/14



It should be noted that the Gini coefficient was developed by the Italian statistician Corrado Gini and is a widely used measure of income inequality. In this example, a figure of 0 would indicate a society in which income was perfectly equally distributed (i.e. everyone received the same income). A figure of 100 would represent total inequality (i.e. a single person would receive all of the income in an economy). Thus a rise in the trend of the graph here indicates rising income inequality.

2 The effects of leisure are ignored

A person's standard of living is likely to be strongly influenced by his or her leisure activities – for example, participating in sport, listening to live music or taking holidays. One way of measuring the amount of leisure enjoyed by a country's inhabitants is the length of the average working week. In the UK the average length of the working week for full-time employees has risen, according to the Trades Union Congress (TUC) – an organisation that represents trade unions in the UK. The TUC's research has shown that working hours in the UK are long in comparison to other European countries. This difference would affect the validity of any comparisons between European living standards.

3 Some economic activity is excluded

National income data only measure some of the economic activity that takes place in an economy. For example, national income accounts exclude the value of unpaid work within the economy, such as carers who are not paid for looking after family members or friends, parents who choose to stay at home and care for children, and even the value of DIY that is carried out. Other omissions include some elements of the so-called 'shadow economy'. The shadow (or underground) economy comprises activities that are not regulated by the government and are often illegal, such as prostitution and the production and sale of drugs. Another activity that is not recorded for the purposes of taxation is work done for 'cash in hand', where the recipient avoids paying tax on these earnings.

The Statistical Office of the European Union (Eurostat) required all EU member states to incorporate estimates of voluntary illegal activities into National Accounts estimates by September 2014. The Office for National Statistics (ONS) included estimates for illegal drugs and prostitution in its National Accounts publications from

30 September 2014. However, the ONS recognises that such estimates are unlikely to be accurate and the amounts included in the UK accounts for these purposes are small: less than 1% of the UK's total GDP. It is likely that a country's GDP is still understated in relation to its shadow economy.

4 Negative externalities

National income data record the private costs of production but tend to ignore any external costs (or negative externalities) that are associated with them. For example, the building of Crossrail in London to provide an East–West railway across the city will add significantly to the UK's GDP as a result of both its construction and its use. Its construction has certainly caused noise, congestion and pollution and thus has imposed costs that will not be reflected in the UK's national income accounts. The existence of negative externalities means that GDP can often overstate living standards.

5 Inaccurate data

Collecting national income data for the UK, or any other country, is an immense task and it is inevitable that errors will creep in. This is one reason why GDP figures in the UK are revised, often for several years, as more accurate information comes to hand.

REALWORLD ECONOMICS 37.1

Britons 'working more than 48 hours a week'

The number of people working more than 48 hours a week has increased, highlighting the 'national disgrace' of Britain's long-hours culture, according to a new report.



Long working hours take their toll on this junior doctor

Almost four million employees are working at least 48 hours a week, 350,000 more than a decade ago, despite a European directive aimed at reducing working time.

Managers and professional staff work the longest hours, while one in 25 men are toiling for at least 60 hours a week, according to the TUC report. 'Britain's long-hours culture is a national disgrace,' said TUC General Secretary John Monks. He further commented: 'It leads to stress, ill health and family strains'.

The TUC said its survey showed how badly work was managed in the UK compared with other countries. The average working week in the UK is now 43.6 hours compared with a European average of 40.3 and limits of just 35 in France.

Firms were urged to organise their work more efficiently so that staff 'work smart' and become more productive while working fewer hours. Mr Monks added: 'Half the country is caught in a vicious circle of low pay, low productivity and long hours, with the other half trapped in their offices and battling ever growing in-trays. Other countries produce more, earn more and work far shorter hours. We should, and can, do the same if employers, unions and government work together.'

Source: Daily Mail

Exercise

- 1 Explain the possible reasons why employees working long hours in the UK might not add to the country's GDP. (8 marks)

Using national income data to compare international standards of living

Comparisons between living standards in different countries are made for a variety of reasons. Some international agencies such as the International Monetary Fund (IMF) may use this as the basis for offering financial support to economies. The European Union (EU) will compare living standards as part of its decision making on which countries to support with investment programmes designed, for example, to improve a country's infrastructure. Governments use this data to measure their relative performance and to draw judgements about comparative living standards. Table 37.1 shows the top ten countries in the world as measured by GDP per capita in 2014. This suggests that the people of Qatar have the highest standard of living in the world. However, as we shall see, this might not be the case.

Table 37.1 Top ten countries rated by GDP per capita 2014 and positions 2013 (converted into American dollars)

Rank, 2014	Country	Per capita income (\$), 2014	Per capita income (\$), 2013 (rank)
1	Qatar	97,519	96,719 (2)
2	Norway	97,363	102,832 (1)
3	Macao, SAR, China	96,038	90,332 (3)
4	Australia	61,887	67,473 (7)
5	Denmark	60,634	59,819 (9)
6	Sweden	58,887	60,365 (8)
7	Singapore	56,287	55,980 (7)
8	United States of America	54,630	52,980 (12)
9	Ireland	53,314	50,470 (16)
10	Iceland	52,111	47,549 (19)

Source: World Bank

The limitations of using national income data for comparing living standards

There are many problems in making judgements about the comparative living standards of the inhabitants of different countries on the basis of national income data. Such comparisons are normally made using GDP data. One initial problem is that countries have populations of different size. This can be overcome by calculating GDP per head of population, or GDP per capita, as in Table 37.1 above. However, other difficulties exist that are not so easily resolved.

1 Inaccuracies in calculating national income

We saw earlier that calculating national income is a complex task and difficult to complete accurately. The degree of inaccuracy will vary between countries as slightly different methods of collecting data are used. Furthermore, the size of the shadow economy will differ: from around 10% of GDP in the UK to up to 30% in some economies in Southern Europe. These factors mean that the data used for comparisons of living standards are not the same, reducing the value of any judgements that result.

2 Differences in the length of the working week

Employees in different countries work for different numbers of hours each week. For

example, employees in the UK generally work a larger number of hours than those in many other northern European countries. Working longer hours to generate a given income is likely to reduce a person's standard of living. This difference can invalidate comparisons of living standards using GDP data to some degree.

3 Changes in exchange rates

In order to compare countries' living standards using national income data, it is necessary to convert GDP data into a common currency. The American dollar is frequently used for this purpose, as in Table 37.1. One fundamental weakness of this method is that a change in exchange rates can occur for a variety of reasons (such as a change in another country's interest rates) that have no significant impact on living standards, yet the consequence is that the apparent relative standard of living of a country's inhabitants may alter. However, in reality it may have remained quite stable. Changes in the value of currencies against the dollar would have been a major cause of some of the quite dramatic changes in per capita incomes (and country ranks) shown in Table 37.1.

4 Exchange rates do not reflect relative prices in different countries

Prices of products that are purchased in many countries can vary and this affects the purchasing power of money. Exchange rates are determined by the demand and supply of currencies and are thus unlikely to reflect these differences. So, differences in incomes may not indicate differences in standards of living. One approach to this problem is to base exchange rates on purchasing power parities. We explore this fully in the section below.

The use of purchasing power parity (PPP) exchange rates

The exchange rates that you commonly see quoted in the media or your local bank are determined by the demand for, and supply of, the currency on the international currency markets. For instance, at the time of writing, £1 = €1.35. It is highly unlikely that products that sell for £1 in the UK would be priced at €1.35 in any single country using the euro, and certainly not in all nineteen. Exchange rates tend to be changed by short-term factors, such as changes in macroeconomic performance, and only in the long term might they reflect differences in the cost of a common basket of goods and services. This means that market exchange rates do not indicate differences in the cost of living in countries and thus do not effectively measure differences in living standards.

The weaknesses of market exchange rates for comparing living standards prompted the development of purchasing power parity (PPP) exchange rates. PPP exchange rates are determined using the relative cost of purchasing a common basket of goods and services in two (or more) countries to determine the exchange rate.

For example, the market exchange rate between the American dollar and the pound sterling might be £1 = \$1.50 at a given point in time. It is possible to calculate a PPP exchange rate for the USA and UK by deciding upon a basket of goods and services that are purchased as widely as possible by consumers in both countries. The cost of this basket of goods might be \$84 in the USA and £42 in the UK. This would give a PPP exchange rate of £1 = \$2 and would suggest that the American dollar is overvalued using market exchange rates. If the market exchange rate were used to

Key terms

Market exchange rate is the price of one currency expressed in terms of another as determined by the demand for, and supply of, currencies.

Purchasing power parity (PPP) exchange rates value currencies against one another based upon the relative costs of a common basket of goods and services.

compare living standards in the USA and UK using these figures, it would overstate the living standards of Americans.

PPP does offer many benefits to economists and to others when comparing living standards in different countries. PPP exchange rates tend to be stable over time while market rates are more volatile. As a result, market exchange rates can suggest that substantial changes in comparative living standards are taking place when this is not the case. Market exchange rates are relevant only for internationally traded goods such as cars, oil or rice. Goods and services that are not traded on international markets tend to be cheaper in low-income than in high-income countries. For example, a taxi journey in London is more expensive than one in Hanoi in Vietnam. Any analysis that fails to take into account these differences in the prices of products such as haircuts across countries will underestimate the purchasing power of consumers in emerging market and developing economies. As a result, it will understate their living standards. Thus PPP generally gives a better measure of comparative living standards.

On the other hand, any worthwhile comparison of prices across countries must consider a wide range of goods and services. This is not an easy task, mainly because of the amount of data that must be collected and the complexities of the comparison process. Consumption patterns can differ enormously between countries, especially between developed and developing countries, making it difficult to construct a common basket of goods and services.

PPP rates tend to be of greatest value when comparing developing (or emerging) economies with developed economies. This is because there is the greatest divergence between PPP exchange rates and market exchange rates for these economies, whereas PPP and market rates tend to be more similar for developed nations. It is noticeable in Table 37.2 that many of the countries at the top of the PPP ranking of income per capita were also high on the ranking based on market exchange rates. This is because there is a smaller difference when comparing developed economies and many of these economies are developed.

Table 37.2 *Top ten countries rated by GDP per capita 2014 and positions 2013 (converted into American dollars using PPP)*

Rank, 2014	Country	Per capita income (\$), 2014	Per capita income (\$), 2013 (rank)
1	Qatar	139,760	132,480 (1)
2	Singapore	80,270	77,840 (4)
3	United Arab Emirates	66,270	63,150 (7)
4	Norway	65,970	66,520 (6)
5	Hong Kong SAR, China	56,570	54,380 (10)
6	United States of America	55,860	54,360 (11)
7	Netherlands	47,660	42,260 (13)
8	Germany	46,840	45,020 (17)
9	Sweden	46,710	46,260 (14)
10	Denmark	46,160	45,350 (15)

Source: World Bank

It is, however, noticeable that the income gap (and therefore difference in living standards) between these countries and developing or emerging economies is much smaller using PPP exchange rates. For example, using market exchange rates the GDP per capita in 2014 for Chile was \$14,528; using PPP exchange rates, this increased dramatically to \$21,310.

**REALWORLD
ECONOMICS** 37.2

The Big Mac Index

The Big Mac Index was invented by *The Economist* in 1986 and measures the extent to which market exchange rates are correct when judged against the cost of a Big Mac burger converted into dollars. It was intended as a light-hearted guide to whether currencies are at their 'correct' level (i.e. that they reflect the relative costs of products in different countries). It is based on the theory of purchasing power parity (PPP) – the notion that in the long run, exchange rates should move towards the rate that would equalise the prices of

an identical basket of goods and services (in this case, a burger) in any two countries. For example, the average price of a Big Mac in America in July 2015 was \$4.79; in China it was only \$2.74 at market exchange rates. So, the Big Mac Index says that the Chinese yuan was undervalued by 43% at that time.

'Burgernomics' was never intended as a precise gauge of currency misalignment, but merely as a tool to make exchange rate theory more digestible. Yet the Big Mac Index has become a global

standard, included in several economic textbooks and the subject of at least 20 academic studies. The index is calculated twice each year to reveal changes in the global prices of Big Macs.

Source: Adapted from *The Economist*

Exercises **Total: 15 marks**

- 1 Why has the Big Mac Index proved to be a popular and enduring measure of relative costs of living in different countries? **(6 marks)**
- 2 What are the weaknesses of this index? **(9 marks)**

Other ways of measuring comparative living standards

Since there are so many difficulties in using national income data to judge living standards or to compare living standards in different countries, other measures have emerged. The United Nations, for example, has developed its Human Development Index (HDI) that is based on the average of three indicators:

- life expectancy at birth in years;
- living standards as measured by real GNP per capita using PPP exchange rates;
- a measure of educational attainment.

The HDI only accords a 33% influence to national income in determining the economic wellbeing of the inhabitants of a country, indicating that it does not believe that it plays the major role in determining living standards. We consider the HDI in more detail in Chapter 54.

The Organisation for Economic Cooperation and Development (OECD) has developed a 'better life index'. The OECD describes this as follows: 'There is more to life than the cold numbers of GDP and economic statistics. This index allows comparisons of wellbeing across countries, based on 11 topics the OECD has identified as essential, in the areas of material living conditions and quality of life'. The index incorporates measures relating to housing, the environment, education, health and work–life balance as well as incomes.

Review questions

Total: 36 marks

- 1 Explain the difference between national income and economic growth. **(5 marks)**
- 2 Explain the difference between the output and income methods of collecting national income data. **(7 marks)**

- 3 Which of the following measures the value of all goods and services produced within an economy over a given time period?
- A Gross national product
 - B Gross domestic product
 - C Net national income
 - D Net property income from abroad
- (1 mark)*
- 4 What is meant by the term 'standard of living'?
- (2 marks)*
- 5 Which of the following is *not* a reason for a government to measure the national income generated by its economy?
- A It helps government assess the rate at which prices are increasing
 - B It helps government to make important decisions
 - C It helps businesses and other organisations to make important decisions
 - D It helps government to make judgements about living standards
- (1 mark)*
- 6 Explain why the existence of negative externalities might cause national income data to be an inaccurate indicator of living standards in a country.
- (7 marks)*
- 7 Explain why the existence of a shadow economy can result in inaccurate national income data.
- (6 marks)*
- 8 What is the difference between a market exchange rate and a purchasing power parity (PPP) exchange rate?
- (5 marks)*
- 9 The market exchange rate between Japan and the UK is £1 = ¥180. At the same time a common basket of goods priced at £25 in the UK sells for ¥5000 in Japan. In these circumstances, which of the following is most likely to be true?
- A The use of PPP would give an exchange rate of £1 = ¥250.
 - B Using market exchange rates would overstate relative living standards in the UK.
 - C The use of PPP exchange rates would show that living standards in the two countries are exactly the same.
 - D Using market exchange rates would overstate relative living standards in Japan.
- (1 mark)*
- 10 Which of the following is a true statement?
- A The use of purchasing power parity (PPP) exchange rates tends to understate living standards in developed economies.
 - B The use of purchasing power parity (PPP) exchange rates tends to understate living standards in developing and emerging nations.
 - C The use of market exchange rates always provides more accurate international comparisons of living standards in all economies.
 - D The use of market exchange rates always provides more accurate international comparisons of living standards between developed and developing nations.
- (1 mark)*

Topic 9 Exam-style questions

A-LEVEL PAPER 2

SECTION A Context: The UK's economic performance

Extract A Macroeconomic data

Macroeconomic indicator	UK	Euro area
GDP growth rate (annual), June 2015	2.4%	1.5%
Unemployment rate, August 2015	5.5%	11.0%
Labour costs, June 2015 (UK base year = 2010, euro area = 2011)	101.1	105.6
Annual inflation rate (CPI), September 2015	0.00%	0.05%
Current account balance, 2014	–£97.90 billion	€113.28 billion
Current account balance as a percentage of GDP, 2014	–5.5%	2.1%
GDP per capita (PPP), December 2014	£37,614	€36,925

Note: In December 2014, £1 = €1.26 Source: Trading Economics

Table A Selected macroeconomic indicators for the UK and the euro area

Extract B Relative performance in terms of economic growth

The UK's growth will exceed the growth of its largest eurozone rivals in the years to come, remaining robust despite difficulties in the global economy. Some economists believe that this provides fundamental evidence that the UK economy is outperforming the euro area. The European Commission (EC) predicted that the UK economy would grow 3.1% in 2014, more than ten times faster than France. Growth in the UK is forecast to slow in successive years, but will continue to be higher than the eurozone's four largest economies – Germany, France, Italy and Spain.

Continued economic difficulties in the euro area have hit the UK economy, as the currency bloc includes a large number of its trading partners. The USA, which trades far less with the eurozone, is now expected to grow more strongly than the UK in 2015 and 2016.

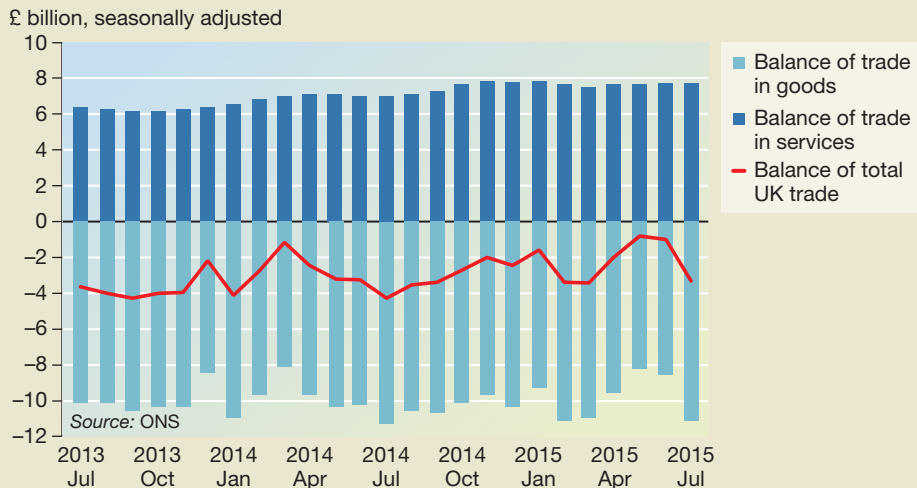
Source: Adapted from the *Daily Telegraph*, 4.11.14

Extract C Trading problems

Concerns are being expressed about the strength of Britain's economic recovery, after declines in both exports and manufacturing output were revealed by recent releases of economic data. These provide evidence that the global economic slowdown may be having an effect on the UK economy's macroeconomic performance. The UK authorities may also be concerned about the possibility of rising GDP and incomes in the UK affecting the volume of goods and services imported.

The UK's deficit on trade in goods and services was estimated to have been £3.4 billion in July 2015, compared to £0.8 billion June 2015. The widening deficit was, in part, the result of exports of goods falling to £22.8 billion, the lowest amount since September 2010. This caused the deficit on goods to widen by 30.6% to £11.1 billion in July.

Figure A Balance of UK trade, 2013–15



Source: Various media sources

The strength of sterling, which makes the UK's exports less competitive in global markets, has been one reason the Bank of England has so far delayed lifting interest rates from their record low of 0.5% to avoid the prospect of rising rates of inflation.

The first estimate for euro area exports of goods to the rest of the world in July 2015 was €185.2 billion. This is an increase of 7% compared with July 2014 (€173.7 billion). Imports from the rest of the world stood at €153.8 billion, a rise of 1% compared with July 2014 (€152.4 billion). As a result, the 19 economies that comprise the euro area recorded a €31.4 billion surplus in trade in goods with the rest of the world in July 2015.

Questions

Total: 40 marks

- Use the information in Extract A to calculate the UK's GDP for 2014. (2 marks)
- Explain how the data in Extract A show that living standards of the inhabitants of the UK are higher than those of the euro area. (4 marks)
- Use the information in Extract C to explain why the UK government's macroeconomic objectives could conflict with one another. (9 marks)
- Some economists believe that the UK's comparatively high rate of economic growth 'provides fundamental evidence that the UK economy is outperforming the euro area'. Using the data in the extracts and your economic knowledge, evaluate the view that high rates of economic growth are always the best measure of macroeconomic performance. (25 marks)

SECTION B Essays

Total: 40 marks

Global trade has increased over recent years – the World Trade Organisation (WTO) is predicting a 5.3% increase in 2015. The importance of trade to many economies, especially the UK, means that price stability remains a very important macroeconomic objective.

- Explain why price stability is the most important macroeconomic objective for the UK authorities. (15 marks)
- Discuss the view that increased international trade and low rates of inflation in many countries mean that it is not necessary to use PPP exchange rates to compare living standards in different countries accurately. (25 marks)



Topic 10

How the macroeconomy works

Aggregate demand & the level of economic activity

Key concepts from Year 1

This chapter builds on Chapter 35 of the Year 1 companion textbook, which examined a number of important economic concepts and issues.

- It established that aggregate demand is the total planned demand for an economy's goods and services at a given price level and over a specific time period. It further explained that demand for an economy's products comes from consumption, investment, government spending and the purchase of the country's exports.
- It set out the factors that determine aggregate demand – the *levels* of consumption, government spending, investment and export sales. These forms of demand are influenced by a range of factors including income levels and expected income (consumption), interest rates and the rate of change of national income (investment), political factors (government spending) and the exchange rate (exports).
- Aggregate demand is an important influence on the level of economic activity in an economy. In most circumstances a rise in aggregate demand will cause a rise in economic activity, as measured by real GDP. It may also lead to an increase in the price level. A fall in the level of aggregate demand would normally provoke the opposite effects on the levels of real GDP and prices. The impact of changes in aggregate demand can be illustrated on an aggregate demand and supply diagram.
- This chapter also introduced the notion of the multiplier effect. This is a process through which a change in any component of aggregate demand results in a magnified change in real GDP or national income. The multiplier can operate within a local or the national economy.

This chapter introduces the notion of the marginal propensity to consume. It further builds on your understanding of the effects of changes in the level of aggregate demand within an economy to consider how the multiplier can be calculated using the marginal propensity to consume. It examines why the size of the marginal propensity to consume determines the magnitude of the multiplier effect.

Key terms

Consumption is planned spending by households on goods and services.

A **consumption function** illustrates the relationship between consumption and real income.

The **circular flow of income** is a model that shows how money flows within a simplified economy, with households and firms as key components.

The **marginal propensity to consume (MPC)** is the proportion of any additional income that is spent on consumption.

The marginal propensity to consume

We saw in Chapters 34 and 35 of the Year 1 companion textbook that income flows around an economy, passing between households and firms. The decisions by firms and households to withdraw money affect these flows for three reasons:

- savings to finance future consumption or investment;
- the purchase of imports of goods and services (which means the money flows out of the domestic economy);
- the payment of taxes to the government or other authorities.

Money is also added to the circular flow as a result of firms taking decisions to invest. These additions are called injections and can take the form of investment, government spending or export sales. The circular flow of income, withdrawals and injections are illustrated in Figure 38.1.

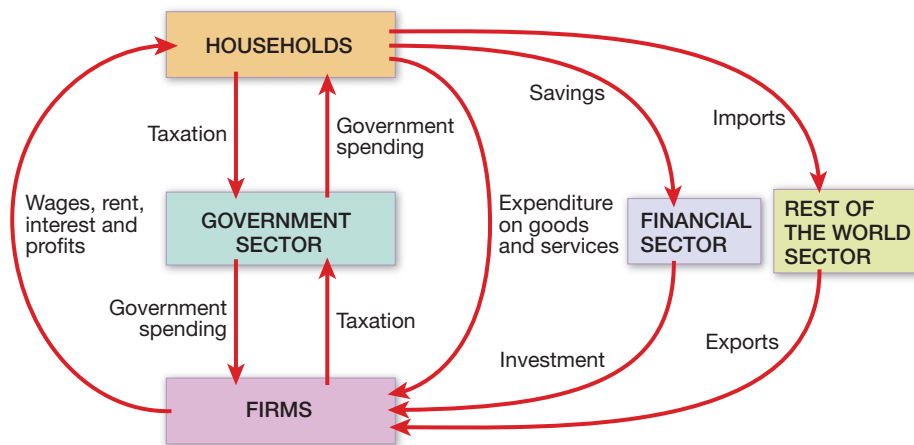


Figure 38.1 The circular flow of income, withdrawals and injections

This model shows that not all income received by households is spent on goods and services – only a part of the income received by households is used for consumption purposes. The remainder is withdrawn from the economy through saving, the purchase of imports or the payment of taxes. The proportion of any increase in household income that is used for consumption (i.e. that is passed on in the circular flow) is called the marginal propensity to consume (MPC).

The consumption function and the MPC

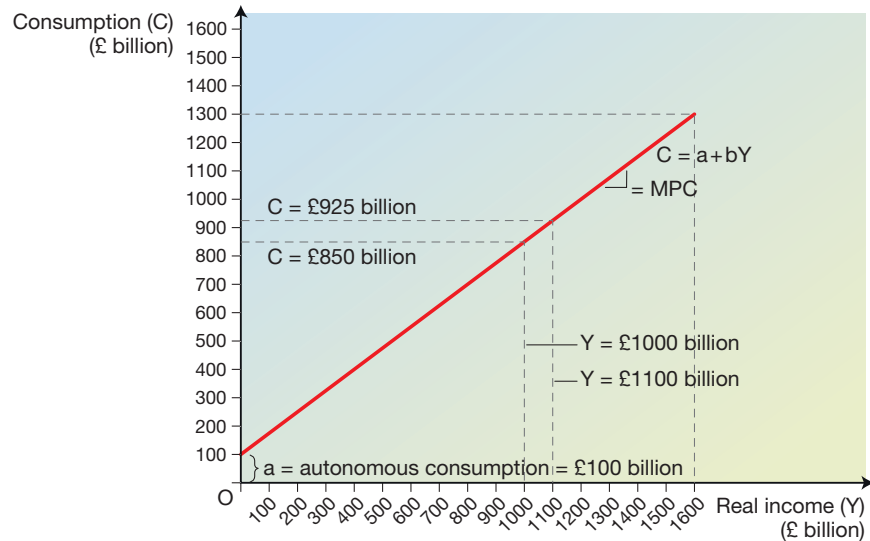
The relationship between the level of consumption and the level of real income in an economy can be shown on a consumption function, as in Figure 38.2. We saw in the Year 1 textbook that a consumption function expresses the relationship between real income and consumption. The consumption function in Figure 38.2 is based on broadly Keynesian ideas. It has a level of consumption that is undertaken by a society irrespective of its level of income. This is necessary to maintain a basic standard of living and is termed ‘autonomous consumption’. Beyond this point, consumption shows a direct relationship with income. An increase in real income provokes a positive, but smaller, rise in consumption. This is the MPC, which can be expressed as $\Delta C/\Delta Y$, where Δ represents the quantity of change.

This consumption function can be expressed using the formula: $C = a + bY$, where ‘a’ is autonomous consumption and ‘b’ represents the MPC. Since the consumption

function in this example is a straight line, it follows that the MPC is constant across this range of real income.

In our example in Figure 38.2, the level of autonomous consumption is £100 billion. A rise of income of £10 billion in this example would result in an increase in consumption of £7.5 billion.

Figure 38.2
A consumption function
and the MPC



Calculating the MPC

The MPC can be calculated by comparing the change in consumption to the change in income that caused it. In Figure 38.2, an increase in real income from £1000 billion to £1100 billion results in consumption increasing from £850 billion to £925 billion. Thus, in these circumstances:

$$\text{MPC} = \Delta C / \Delta Y = £75 \text{ billion} / £100 \text{ billion} = 0.75$$

The value of the marginal propensity in this example would be the same (0.75) throughout the range of real income that is shown on this graph.

Influences on the value of the MPC

1 Households' expectations

If the consumers within a household are confident that their futures are secure and that they can expect to maintain, or even increase, their current levels of disposable income, then this may prompt them to increase the proportion of income that they spend on consumption and to decrease their rate of savings. At the time of writing, many UK households are positive about the security of their future incomes and willing to reduce their level of savings. In other circumstances, when economic conditions are less favourable, households may be reluctant to reduce their level of savings.

2 Households' income levels

Income is a flow of money paid to households over a period of time. Those households that receive lower incomes tend to have a higher MPC. This is because much of their expenditure is on essential items such as food, with comparatively little spent on

luxury products – the type of expenditure that can more easily be withheld. Thus low-income households are likely to spend a high proportion of any increase in real income. In contrast, high-income households tend to spend a lower proportion of their income on essential products and will typically save a greater proportion of their income. As John Maynard Keynes first argued, rich households will have a lower MPC.

3 The rate of interest

The MPC of most, if not all, households will be determined to some extent by the rate of interest. Increases in the rate of interest could be expected to reduce the level of an economy's MPC for a number of reasons.

- A rise in interest rates makes saving more attractive because it increases the returns, thereby increasing the opportunity cost of consumption.
- Some forms of consumption, such as the purchase of property and vehicles, are normally financed through borrowing. Higher interest rates will effectively increase the cost of purchasing these products for many consumers and may reduce the MPC.

However, the population in the UK is steadily ageing, resulting in more people relying on retirement incomes to fund their consumption. Some of this income will come from a variety of types of saving, including annuities, which will provide a higher income when interest rates rise. This may boost the MPC amongst some groups in society.

Falling interest rates could be expected to have the opposite effects on MPC to those described above.

The multiplier effect

We encountered the multiplier in Chapter 35 of the Year 1 companion textbook. We saw that a rise in injections within the circular flow model will result in an increase in national income or real GDP. The relationship between the rise in the level of injections and the increase in national income that results is more complex than might be imagined. The rise in injections will result in a magnified increase in the level of national income as a result of the multiplier process. A rise in any of

Key term

The **multiplier** is a process through which any change in a component of aggregate demand results in a magnified change in real GDP or national income.

REALWORLD ECONOMICS 38.1

British retailers buoyed by September spending surge

British retailers are reporting the best business conditions since before the financial crisis erupted in 2007 as low inflation and rising wages prompt a surge in consumer spending. It is likely that the rise in real incomes has increased the marginal propensity to consume amongst many UK households.

The results of a survey by the

Confederation of British Industry (CBI), an employers' organisation, revealed that high street and online sales volumes were well above normal for the time of year. Retailers selling food and clothes enjoyed particularly strong sales.

The CBI stated that 41% of the 119 retailers surveyed reported that business was above average



Boom times return for food and clothing retailers, September 2015

for September, against 6% who said they were below average. The balance of +35 percentage points

Month and year	Index of real average weekly earnings (2000 = 100)
December 2013	109.1
March 2014	108.3
June 2014	108.5
September 2014	109.4
December 2014	110.2
March 2015	111.4
June 2015	111.6
September 2015	111.2

Source: ONS

Table 38.1 Index of UK real average weekly earnings, 2013–2015

was the highest since April 2007. Retailers are planning to increase orders with suppliers at the fastest rate in almost five years.

Rain Newton-Smith, the CBI's Director of Economics, said: 'Low inflation and the recovery in wage growth are helping to stimulate consumer demand, but the slowdown in the global economy and tight margins mean retailers won't get ahead of themselves as we head into autumn... Clothing and the food and drink sectors have led the way, and with the general outlook for the UK economy remaining positive, the expectation is that there will be further encouraging results next month'.

The CBI survey found that 60%

of retailers questioned indicated that sales volumes in September were up on the same month in 2014, against 11% that reported a decline.

At the same time weekly earnings from total pay have shown growth in real terms, as illustrated in Table 38.1.

Source: Adapted from the *Guardian* (Larry Elliott), 29.9.15

Exercises **Total: 16 marks**

- 1 Explain the possible factors that may have caused an increase in the marginal propensity to consume in the UK. **(7 marks)**
- 2 Do you think that the marginal propensity to consume in the UK is likely to continue to rise? Explain your reasoning. **(9 marks)**

the injections (investment, government spending or exports) will start the multiplier process. We assume here that we are considering an open economy with international trade taking place as well as a government sector. This means that three types of withdrawal exist whereby income is not passed on in the circular flow – saving, imports and taxation.

Figure 38.3 An example of the multiplier process

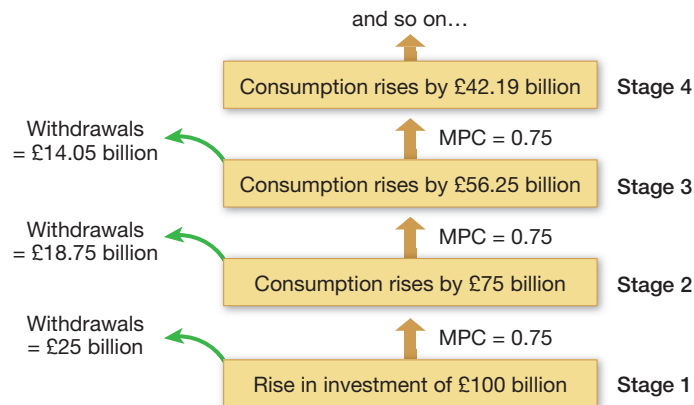


Figure 38.3 illustrates the multiplier process using a simple example. A rise in injections (we have assumed this is an increase of £100 billion in the level of investment) initiates the multiplier process. The investment funds find their way into the economy, possibly in the form of wages to the employees of businesses engaged in the investment activity. They spend a proportion of the incomes they receive on goods and services. We have assumed that households in this economy have a marginal propensity to consume of 0.75. Thus the £100 billion injection would result in a rise in consumption of £75 billion in the second stage of the multiplier process.

Simultaneously, withdrawals would increase by £25 billion as some of the income received is not consumed, but is saved or paid in taxes and thus not passed on in the circular flow.

The rise in consumption in Stage 2 of the multiplier process means that £75 billion is passed on within the circular flow and is received as incomes by households. This will generate another stage of the process in which 0.75 of the £75 billion is consumed and 0.25 withdrawn from the circular flow. Thus consumption will rise by a further £56.25 billion and withdrawals by £18.75 billion. This process will continue, with the amounts at each subsequent stage becoming smaller.

The operation of the multiplier means that the impact of a rise in an injection, such as investment, will result in a higher level of aggregate demand and a higher level of real GDP or national income.

Calculating the size of the multiplier

It is apparent from the example in Figure 38.3 that the eventual rise in income will be greater than the rise in investment that caused it. This is true because the initial £100 billion addition to the circular flow is supplemented by further amounts of £75 billion, £56.25 billion and so on. How much will the eventual increase in national income be?

There are two formulae by which the value of the multiplier, and hence the effect of an increase in injections on national income, can be calculated.

The multiplier can be calculated using the following formula (in which the value of the multiplier is shown by the letter 'K'):

$$K = 1/1 - MPC \text{ or } K = 1/MPW$$

where MPC is the marginal propensity to consume and MPW is the marginal propensity to withdraw. If we assume that consumption only relates to domestically produced goods and services and excludes sales taxes, then these two formulae should give the same value for the multiplier.

Thus, in our example set out in Figure 38.3 the calculation of the multiplier would be as follows:

$$K = 1/1 - 0.75 = 1/0.25 = 4$$

So, the multiplier has a value of 4. This means that the final increase in national income will be four times the size of the original increase in investment. Therefore:

$$\Delta Y = K \times \Delta J \text{ where } \Delta \text{ represents change}$$

Thus, the change in national income in our example will be $4 \times £100 \text{ billion} = £400 \text{ billion}$.

Equally we could assume that, if the consumption relates to domestically produced goods and excludes sales taxes, an MPC of 0.75 would give an MPW of 0.25. Thus, using the other formula, $K = 1/0.25 = 4$.

Number crunching

Assume that the value of the MPC was 0.6 in the example in Figure 38.3. What would be the final effect on national income? By how much would this differ from the outcome when the MPC was 0.75?

Key term

The **marginal propensity to withdraw (MPW)** is the proportion of any additional income that is not passed on within the circular flow.

The formula based on the MPW does allow the calculation of the multiplier by reference to the proportion of income not passed on, and can deal with data presented in a different form.

For example, assume a rise in injections of £80 billion takes place in an economy in which the following marginal propensities exist:

- marginal propensity to save = 0.25;
- marginal propensity to import = 0.2;
- marginal propensity to pay taxes = 0.15.

In these circumstances the $MPW = 0.25 + 0.2 + 0.15 = 0.6$.

Thus the value of the multiplier = $1/0.6 = 1.67$. So, an increase in injections of £80 billion in this five-sector economy will lead to an increase in national income of £133.6 billion (£80 billion \times 1.67). If we hold our assumptions about consumption, this would equate to a situation in which the $MPC = 0.4$.

Author tip

Although we have looked at two different ways of calculating the value of the multiplier, AQA examinations will only require you to use the $K = 1/1 - MPC$ formula. However, it is important that you understand the importance of the marginal propensities (collectively the MPW) to understand what determines the value of the multiplier.

What determines the size of the multiplier?

Spare productive capacity tends to boost the value of the multiplier as it makes it possible for an economy to respond to increases in aggregate demand that occur as part of the multiplier process. If an economy has limited spare capacity, any increases in demand will be likely to provoke increases in prices rather than in real GDP.

Clearly, the value of an economy's MPW determines the size of its multiplier: the higher the value, the lower will be the value of the multiplier. This is the case because if more income is withdrawn from the circular flow at each stage of the multiplier's operation, the overall impact on national income will be reduced.

Therefore in economies where the rate of taxation is high, as is the case in Scandinavian countries, the value of the multiplier will be lower than would otherwise be the case. In other economies the rate of saving is much higher than in the UK, often for cultural reasons. In China, for example, consumers save approximately 30% of their disposable income, compared with around 5% in the UK. This high rate of withdrawal depresses the value of the Chinese multiplier.

Finally, the propensity to import is an important determinant of the value of the multiplier. The UK tends to spend relatively heavily on imports, which depresses the value of its multiplier.

The reverse multiplier

The multiplier can also operate in the opposite direction, leading to a fall in national income. This is called the reverse or downward multiplier. It can be caused by a fall in the level of injections whereby the process illustrated in Figure 38.3 works in the opposite direction, resulting in lower levels of consumption.

Review questions

Total: 30 marks

- 1 Explain the difference between consumption and a consumption function. **(5 marks)**
- 2 State three reasons why income may be withdrawn from the circular flow. **(3 marks)**
- 3 Which of the following best describes the marginal propensity to consume?
 - A The level of consumption by households following a rise in real income
 - B The proportion of any additional real income that is spent on consumption
 - C The percentage of real income that is represented by consumption
 - D The amount by which consumption increases following a rise in real income**(1 mark)**
- 4 An economy's real income rises from £1500 billion to £1650 billion. As a consequence, consumption expenditure increases by £105 billion. Which of the following statements is true?
 - A The economy's MPC = 0.75
 - B The economy's MPC = 0.8
 - C The economy's MPC = 0.85
 - D The economy's MPC = 0.7**(1 mark)**
- 5 Explain how differences in households' incomes may influence the value of their marginal propensities to consume. **(6 marks)**
- 6 Explain why increases in the rate of interest might result in a fall in the marginal propensity to consume. **(6 marks)**
- 7 Define the term 'the multiplier'. **(2 marks)**
- 8 An economy has experienced the following changes: real national income has risen by 12% from £550 billion; and consumption expenditure has increased by £39.6 billion as a consequence. Which of the following statements is true?
 - A The value of the multiplier would be 1.67
 - B The marginal propensity to withdraw in this economy is 0.7
 - C The value of the multiplier would be 2.5
 - D The marginal propensity to consume is 0.75**(1 mark)**
- 9 If the value of the marginal propensity to consume in an economy rises from 0.6 to 0.65, which of the following will be true?
 - A A rise in investment of £100 million will lead to a rise in national income of £167 million following the change in the MPC
 - B The original value of the multiplier was 1.54
 - C The effect of the change in MPC means that a rise in investment of £100 million will increase national income by £13 million less than would have been the case
 - D Any rise in investment will result in a larger final change in national income**(1 mark)**
- 10 Describe the relationship between the marginal propensity to withdraw and the size of the multiplier. **(4 marks)**

The determinants of long-run aggregate supply

Key concepts from Year 1

This chapter builds on Chapter 36 of the Year 1 companion textbook, which looked at the determinants of aggregate supply.

- It introduced the concept of aggregate supply as a partner for aggregate demand. It also examined the factors that determine aggregate supply in the short and long runs. Key determinants of aggregate supply in the short run include the price level in the economy (as a higher price level makes supplying goods and services potentially more profitable), the cost of employing labour and acquiring other resources such as raw materials, the level of taxes paid by businesses in the UK, and the productivity of an economy's factors of production.
- We saw that in the long run, aggregate supply is determined by the quality (and hence the productivity) of factors of production and the quantity of factors of production. We made an assumption that the long-run aggregate supply curve was vertical and thus independent of the price level. We shall relax this assumption in this chapter.

This chapter investigates the effect that institutions within the UK economy, such as the banking system and trade unions, can have on its long-run aggregate supply. It also explores the Keynesian long-run aggregate supply curve and the factors that explain its shape, as well as its implications for the economy.

Long-run aggregate supply

Long-run aggregate supply (LRAS) is the total quantity of output that producers in an economy are willing to supply over a period of time during which the prices of all factors of production can vary. In the long run it is possible for producers to adjust the quantity of all types of resources that are used in production.

The long run has a specific meaning in economics. This meaning was clearly defined by two American economists, Rodney Mabry and Holley Ulbrich: 'The long run is the time period in which anything can be changed, or in which individuals and firms are fully able to respond to economic incentives and take advantage of economic opportunities. The long run has no specific time frame; it is simply the time period that is long enough to allow full response to changing incentives' (*Introduction to Economic Principles*, 1989).

The significance of this for an economy's aggregate supply is apparent: in the long term, firms and other decision makers can take full advantage of the opportunities afforded by factors such as changes in the price of factors of production or final goods and services by adjusting production.

Key terms

Aggregate supply is the total quantity of output that producers in an economy are willing to supply at specific price levels over a period of time.

Productivity measures the efficiency with which inputs are transformed into outputs.

Aggregate supply and the institutional structure of the economy

Institutions are mechanisms that govern the behaviour of individuals and organisations within an economy. Those institutions that have an impact on the aggregate supply within the UK include the government and its agencies, such as the Bank of England. These influence the behaviour of firms and households within the UK by offering incentives and imposing rules.

There are a number of aspects of the institutional structure of the UK economy that can have an impact on the aggregate supply of an economy by affecting the quality and quantity of factors of production that are available to businesses.

1 The banking system

The UK's banking sector is very large relative to the rest of the economy. It focuses heavily on more risky types of banking – often referred to as ‘investment banking’. Its size and importance to the UK economy meant that the consequences of the financial crisis in 2008 were severe. The cost of rescuing UK banks (principally Lloyds Banking Group and RBS) from failing was £66 billion. The figure was this high because of the amount of debt they had incurred from highly speculative lending.

Since the financial crisis, the UK authorities have tightened the regulations concerning banking. One new regulation requires UK banks to hold higher amounts of reserves to deal with bad debts in the event of another financial crisis or recession. The intention is that they should have sufficient funds to avoid needing support from the UK government. Following the financial crisis, UK banks were required to hold reserves equivalent to 3% of their loans – known as the ‘leverage ratio’. This has reduced lending to business because UK banks have sought to increase their capital reserves in relation to loans in order to meet this new requirement. The effects of this are shown in Table 39.1.

Key term

The **leverage ratio** is the minimum amount of capital (expressed in percentage terms) that banks must hold in comparison to the amount they have lent. The current figure in the UK is 3%.

Time period	Net flow (£ billion)
2003–08	3.8
2012	–1.5
2013	–0.7
2014	–0.3
2015 Quarter 1	1.2
2015 Quarter 2	0.0
2015 Quarter 3	0.3

Source: Credit Conditions Review, August 2015 (Bank of England) Note: All figures are monthly. Net flow of lending is loans made less payments received. A negative figure indicates that repayments exceeded new loans advanced.

Table 39.1 Net monthly flow of lending to UK businesses that produce goods and/or supply non-financial services

Table 39.1 shows that bank lending to UK businesses (which is likely to be used to fund investment) has been low, with few exceptions, since 2012. Tighter regulations are possibly in the pipeline as the Bank of England has suggested that the leverage ratio might rise to 4.95% from 2019. The consequence of such a change for banks is that they would be required to have £1 of reserves for each £20 of loans they make, rather than the current £1 for every £33 of loans.

The UK authorities have implemented a series of policies intended to free up credit

REALWORLD ECONOMICS 39.1

Bank lending finally shows signs of recovery

Two and a half years after the government unveiled a flagship Bank of England loan scheme, banks are finally growing their lending to small businesses

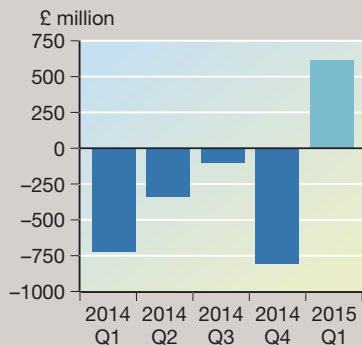


Figure 39.1 Net lending by banks to small and medium-sized business in the UK, 2014–15

(Figure 39.1), in a sign that the UK's shortfall of business investment is coming to an end.

Data released by the Bank of England showed that net lending – the difference between new loans being granted and companies paying off debt – from banks participating in the Funding for Lending Scheme grew by £635 million in the first three months of 2015. This is the first time that net lending has grown since the FLS, which allows banks to borrow cheaply from the Bank of England as long as they lend the money out, was refocused to target business lending from the start of 2014.

The £635 million increase in the first quarter of 2015 compared to an

£811 million decline in the final three months of 2014, and a £2 billion fall in the whole of 2014. The figures do not represent all bank lending, with the likes of Barclays and HSBC not participating in the scheme, but are further evidence of a crucial rebound in business lending as economists hope for improvements in investment and productivity.

Source: *Daily Telegraph*, 28.5.15

Exercise **Total: 14 marks**

- 1 Explain why an increase in bank lending might result in a rise in aggregate supply. **(6 marks)**
- 2 Why was it necessary for the government to intervene to increase lending to businesses by banks? **(8 marks)**

in the banking system and to boost lending to businesses for investment purposes. The latest and most ambitious is the Funding for Lending Scheme (FLS), which was launched in July 2012 and currently runs until 2016. The FLS allows banks to borrow cheaply from the Bank of England so that they can pass on these loans, at low interest rates, to businesses. It is hoped that the FLS will increase bank lending by up to £70 billion.

The nature and scale of the UK banking industry has meant that its operation in recent years has not supported an increase in aggregate supply as fully as might have been possible. This has been one factor holding back business investment, although economists recognise that demand for loans for business investment have been low at a time when many firms have not enjoyed rising demand for their products.

2 The labour market

Labour is a major resource used by firms in the UK economy, particularly in industries that supply services. The quantity and quality of labour available to firms is a significant determinant of their ability to produce goods and services and therefore of aggregate supply.

There are a number of institutional aspects of the UK economy which impact on labour markets and aggregate supply.

The declining role of trade unions

The power and importance of trade unions within the UK economy has declined

since the late 1970s. In 1979 about 13 million workers were members of a trade union. This figure fell to 8.4 million in 1995 and, by 2014, the number of members had fallen to 6.4 million. This decline occurred despite a substantial growth in the number of people working in the UK over this period. In 1995, 34.4% of employees were members of a trade union; by 2014 the percentage had fallen to 25.0%.

Number crunching

Use the information in the paragraph above to calculate:

- 1 the number of people employed in the UK in 1995 and 2014;
- 2 the percentage change in the number of people working in the UK over this period.

This change has come about for two principal reasons. First, there has been a decline in those industries that were previously bastions of trade unionism, such as coal mining and steel manufacture. Secondly, UK governments have passed a series of laws designed to limit the power and influence of trade unions. As a result, the UK has developed a more flexible workforce made up of greater numbers of part-time and temporary employees, which enables firms to increase production swiftly in response to changes in demand. This flexibility has helped to make the country an attractive location for overseas producers and to persuade many multinational businesses (Nissan and Hitachi, for example) to invest in the UK. In 2013, nearly 1800 projects were funded by overseas investors, creating close to 67,000 new jobs, according to a report from UK Trade & Investment. In fact, the UK attracts more foreign investment than any other European country and the flexibility of UK employees is one reason for the country's popularity among foreign investors.

The increasing use of migrant labour

UK businesses have used large amounts of migrant labour over recent years, making extensive use of the European Union's right for its citizens to live and work in any member country. Many UK employers automatically look overseas for employees, seeking to reduce training costs as well as other employment costs. However, recent decisions by the government may lead to a change in this approach because it is more difficult for UK firms to employ foreign workers who are not EU citizens.

According to research by the Migration Observatory, the number of foreign-born people of working age in the UK more than doubled from 2.9 million in 1993 to just over 6 million in 2013. At the same time the percentage of those in employment who are foreign born has risen sharply to about 15% of the workforce. Migrants comprise about 9% of the EU's workforce as a whole, showing that the UK is more open to foreign-born workers forming part of the country's workforce. The implications for its aggregate supply are apparent.

Although many of the migrant workers are employed in low-skill occupations in manufacturing and the catering industry, for example, others have filled highly skilled professional roles in the National Health Service and as senior managers in UK companies.

Education and enterprise

The UK has a culture that is enterprising and inventive, with institutions that offer educational support to enhance this. Enterprise education is an important

Key terms

Trade unions are organisations that operate with the intention of protecting and enhancing the working conditions and financial positions of their members.

Flexible workforces exist when firms make less use of permanent full-time employees and rely more on part-time and temporary employees to supply goods and services.

part of the curriculum in many schools and colleges. It enables students to acquire relevant business-related knowledge, skills and competences for enterprise and entrepreneurship. Furthermore, students receiving enterprise education are more likely to have positive attributes, such as risk taking, and positive intentions, such as becoming self-employed or being entrepreneurial as an employee within a business.

Research by the UK government reveals that there are positive statistical relationships between various enterprise and entrepreneurship learning activities in education and economic impacts. Impacts include starting a new business, increased employability and earnings, and contributing to the growth of businesses (especially for graduates entering small businesses). These suggest that enterprise and entrepreneurship education is a positive stimulus for economic growth and aggregate supply.

The UK government has given considerable emphasis to enterprise education within the curriculum. The views of the government minister responsible for this type of education make clear the positive implications of enterprise education for production in the UK economy: ‘Enterprise education is a high priority for the Government and something I take a close personal interest in. I’m therefore delighted that this guidance has been produced, exemplifying the sort of good practice I’ve seen myself up and down the country, often embedded in the 14–19 reforms. Young people today often say they want to be their own boss, to start their own company or to make a living from a personal passion. Enterprise education will help them to do that successfully, to their own benefit and that of the economy and their local community’ (*A Guide to Enterprise Education*).

Enterprise education does not stop when students leave school. It is offered to students at over 90% of universities in the UK and part of 75% of all courses delivered. Both the government and the universities recognise that enterprise education for all students is vital to promote the right attitudes and to impart the skills necessary to fuel economic growth.

The Keynesian long-run aggregate supply curve

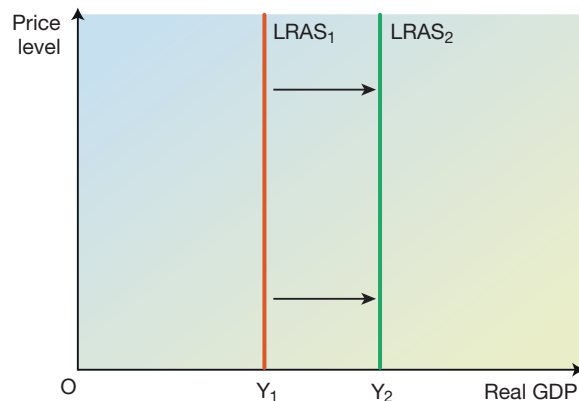


Figure 39.2 A shift in a vertical LRAS curve

Key term

Keynesian economics

refers to economic theories based on the work of the British economist John Maynard Keynes (1883–1946).

Economists hold diverse views about the shape of an economy’s long-run aggregate supply (LRAS) curve. In our Year 1 companion textbook it was assumed that the LRAS curve was vertical, as shown in Figure 39.2. A vertical LRAS implies that aggregate supply is not determined in the long run by the general price level but by other factors such as advances in technology and the quantity and productivity

of all the economy's factors of production. Thus an increase in the level of foreign investment into the UK, as occurred in 2013–14, would shift the LRAS to the right. However, a change in the general price level would have no effect.

Key note

The AQA AS specification states that 'it is assumed that in the long run the AS curve is vertical'. However, this assumption is relaxed for the full A-level, which opens up an important area of debate that we will explore over the next few pages.

Classical economists and the LRAS curve

The fundamental reason why classical economists (and their followers, the neoclassicists) believe that the LRAS is vertical, as in Figure 39.2, is that wages and prices are flexible – downwards as well as upwards. Therefore, in the long run there cannot be any unemployment because the wage rate will adjust to create a situation where any worker who is seeking a job will be able to find one. This assumption that labour and other markets operate efficiently results in a situation in which an economy always produces close to, or at, its maximum capacity. For example, a short-term fall in aggregate demand would reduce the general price level. As a consequence, workers would be willing to accept lower wages without reducing their living standards, pricing themselves back into the labour market and maintaining the full-employment, long-run equilibrium level of output.

Classical economists argue that in the long run, firms will supply the maximum potential output of the economy. The result is that a change in aggregate supply can only be brought about by a change in the productive capacity of the economy.

Keynesian economists and wage 'stickiness'

Keynesian economists take a different view of how an economy adjusts to changes in aggregate demand and this is based on the notion of the 'stickiness' of real wages. Many of the economists in this school of thought believe that unemployment can exist in the long run because wages do not adjust downwards. In fact, they argue that at times of unemployment, employees will negotiate wage rises. As a consequence, real wages will be above the free market equilibrium (they will not fall – they are 'sticky') and the market will not clear. The result will be unemployment in the long run, since there is a lack of demand in the economy. This is termed 'demand-deficient unemployment'.

Keynesian economists divide the LRAS into three distinct sections, as explained below.

1 Low levels of output

The first section is labelled as (a) in Figure 39.4. Over this element of the LRAS curve supply is highly price elastic, allowing output to be increased without the price level rising. Over this range of output, since unemployment is very high and workers are nervous about job security, expanding supply does not push up wages. Additional employees can be hired at the current wage rate. This may occur in a recession and this is represented by a horizontal curve and reflects the existence of unused or spare productive capacity.

REALWORLD ECONOMICS 39.2

Real wages stagnate in the UK

The UK economy continues to do pretty well, despite the weakness of the European economies. The UK's GDP exceeded forecasts in 2014, and in the second quarter was 3.1% larger than a year ago. Yet working Britons are not feeling the benefit. Real wages have fallen for seven consecutive years and are 6.9% below their 2007 level. Britain is experiencing its longest period of pay stagnation since records began in 1855 (Figure 39.3).

In October, 90,000 workers took to London's streets to join a protest organised by the Trades Union Congress (TUC). Participants complained about not just pay stagnation but also inequality: chief executives of the UK's largest firms now typically earn 120 times as



Workers march through London protesting against austerity, October 2014

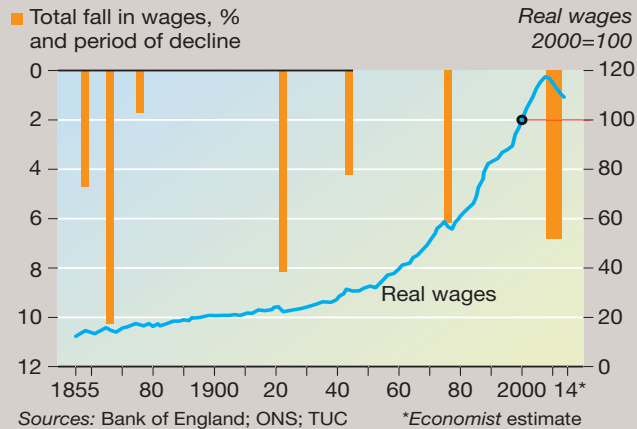


Figure 39.3 UK real wages 1855–2014

much as their average employee, up from 47 times as much in 2000, according to Incomes Data Services, a research company.

As well as pay rises for workers, the TUC wants more companies to pay the living wage – £7.65 (\$12.30) an hour, or £8.80 in London – to their lowest earners, and a crackdown on 'excessive' executive pay.

Low pay reflects Britain's dismal record since the recession when it comes to productivity: output per hour worked remains 2.2% below its pre-crisis peak and 16% below its pre-crisis trend. The recession caused less unemployment than elsewhere,

but real wages fell instead as productivity tanked. That is fortunate in some ways, as unemployment has acute and concentrated costs, but those in work are now in need of some respite.

Source: *The Economist*, 25.10.14

Exercises Total: 13 marks

- 1 What impact might a fall in labour productivity have on the UK LRAS curve? Explain your reasoning. (5 marks)
- 2 Explain whether the information in this case study suggests that the UK economy's LRAS curve is vertical? (8 marks)

2 Intermediate levels of output

The second stage is labelled as (b) in the diagram. Over this range of output, increases in production can only be achieved at the expense of rises in the general price level because resources become relatively scarce and spare productive capacity diminishes. Workers feel more secure as unemployment diminishes and they feel able to negotiate rises in real wages, despite the continuing existence of some unemployment. As output approaches the full employment level, a greater proportion of workers will be able to negotiate real wage rises, and the aggregate supply curve becomes more vertical.

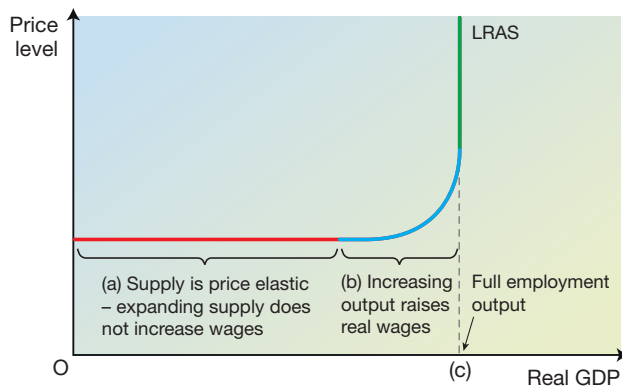


Figure 39.4 The Keynesian LRAS curve

3 Full employment output

Once the full employment level of output has been attained, labelled (c) in Figure 39.4, attempts to increase aggregate supply by increasing aggregate demand will only result in a rise in the general price level. At this level of output, Keynesian and classical economists agree about the shape of the LRAS curve.

The implications of the Keynesian LRAS

By accepting that the LRAS is the shape illustrated in Figure 39.4, rather than vertical, Keynesian economists contend that unemployment can exist in the long run because labour markets do not always clear. This means that a long-run cause of unemployment could be a lack of demand if the aggregate demand curve is at a position in which it intersects the LRAS curve in sections (a) or (b). By managing the level of demand, Keynesian economists argue, it is possible to reduce long-run unemployment, though this does tend to cause prices to rise as a consequence.

At the full employment level of output, point (c) in Figure 39.4, Keynesian and classical economists agree that the LRAS curve is vertical.

Review questions

Total: 22 marks

- 1** Define the term 'aggregate supply'. **(2 marks)**
- 2** State the key distinction between the short and long run in economics. **(2 marks)**
- 3** Which of the following factors would be most likely to have resulted in a reduction in lending to businesses by banks in the UK? **(1 mark)**
 - A A fall in the leverage ratio
 - B A rise in the leverage ratio
 - C The implementation of the Funding for Lending Scheme (FLS)
 - D The Bank of England holding bank rate at 0.5% for more than six years
- 4** Explain the likely impact of a fall in bank lending to businesses on the aggregate supply of an economy. **(5 marks)**
- 5** Define the term 'flexible workforces'. **(3 marks)**

- 6** Which of the following statements about a vertical LRAS curve is correct?
- A It shows that supply is perfectly price elastic
 - B It is vertical because wages and prices are only flexible upwards, not downwards
 - C The wage rate adjusts so that all workers seeking a job are successful
 - D The level of aggregate supply depends upon changes in the general price level
- (1 mark)*
- 7** Which of the following is a consequence of 'sticky' real wages during a recession?
- A Real wages are lower than those necessary for the market to reach equilibrium
 - B There is excess demand for labour
 - C The economy maintains its full-employment level of output
 - D The labour market fails to clear
- (1 mark)*
- 8** Keynesian and classical economists agree over which of the following?
- A That the LRAS curve is perfectly price inelastic at full employment output
 - B That the labour market fails to clear in the long run
 - C That the LRAS curve is perfectly price elastic at full employment output
 - D That real wages fall to allow workers to price themselves back into employment
- (1 mark)*
- 9** Draw and label fully the Keynesian LRAS curve.
- (5 marks)*
- 10** Which of the following is an implication of the Keynesian LRAS?
- A That governments should impose minimum wages to maintain living standards
 - B That any attempt to increase aggregate demand will only result in a rise in the general price level
 - C That employment can always be increased without any effect on the general price level
 - D That it is necessary for governments to manage aggregate demand to maintain full employment
- (1 mark)*

Topic 10 Exam-style questions

A-LEVEL PAPER 2

SECTION A Context: Productivity, investment and consumption in the UK

Extract A Productivity data

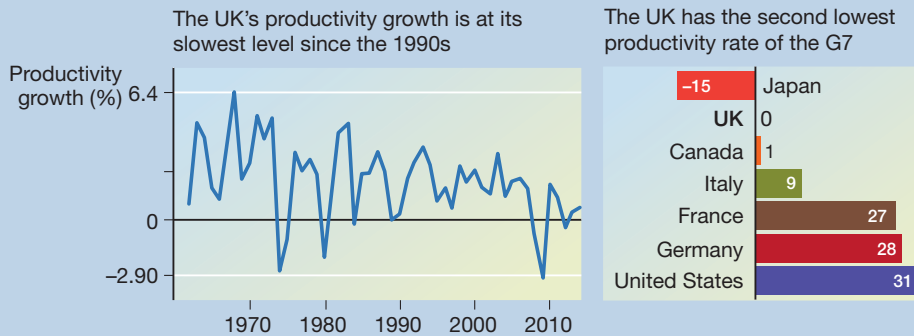


Figure A UK labour productivity growth, 1970–2014 (%)

If the UK matched the productivity rate of the USA, its economy would grow by 31%

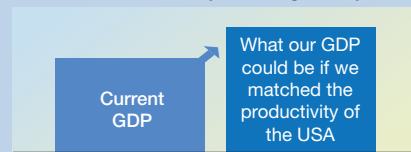


Figure C The effect of higher labour productivity rates

Figure B Comparative data for labour productivity (%)

The three charts reveal just how poorly the UK's workforce has performed over the last 30 years in comparison to countries such as France, Germany and the United States. Figure C shows the potential benefits if the UK can start producing more for every hour worked. If the UK matched the productivity rate of the United States, its GDP would increase by 31%, equivalent to £21,000 a year for every household in the UK.

Sources: Adapted from the *Independent*, 10.7.15; ONS data

Extract B Investment in UK infrastructure

A report by Standard & Poor (an American financial services company) has stated that increasing UK investment in infrastructure by 1% of GDP would create more than 200,000 new jobs in the same year. The 'Building For Growth' report claimed that each additional £1 that the UK spends on infrastructure in one year would increase the country's GDP by £1.90 over a three-year period. The UK's investment in infrastructure, the report argued, has lagged behind several other countries in the Organisation for Economic Cooperation and Development (OECD).

It is not surprising, therefore, that the head of the International Monetary Fund has supported plans to improve the UK's infrastructure, describing it as a 'win-win' option to boost growth. Christine Lagarde argues that targeted investment in infrastructure will create jobs in the short term and lift future growth prospects.

Source: Adapted from the *Daily Telegraph*, 8.10.15

Extract C **Rising consumption boosts UK aggregate demand**

Consumer spending accounts for over 60% of the UK's GDP and is the most important component of aggregate demand. PWC, an accountancy firm, believes consumption could grow by around 2% per annum in real terms on average to 2030, but this could vary from around zero real growth for food, clothing, and alcohol and tobacco to around 2.5–3% for housing and utilities, health, and recreation and culture.

UK retailers are reporting the best business conditions since before the 2008 financial crisis. Consumer confidence is growing as the UK's inflation rate hovers around zero and a rise in real wages is finally taking place. However, consumers will remain more price conscious as online retailers and high street discount stores continue to take an increasing share of the market even after the economy recovers.

The Confederation of British Industry (CBI), an employer's organisation, said that high street and online sales volumes were well above normal for the time of year, with food and clothes outlets notching up particularly strong performances. The CBI reported that a significant majority of retailers in the UK believe consumers will continue to spend freely and are planning to increase orders with suppliers at the fastest rate in almost five years.

Questions

Total: 40 marks

- 1 Use the information in Extract A to calculate the current GDP per household for the UK. **(2 marks)**
- 2 Explain the implications of the data in Figures A and B for aggregate supply in the UK since 2010. **(4 marks)**
- 3 Extract B states that investment in UK's infrastructure is 'a "win-win" option to boost growth'. With the help of an aggregate demand and supply diagram, explain why this type of investment can have such a positive impact on growth. **(9 marks)**
- 4 Using the data in the extracts and your economic knowledge, evaluate the view that increases in aggregate demand will not result in rises in the general price level. **(25 marks)**

SECTION B Essays

Total: 40 marks

In July 2015 UK exports of manufactured goods fell by 9% (or £2.3 billion) compared to the previous month; inflation in July was 0.1%. In the first half of 2015 consumption expenditure rose by 4.5% compared to 2014, and the rate of economic growth in the UK accelerated to nearly 3% annually.

- 1 Explain the possible consequences of a substantial fall in exports on the level of inflation in the UK. **(15 marks)**
- 2 Discuss the view that the level of consumption is the most important component of aggregate demand in determining the rate of economic growth in the UK. **(25 marks)**



Topic 11

Economic performance

Economic growth & the economic cycle

Key concepts from Year 1

This chapter builds on Chapter 38 of the Year 1 companion textbook, which looked at the ways in which economic performance can be judged through the analysis of data on economic growth in both the short and long runs.

- It distinguished between the two types of economic growth: short-run economic growth is an increase in output without an economy acquiring additional factors of production, whereas long-run growth is an increase in the economy's productive potential.
- It defined the trend rate of economic growth as one that is sustainable over time, and examined the nature of positive and negative output gaps which are deviations in economic growth away from this long-run trend.
- It also introduced the concept of the economic cycle, which occurs when an economy's actual level of GDP shows a regular pattern in its level of activity. We described the phases or stages of the economic cycle (such as peak and trough) and how these can be identified from the use of economic indicators.
- Finally, we explained the nature of economic shocks and analysed how they might affect an economy.

In this chapter we shall consider the benefits and costs that are associated with economic growth. In particular we will examine how growth can affect individuals, the whole economy and the natural environment. The final section will revisit the concept of the economic cycle and investigate the factors that cause economies to move through the various stages of this cycle.

Economic growth

Economic growth is normally measured by an increase in an economy's gross domestic product or GDP over some time period, normally one year. Measuring an economy's GDP is not a straightforward task, as we saw in Chapter 37. Figure 40.1 shows the annual growth in the UK's real GDP from Quarter 1 in 2007 (i.e. from the

Key terms

Economic growth occurs when over time an economy expands its capacity to produce goods and services.

Short-run economic growth is an increase in production of goods and services because the available resources within an economy are used more fully.

Long-run economic growth refers to an increase in an economy's productive potential and is what is normally meant by the term 'economic growth'.

start of January 2007) until Quarter 2 in 2015 (i.e. the end of June). It is apparent that data that have been collected require revisions, as adjustments (both up and down) have been made throughout the period.

Despite these adjustments, Figure 40.1 shows clearly that since the middle of 2013 the annual rate of economic growth in the UK, as measured by growth in real GDP, has exceeded 2%. In the first six months of 2015 the rate of growth slowed. This will not necessarily be entirely bad for the stakeholders of the UK economy, as economic growth brings costs as well as benefits.

Discussion point

What problems might be caused for the government and firms in an economy if data relating to real GDP are significantly inaccurate?

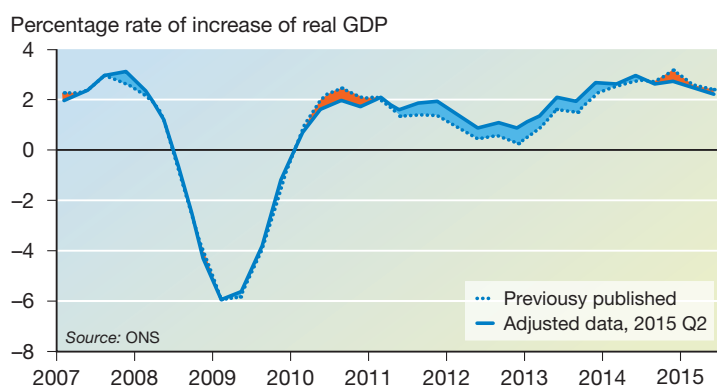


Figure 40.1 UK's previously published and adjusted real GDP growth rates, 2007–15

The benefits of economic growth

Economic growth brings a number of benefits, which is why it is an important macroeconomic objective for most governments.

1 Higher living standards

Arguably, this is the most important benefit of economic growth. A rise in real GDP can result in higher incomes for the people who live and work in the economy. In developed countries, which already have comparatively high incomes, rises in real GDP can provide access to greater quantities of goods and services for their inhabitants, for example motor cars and foreign travel, which can increase welfare levels. In these countries, rising incomes offer consumers increased access to an expanding range of luxury products.

Economic growth brings possibly even greater benefits to developing countries. It may allow the citizens of these economies access, or greater access, to vital personal services such as health and education, and well as to improved housing and diets. Crucially it can lift huge numbers of people out of poverty.

Economic growth in China has removed more people out of poverty than anywhere else in the world, principally because it has enjoyed exceptionally rapid rates of economic growth and has a large population. Its per capita income increased by a factor of 25 between 1990 and 2010 (from \$200 to \$5000). This has resulted

in China becoming a middle-income country. Data released in 2015 revealed that China has effectively eradicated urban poverty. For a country with huge numbers of poor people, this is a great success story. It has been achieved, in large part, because economic growth has provided the government with sufficient revenue to be able to subsidise urban dwellers to bring incomes up to a minimum level of 4476 yuan (about £450).

However, GDP per capita is an average figure and has many weaknesses, as we saw in Chapter 37. One danger is that it masks increasing levels of inequality in terms of income and wealth. Although average incomes in China have risen remarkably quickly, so has income inequality. A comparatively small number of very rich people have benefited disproportionately from China's remarkable economic growth.

2 Increased levels of employment

In the long-term economic growth entails an increase in the productive capacity of an economy. This means that governments and firms invest in roads, railways, offices and factories. Such investment creates new employment opportunities. Furthermore, rising incomes increase demand for services, creating a derived demand for employees in industries such as education, health, catering and hotels. In the UK in 1951 just over 40% of the workforce was employed in the service sector; by 2011 the figure had reached 80%. Over the same period the number of people in employment in the UK rose from 23.5 million to 29 million.

Number crunching

What was the increase in the number of people employed in the service sector in the UK between 1951 and 2011?

3 Improved public finances

Economic growth offers the potential to improve government finances, though this is not always realised. Growth provides governments with rising revenues from taxation as firms and households pay higher taxes on earnings and expenditure. In turn this allows governments to spend more heavily on merit goods such as education and healthcare. This can have a highly positive impact on the welfare of the country's inhabitants. At the same time the government's need to borrow to finance day-to-day expenditure may be reduced.

The disadvantages of economic growth

There has been much publicity given to the problems associated with high rates of economic growth, notably pollution. We will consider some of the major disadvantages that can occur.

1 Negative externalities

Increased levels of production that form the basis of economic growth have the potential to impose costs on third parties. Pollution of the natural environment is a common consequence and is possibly most damaging in countries in which legal controls are least effective. Pollution of land, air and water supplies can all impact heavily on the quality of life of local inhabitants.

Key term

Negative externalities

describe the problems experienced by third parties as a consequence of an economic activity. These problems can be passed on as a result of the consumption or production of a product.

Other negative externalities associated with economic growth can include congestion and noise as cities become crowded and as transport becomes more difficult and time consuming. The problems of environmental damage that can be associated with economic growth have led to many governments seeking to achieve sustainable growth rates. These are rates of economic growth that do not damage the prospects of future generations by, for example, causing irreparable environmental damage. The notion of sustainability also refers to rates of economic growth that allow governments to achieve other macroeconomic objectives simultaneously, as discussed below.

2 The possibility of conflicts arising with other macroeconomic objectives

High rates of economic growth can make it more difficult for governments to achieve other macroeconomic objectives. For example, high rates of growth may result in resources such as skilled labour becoming scarce. The outcome is likely to be increasing rates of inflation, which may damage the economy's international competitiveness as well as threatening the standards of living of its inhabitants. For much of the 1980s and early 1990s the UK economy experienced a series of 'stop-go' cycles. These were characterised by bouts of rapid growth that resulted in rising inflation, provoking governments into implementing fiscal and monetary policies to slow growth. The consequence was rising unemployment as inflation fell. Hence the economy continually stopped and started. From the mid-1990s onwards the UK economy enjoyed more sustainable levels of economic growth, with much more stable rates of inflation and unemployment.

Another problem that may result from high rates of economic growth is a worsening deficit on current account. In part this can arise because as a country's inhabitants enjoy rising real incomes, they will be able to afford more products that are produced overseas – often luxury products. Simultaneously, foreign governments, firms and individuals may be attracted by the prospect of high returns to invest in the economy. The payment of increasing amounts of dividends and interest overseas can also add to a current account deficit. The UK is susceptible to current account deficits during rapid economic growth as the economy has a high marginal propensity to import.



Air pollution in Delhi – one of six Indian cities in the top ten of the World Health Organisation's list of the most polluted cities in the world

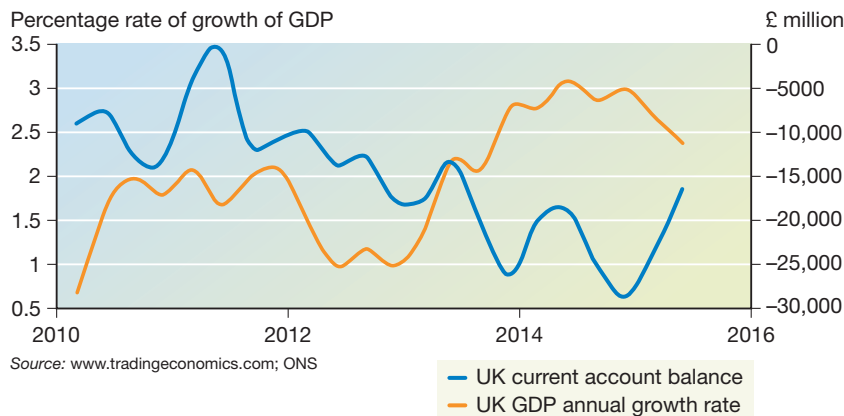


Figure 40.2 UK annual growth rate and current balance, 2010–15

Figure 40.2 shows the relationship between the UK's rate of growth of GDP and its current account balance. This suggests that as the UK's rate of economic growth has risen since 2012, the current account deficit has deepened. In part this could be due

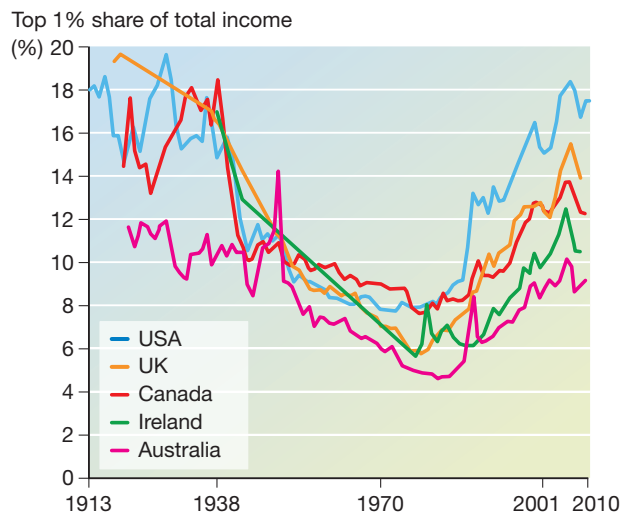
to increasing levels of foreign investment in the strengthening UK economy leading to rising outflows of dividends and interest payments as well as UK consumers and firms purchasing more from overseas.

3 Increasing income inequality

Rapid economic growth increases average incomes quickly, but this does not mean that all of a country's inhabitants benefit equally, or at all. It is not unusual for the benefits of economic growth to be concentrated in the hands of a relatively small proportion of people in an economy. The groups that benefit most tend to be skilled workers, who can negotiate substantial pay rises, and those with assets such as property, which tend to increase in value.

Figure 40.3 shows that income inequality, as measured by the proportion of total income received by the top 1% of earners, has risen sharply in a number of economies since the 1980s. This suggests that income inequality is an increasing problem and one that reduces the benefits of economic growth, at least in developed economies.

Figure 40.3 The share of income received by the top 1% of earners in a selection of economies, 1913–2010



Source: Guardian, 27.3.15

REALWORLD ECONOMICS 40.1

China's government aims for slower economic growth

China's government will target a rate of economic growth of approximately 7% in 2015 – a reduction on the growth rate of 7.5% that was pursued in 2014 and in fact the slowest rate since 1990. The country's Premier Li Keqiang made the announcement at the opening of the annual meeting of the National People's Congress.

The Chinese authorities are now keen to pursue higher quality economic growth. They have concerns about the high levels of pollution experienced by people living in certain parts of the country. Indeed, there have been outbursts of public anger over the effects of water pollution and poor quality air. The government has committed itself to the stricter enforcement

of Chinese laws on environmental protection.

Exercise Total: 13 marks

- 1 Explain what is meant by the term 'higher quality economic growth'. (6 marks)
- 2 Why do you think that the Chinese government has taken this decision at this time? Explain your reasoning? (7 marks)

The impact of economic growth

Economic growth can have a significant impact on many groups in society as well as on the environment.

The impact on individuals

Economic growth has the potential to deliver people from poverty, as we saw earlier when discussing its consequences for Chinese citizens. It offers increased opportunities for employment and the prospect of earning a level of income that allows people to enjoy some financial security as well as rising levels of welfare. In addition it gives governments greater tax revenues that can be spent on merit goods such as healthcare, further improving living standards as well as life expectancy.

In developed economies, economic growth can give people access to a range of luxury products, such as designer clothing and second homes, as well as to services, such as foreign travel.

However, these benefits are not certain for all. Some may miss out due to the unequal way in which the benefits of economic growth may be distributed amongst an economy's inhabitants. We saw in Figure 40.3 that there is evidence that the distribution of income in developed economies has become less equal since 1980. Damage to the environment can also blight the lives of those on lower incomes who are forced to live and work in industrial zones.

The impact on economies

Economic growth can transform economies. It provides funds to develop an effective infrastructure, such as a road and rail network, as well as effective communication systems. It provides funds for the government to build high-quality education and healthcare systems and to provide financial support to nations that are less developed.

South Korea is an economy that has enjoyed the benefits of a period of rapid economic growth. It was one of the world's fastest growing economies from the early 1960s to the late 1990s and has achieved impressive rates of growth more recently. In 1980 the South Korean GDP per capita was \$2300, which was about 30% of the figure for developed Asian economies such as Singapore and Japan. Since then, South Korea has become a developed economy, attaining GDP per capita of \$25,000 in 2014.

Such rapid economic growth does bring problems. Economic crises can occur. For example, the South Korean economy suffered a liquidity crisis in the 1997 Asian

Key terms

The economic cycle occurs when an economy's actual level of GDP shows a regular pattern of variation compared with its long-term trend.

The **accelerator theory of investment** states that the level of planned investment depends upon the rate of change of national income.

Herding instinct occurs when those taking financial decisions operate without sufficient thought and act in the belief that a large group of people are unlikely to be wrong.

Speculation is a high-risk financial activity in pursuit of a potentially substantial gain.

Demand-side shocks are unexpected factors that affect aggregate demand negatively or positively – for example, a significant reduction in income tax rates.

Supply-side shocks are unexpected factors that affect aggregate supply negatively or positively – for example, a major fall in the price of oil.

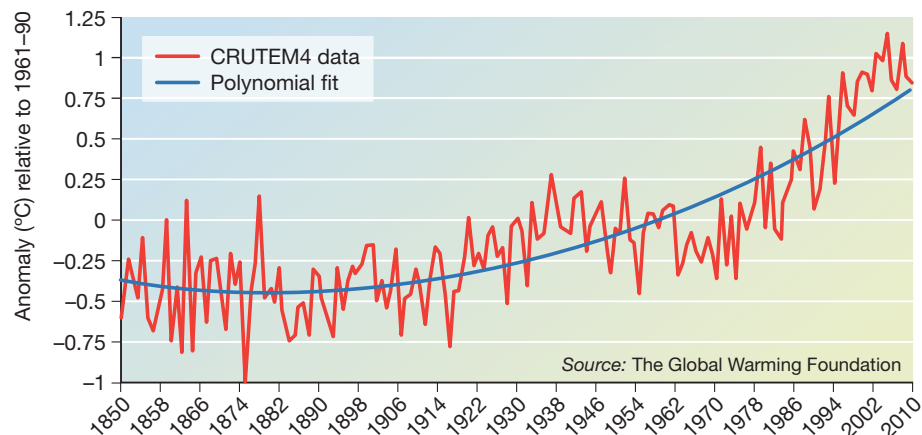
financial crisis and relied on a bailout from the IMF. However, the economy was restructured and modernised and has since recovered. Other problems can include balance of payments problems on current account, as rising incomes and production 'suck in' increased volumes of imports, and bouts of inflation if growth is too rapid.

The impact on the environment

The effects of economic growth on the natural environment have received a lot of coverage in the media. The pollution of oceans by chemicals and plastics, deforestation in the Amazon and Borneo, and the dangers of damage to the fragile arctic ecosystem from oil spillages have all attracted a lot of attention. Some of these problems have been resolved to a limited extent, but the threat to the environment remains.

Probably the biggest issue relating to economic activity and the natural environment is global warming. Most scientists, though not all, accept the threat of rising temperatures because of the emission of industrial gases, principally carbon, into the atmosphere. Economists are playing a role in attempting to alleviate the worst of the effects of economic activity on the environment through developments such as carbon trading schemes, which attempt to use market forces to control emissions of gases such as carbon. Despite these and other efforts to limit global temperatures, they appear to be rising at an increasing rate if the evidence in Figure 40.4 is correct. This shows that average global temperatures compared to those experienced in the period 1961–90 have been rising steadily since the start of the twentieth century.

Figure 40.4 Changes in global temperatures, 1850–2010



The stages of the economic cycle

The rate of economic growth in the UK, as in other economies, fluctuates over time. Since 1949 annual rates in the UK have ranged from 6.3% to –4.2%. The variations have discernible patterns. One regular pattern occurs over a period of between five and 11 years. This is the economic cycle, which has four stages as set out below.

- **Peak:** At this stage, rates of economic growth are high, which is reflected in high levels of production and consumer expenditure. Inflationary pressures will emerge.
- **Downswing:** This occurs when incomes and output start to fall along with profits and levels of investment. If output falls for more than six months, this is termed a recession.

- **Trough:** At this stage in the economic cycle, the economy's output is at its lowest. Unemployment is likely to be high, but there will be little inflationary pressure. Prices may be falling.
- **Recovery:** During this stage of the cycle, output and employment begin to recover. Households will increase consumption expenditure and spare productive capacity will be utilised before investment begins to rise.

A smoothed version of the operation of the economic cycle is shown in Figure 40.5.

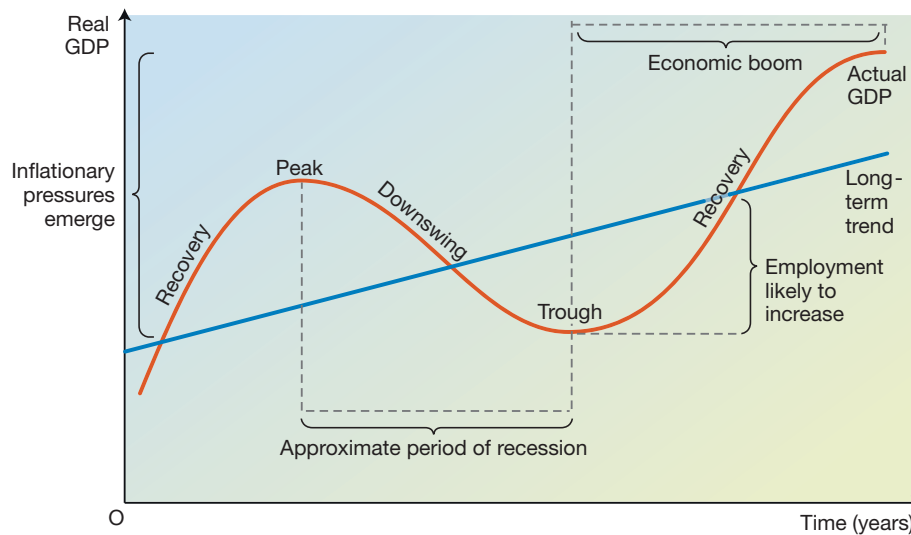


Figure 40.5 The pattern of economic activity associated with the different stages of the economic cycle

The causes of changes in the economic cycle

An economy moves into a different stage of the economic cycle due to a range of factors (Figure 40.6). Some of these originate within the economy itself and are the result of the government's economic policies or other domestic events. These are termed endogenous causes. However, as the process of globalisation makes the world's economies more interdependent, these causes are increasingly likely to be external or exogenous, such as demand-side shocks.

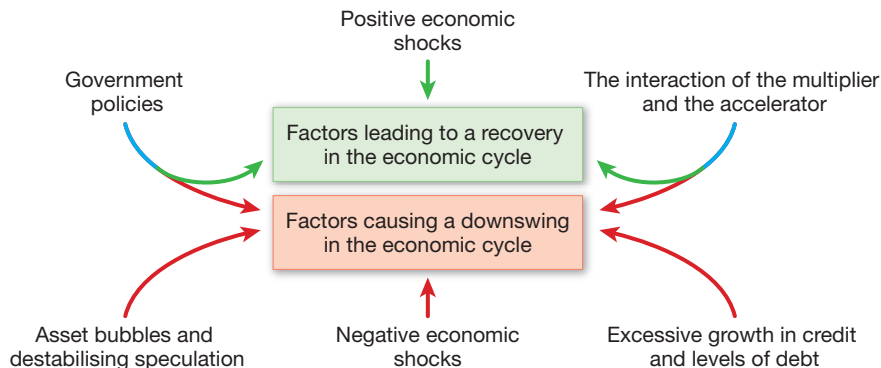


Figure 40.6 The causes of changes in the economic cycle

1 Economic shocks

Economic shocks take a number of different forms and have the potential to have a major impact on the economic performance of economies, not least on the rate of

economic growth. Demand-side shocks impact on the level of aggregate demand, whereas supply-side shocks affect aggregate supply.

Negative shocks

Negative demand-side and supply-side shocks have the potential to slow rates of economic growth or even to provoke a recession. The financial crisis in 2008 was a global demand-side shock that affected many economies. The crisis in the UK comprised several factors that included falling house prices, a mortgage crisis and declining household wealth, which led to a dramatic drop in consumer spending and precipitated the onset of a deep recession. Some negative demand-side shocks might be primarily constrained to a single economy, though the effects might ripple outward through changing patterns of international trade and investment. Such a domestic demand-side shock might be caused by a substantial rise in tax rates. For instance, in Greece the rate of sales tax (its equivalent of VAT) was increased from 13% to 23% on many products in July 2015 as the government attempted to improve its budget balance. Levels of consumption in Greece have fallen as a consequence.

Negative supply-side shocks can come about from a range of causes, including sudden increases in the price of commodities as a result of natural disasters. These reduce aggregate supply and can lead to an economic downturn.

Positive shocks

On the other hand, positive demand-side shocks can contribute to an economy experiencing high rates of economic growth by providing a substantial boost to aggregate demand. These may be caused by a range of factors, including unexpectedly high rates of economic growth in a major economy and trading partner. This could lead to a global economic shock because a number of economies are likely to benefit. Economic growth rates would be expected to increase, moving economies into the upswing stage of the trade cycle.

Positive supply-side shocks are less common but could result from a sudden and significant decline in commodity prices, such as oil. This is an example of a global supply-side shock.

2 Asset price bubbles and destabilising speculation

High rates of economic growth can often result in rapid rises in the price of assets such as property or company shares. Over a period of time, these assets increase in price above their real values as households and firms have exaggerated expectations of their future worth. These unrealistic expectations increase the volume of sales, buyers are more numerous than sellers, and prices rise above the levels suggested by any objective analysis. This behaviour is called 'herding' or 'herd behaviour' and contributes to large changes in the level of national income within economies.

Eventually, of course, the bubble bursts. The buyers of the assets disappear and sellers find that prices fall sharply, or possibly collapse. Households and firms that hold assets that have fallen sharply in value may cut spending, and banks may have advanced loans to purchase these assets that cannot be repaid. As a result, spending in the economy declines and the economic cycle moves into a downturn.

A well-known example of herd behaviour took place over the period 1997–2000

when businesses and individuals invested heavily in internet-based companies. These companies were often relatively small and did not own many assets or have viable business models. The so-called ‘dot-com’ bubble collapsed in 2000–01. Some companies failed entirely and ceased trading; others saw their share prices collapse (Amazon’s fell from \$107 to \$7) but survived.

Speculation is a high-risk financial activity in pursuit of a potentially substantial gain. Investors have frequently invested in what are risky enterprises in the hope of earning high returns. However, such speculation can destabilise markets by driving prices strongly upwards or downwards and thus contribute to upswings or downswings. At the time of writing, stock markets across the globe have seen their heaviest falls in share prices for four years. This might be a necessary correction, or it could be that the fall in prices is too great, reflecting unwarranted pessimism by investors, and it may contribute to a downturn in the global economy.

Why do people behave in this way? Theories of behavioural finance suggest that people invest so as to avoid missing out on a great opportunity that others are enjoying. Once the high prices fall, as inevitably is the case, most short-term investors leave the market and behaviour becomes more normal and based on prudent analysis once more.

John Maynard Keynes used the term ‘animal spirits’ in *The General Theory of Employment, Interest and Money* in 1936 to describe the human emotions that drive financial decisions such as those to invest in a company or other asset: ‘Even apart from the instability due to speculation, there is the instability due to the characteristics of human nature... that depend on spontaneous optimism rather than mathematical expectations... Most, probably, of our decisions to do something positive... can only be taken as the result of animal spirits – a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities’.

3 Excessive growth in credit and levels of debt

Excessive growth in credit and levels of debt can be expected to lead to a situation whereby some households and firms are unable to repay what they owe. This can result in banks suffering a liquidity crisis and they may be unable to meet the needs of customers who wish to withdraw cash. Firms and households will cut spending as they are unable to receive credit and, at worst, a loss of confidence in the banking system can occur.

A global financial crisis preceded the severe and prolonged recession experienced by many economies in the period after 2008. It was caused, in part, by rapid increases in the amount of credit offered to people who were unable to repay the debt. The expansion of ‘subprime debt’ was used to provide mortgages to homeowners on low incomes in the United States. This debt was attractive to banks in the USA and elsewhere because of the high interest rates charged. However, it was also the cause of the crisis in that low-income households could not afford the repayments and many defaulted. This led to a banking crisis in 2008, with significant disruption of the flow of credit to businesses and consumers.

This subprime debt was sold on to banks and financial institutions in other countries, resulting in the spread of the crisis through many parts of the global economy. The outcome was the onset of a severe global recession.

Review questions

Total: 32 marks

- 1 Explain the difference between short-run and long-run economic growth. **(5 marks)**

- 2 Which of the following is least likely to be a benefit of economic growth?
 - A An increase in the standard of living
 - B Increased investment in environmental protection
 - C Increased income equality
 - D Increased levels of employment**(1 mark)**

- 3 Explain why increasing GDP per capita might not result in rising living standards for all of a country's inhabitants. **(5 marks)**

- 4 Which of the following is the most likely consequence of high rates of economic growth in an economy?
 - A An increase in the economy's rate of deflation
 - B An increase in the surplus on the current account of the economy's balance of payments
 - C An increase in the economy's budget deficit
 - D An increase in the economy's employment rate**(1 mark)**

- 5 Define the term 'negative externalities'. **(3 marks)**

- 6 Explain one benefit to an individual person of an increase in the rate of economic growth in the country in which they live. **(4 marks)**

- 7 Distinguish, with the aid of examples, between positive demand-side shocks and positive supply-side shocks. **(5 marks)**

- 8 Explain how the interaction of the multiplier and accelerator models can be a cause of the economic cycle. **(6 marks)**

- 9 Which of the following can result in asset bubbles occurring?
 - A Herding instinct
 - B A loss of consumer confidence
 - C A loss of business confidence
 - D Positive supply shocks**(1 mark)**

- 10 Which of the following is regarded by many economists as the major cause of the financial crisis of 2008 and the subsequent recession in many countries?
 - A The interaction of the multiplier and accelerator
 - B Excessive growth in credit and levels of debt
 - C A negative supply-side shock
 - D Government policy with regard to taxation**(1 mark)**

Employment & unemployment

Key concepts from Year 1

This chapter builds on Chapter 39 of the Year 1 companion textbook.

- It looked at the ways in which governments can measure the level of unemployment. In the UK the claimant count and the Labour Force Survey methods are used for this purpose.
- It identified and described the different types of unemployment that can occur including seasonal, frictional, structural and cyclical – we shall add to these in this chapter.
- Finally, it examined the ways in which supply-side and demand-side factors can determine the levels of employment and unemployment and how changes in the rest of the world can affect employment in the UK.

This chapter distinguishes between the concepts of voluntary and involuntary unemployment. It also introduces the notion of real wage unemployment and the natural rate of unemployment, and examines the factors that determine them. Finally, we consider the consequences of unemployment for individuals as well as for the performance of the economy.

Involuntary unemployment

The types of unemployment set out in the key terms list on this page are normally all involuntary. This occurs when workers lose their jobs and leave the labour market against their wishes. There are a number of different types of involuntary unemployment, each with different causes. It is important for governments to identify the precise cause of unemployment, as far as is possible, in order that the appropriate solutions can be applied. Thus, for example, if an economy is suffering from high levels of frictional unemployment, the government may seek to improve information about vacancies and available employees in order to assist people in finding jobs and firms in finding employees.

The concept of involuntary unemployment was first identified by Keynes in *The General Theory of Employment, Interest and Money* in 1936, though its existence is

Key terms

Unemployment exists when people are seeking work but are unable to find it.

Cyclical unemployment is unemployment arising from a fall in aggregate demand.

Frictional unemployment exists when workers are in the process of moving to a new job.

Involuntary unemployment occurs when workers leave

the labour market against their wishes.

Structural unemployment is the loss of jobs resulting from the long-term decline of specific industries.

Seasonal unemployment is unemployment that exhibits regular and predictable fluctuations throughout the year.

not accepted by all economists. Classical economists believe that, in the long term, unemployment will solely comprise voluntary unemployment because all those people in an economy who want a job at the current wage rate in the labour market will find one.

Voluntary unemployment

Sometimes unemployment may occur because people do not wish to find a job, choosing not to enter the labour market. They therefore remain unemployed. Voluntary unemployment can occur because the current wage rate is not judged to be sufficient compensation for giving up leisure time.

Voluntary unemployment can also occur because welfare benefits available to unemployed workers are too high in relation to the wage rates currently on offer. Some economists argue that governments should reduce the level of voluntary unemployment by lowering the benefits paid to unemployed workers to reduce the disincentive to entering employment. Research in 2011 by Andrew Dunn of the University of Lincoln suggested that some long-term unemployed in the UK may not be prepared to apply for jobs that they consider ‘unattractive’, preferring to claim Jobseeker’s Allowance (JSA) – the unemployment benefit which was paid at the time. He states: ‘All of the 40 employees of welfare-to-work organisations I interviewed in 2011 said that many of their long-term (i.e. over 6 months, often several years) JSA claimant clients remained unemployed because they were very “choosy” in the jobs they were willing to do; most of the 40 said they believed that a majority of these clients would enter employment within two months if they applied for a range of relatively unattractive jobs’ (LSE Blog, 22.7.15).

The UK government has recognised this cause of voluntary unemployment. It has increased the number of conditions attached to receiving unemployment benefits and has also introduced a new benefit for the unemployed, which is called ‘Universal Credit’.

Real wage unemployment

Real wage unemployment exists when the real wages for workers in an economy are too high and above the equilibrium or market clearing wage rate. The situation results in firms being unwilling to employ every person looking for a job. This type of unemployment is also called classical unemployment.

When real wages are too high, it means that the cost of employing an extra worker is higher than the benefit received from the value of output the worker produces. So, when real wages are too high in an economy, firms cannot profitably employ all the labour on offer. As a result, some of the economy’s supply of labour is not used. The consequence is real wage or classical unemployment, and this type of unemployment

Key terms

Voluntary unemployment occurs when workers take a decision not to enter the labour market at the current wage rate.

Real wage unemployment exists when the real wages for workers in an economy are too high, leaving firms unwilling to employ everyone who is looking for a job.

A **trade union** is an organisation formed with the objective of enhancing and protecting the working conditions and economic position of its members.

The **trade union wage premium** is the percentage difference in average gross hourly earnings between union members and non-members.

is illustrated in Figure 41.1. With a real wage rate of W_2 , the supply of labour is Q_s , while the demand is lower at Q_d . The difference between these two levels ($Q_s - Q_d$) represents the level of real wage or classical unemployment. This unemployment could be eliminated if the real wage rate fell to W_1 at which point the demand for and supply of labour would be equal.

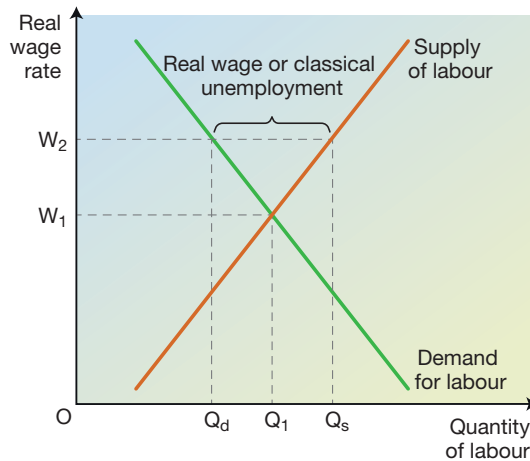


Figure 41.1 Real wage or classical unemployment in a labour market

Classical economists place great emphasis on this type of unemployment and emphasise that if labour markets were to operate more freely, then it could be reduced.

The causes of real wage unemployment

1 Wage 'stickiness'

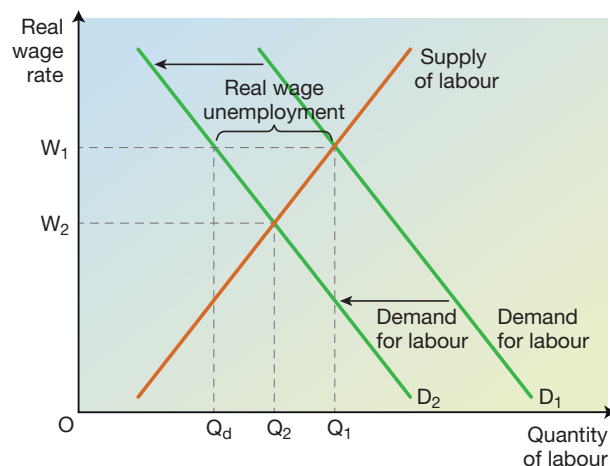


Figure 41.2 Wage stickiness and real wage unemployment

The fundamental cause of this type of unemployment is wage 'stickiness' – the inability of real wage rates to adjust to changes in the demand for or supply of labour. This lack of flexibility is most significant in a downward direction. The demand for labour is a derived demand. Workers are only needed because there is a demand for the goods and services that they can produce. If there is a fall in demand for goods and services, for instance as a result of a negative demand-side shock, the demand for labour will be lower. This is illustrated in Figure 41.2 by the shift in the demand curve for labour from D_1 to D_2 . In an entirely free market, real wages would adjust

smoothly from W_1 to W_2 , ensuring that the market clears and that no real wage unemployment results. However, if wages are not flexible downwards, then real wage unemployment is likely to occur. In this case it would be $Q_1 - Q_d$, if wages did not adjust downwards to any extent.

Prime causes of real wage 'stickiness' are the activities of trade unions and the implementation of minimum wages.

2 The activities of trade unions

Trade unions negotiate wage rates on behalf of their members. They have professional negotiators, bargain on behalf of a large number of people (termed collective bargaining) and may threaten strike action. This gives them considerable power to increase wage rates above the level that would operate in their absence.

In 2013 the Department for Business, Innovation and Skills in the UK estimated that the trade union premium (i.e. the percentage difference between the wage rates of union and non-union employees) was 19.8% for employees in the public sector and 7.0% for private-sector employees. This demonstrates that trade unions do have the potential power to hold wages above the free market level and can cause real wage unemployment. Trade unions will also play a significant part in preventing wages from falling at a time when the demand for labour decreases, or its supply increases.

3 The imposition of minimum wage rates

Many governments in developed economies have passed laws to enforce minimum wage rates. These are often expressed in hourly terms. The UK is no exception in this regard. A minimum hourly wage rate was established in April 1999 and has been raised each October since then.

In 2015 the Chancellor of the Exchequer, George Osborne, announced the introduction of the living wage. This was introduced in April 2016 at a rate of £7.20. It operates alongside the minimum wage but applies to employees aged over 25. It represents a significant premium on the minimum wage, which continues to be paid to those aged 21–25 at a rate of £6.50 an hour. The living wage in the UK will be raised in increments to £9 an hour by 2020.

The Office of Budget Responsibility has predicted that the introduction of the living wage will lead to a loss of 60,000 jobs in the UK. This figure may provide some indication of the extent of the real wage unemployment associated with the government's intervention in the UK's labour market.

Key terms

The **natural rate of unemployment** is the level of unemployment that exists when the labour market is in equilibrium.

Labour mobility

measures the ease with which labour moves from one productive use to another.

The natural rate of unemployment

In all economies, even ones that are growing quickly and evolving new industries, some level of unemployment will always exist. This level of unemployment is made up of two types of unemployed people.

- **Structurally unemployed:** Some industries may go into terminal decline, perhaps because of technological change, or owing to a lack of competitiveness. Their demise will create structural unemployment as workers with obsolete skills are released onto the labour market.
- **Frictionally unemployed:** There are always some people who are between jobs, possibly taking a break, and are thus frictionally unemployed.

Based on the notion that some unemployment is always likely to exist, Milton Friedman and Edmund Phelps, two American economists, developed the theory of the natural rate of unemployment. The natural rate of unemployment is the level of unemployment that exists when the labour market is in equilibrium. This is illustrated in Figure 41.3.

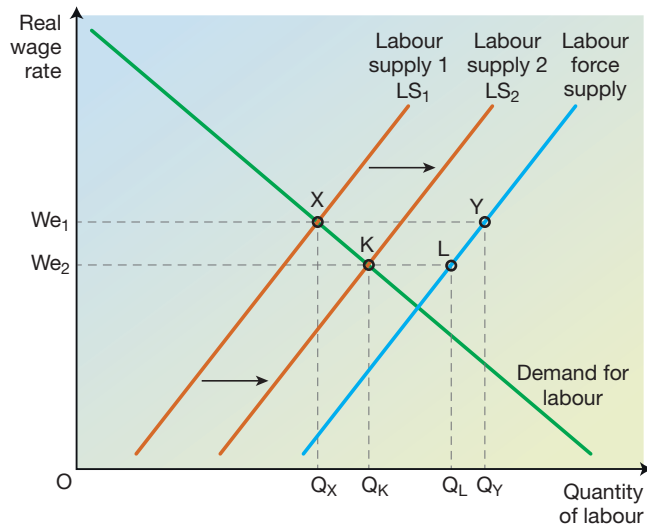


Figure 41.3 The natural rate of unemployment and the labour market

In Figure 41.3 the labour force supply curve represents the level of labour supply if all of the workforce is to be employed. This would occur if there were no frictional or structural unemployment and would mean that the natural rate of unemployment is zero. This is not achievable, but it does provide a yardstick against which to measure employment levels. If the current labour supply in the economy is LS_1 , then the natural level of unemployment is shown by XY , or by Q_xQ_y on the x-axis. In order to express this as a rate, it would be necessary to calculate the natural level of unemployment as a proportion of the labour force.

Factors determining the natural rate of unemployment

The natural rate of unemployment can be determined by a range of factors that affect the efficiency with which the labour market works. Those economists who believe in the absolutely free working of markets would argue that unemployment is a signal of market failure. Thus, if labour markets worked perfectly, then wages would fall to a low enough level to eliminate unemployment.

1 The extent of labour mobility

Labour mobility measures how easily labour moves from one use to another. This might entail the movement of employees from one occupation to another, and therefore reduce the amount of structural unemployment that exists within an economy. It may also refer to geographical mobility, whereby it measures how easy it is for workers to move from one area to another to take up new employment.

If labour in an economy is immobile, this can contribute to structural unemployment as workers leaving a declining industry experience difficulties in finding new employment. It will also add to frictional unemployment because workers take longer to find new

jobs. Inadequate training schemes may result in labour immobility, as unemployed workers cannot acquire the skills needed in those industries where employment is growing. High house prices in many parts of the UK are also seen as contributing to the geographical immobility of labour because employees may be unable to afford housing costs in locations such as the Southeast, where job vacancies may exist.

2 Inadequate labour market information

If unemployed workers are unaware of vacancies, or firms seeking workers do not have information about suitably skilled employees, then the level of frictional unemployment is likely to be higher than otherwise might be the case. Most governments provide extensive labour market information as they benefit from lower rates of unemployment and, of course, from a reduction in the natural rate of unemployment. The UK's Department for Work and Pensions provides labour market information through its network of Jobcentre Plus offices and its Universal Jobmatch, which is an online service.

3 Lack of training and suitable skills

If employees in an economy lack the skills required by firms that have vacancies, then they are unlikely to be hired unless the firms concerned are willing to accept potentially high training costs. Having an efficient national system of training can help to move employees relatively swiftly and smoothly from industries that are declining to those that require more employees. Thus the levels of structural and frictional unemployment can be reduced.

4 The activities of trade unions

Trade union activities can restrict the supply of labour into certain industries and therefore shift the overall labour supply to the left. They may negotiate to shorten weekly working hours or organise strikes or other forms of industrial action that reduce the labour supply, which shifts the supply curve for labour leftwards.

Reducing the natural rate of unemployment

If governments could make labour markets operate more efficiently and reduce factors that contribute to labour market failure, then the result would be a reduction in the natural rate of unemployment as well as a fall in the wage rate. In Figure 41.3 this is shown by a shift in the supply curve of labour from LS_1 to LS_2 . This would reduce the real wage rate to We_2 and the natural level of unemployment is reduced to $Q_K Q_L$, which would represent a lower natural rate of unemployment.

NAIRU

NAIRU is an acronym that stands for the non-accelerating inflation rate of unemployment. It refers to a rate of unemployment that is associated with stable rates of inflation. NAIRU was first proposed in 1975 as an improvement on the concept of the natural rate of unemployment. We shall look at NAIRU in more detail in Chapter 43.

The consequences of unemployment

A low rate of unemployment is an important macroeconomic objective for most

governments because it has a number of undesirable consequences for individuals, firms and macroeconomic performance.

Consequences for individuals

There are a number of severe social consequences for individuals arising from periods of unemployment, especially if unemployment rates are high and persistent. One primary consequence is that unemployed people suffer a fall in income, which potentially can be very large. This leads to a substantial fall in living standards and can widen inequalities in income and wealth between different groups in society. The consequences of unemployment for individuals will be much more severe if they are out of work for a long period of time.

High rates of unemployment also have consequences for those who are in employment. In such circumstances employees and trade unions are in a weaker bargaining position because a falling demand for labour will tend to depress real wage rates with adverse effects on those in work. Real wages fell for a sustained period in the UK after the financial crisis in 2008 at a time when the rate of unemployment rose from about 5% to 8.5%.

If unemployment is concentrated in a single area, which is often the case with structural unemployment, high levels of deprivation can result. High levels of unemployment and deprivation in particular locations in the UK have been linked with increasing levels of crime, poorer standards of health and a decline in life expectancy. Those individuals who worry about their financial security as a result of becoming unemployed frequently suffer poor health. The long-term unemployed often find that their skills become outdated and that they lose confidence in their ability to perform effectively at work.

REALWORLD ECONOMICS 41.1

Unemployment causes severe psychological damage

The psychological damage caused by unemployment is greater than previously thought, according to a study led by researchers at the University of Stirling. Personality is typically considered stable across time, but the researchers found that the experience of unemployment led to reduced levels of conscientiousness, agreeableness and openness, signifying that individuals lose motivation, become less considerate and sympathetic, and less curious about the world around them. These changes were greater the longer an individual spent unemployed.

Participants in the study completed personality tests at two time-points, four years apart. All participants were employed at the time of the first test. At the time of the second test, they had either remained in employment, been unemployed for one to four years, or were re-employed after a period of unemployment. The results showed that compared with those who had remained in employment, unemployed people experienced significant patterns of change in their agreeableness, conscientiousness and openness. Re-employed individuals

experienced limited change.

The study suggests that the effect of unemployment across society is more than just an economic concern. The unemployed may be unfairly stigmatised as a result of unavoidable personality change, potentially creating a downward cycle of difficulty in the labour market.

Source: University of Stirling

Exercise

Explain how the psychological effects of unemployment might damage the macroeconomic performance of an economy.

(7 marks)

Research by the University of Zurich has suggested that there is a link between unemployment and the rate of suicide. It found that between 2000 and 2011 across 63 countries, one in five of an estimated 233,000 annual suicides were linked to unemployment.

There can be positive consequences of unemployment for some individuals, however. It may be that they are prompted by redundancy to retrain for jobs that provide more security and higher incomes in the long term. They may find that their retraining is funded by the UK government or by the business as part of the redundancy package. Other unemployed individuals may use redundancy payments to fund the start-up of a small business. If successful, this can confer considerable benefits on the entrepreneur, such as the pleasure of being one's own boss.

Author tip

It is very easy to think about the consequences of unemployment in just negative terms. However, there are positive aspects and examination questions on this topic can be evaluative, requiring you to offer a balanced account.

Consequences for the performance of the economy

Periods of unemployment have profound negative implications for an economy. They mean that the economy is not utilising all the resources that are available to it and that the level of national income is reduced as a consequence. The economy will be operating within its production possibility boundary. The impact of a severe bout of unemployment can reduce consumption and investment as households and firms experience a loss in confidence. The outcome is likely to be a fall in aggregate demand.

The government's finances are also likely to be damaged by a period of unemployment. Its receipts from taxation will decline as fewer people are paying income tax and national insurance on their earnings. A fall in consumption will reduce revenues from value added tax (VAT). Simultaneously, the government's expenditure will rise as more people will claim benefits (such as universal credit) and it may also have to increase its spending on training.

Key term

Hysteresis describes a situation in which periods of high unemployment tend to increase the rate of unemployment, below which inflation begins to accelerate, commonly referred to as NAIRU.

The word hysteresis is derived from the Greek term meaning 'a deficiency'. In economics, hysteresis is a situation that arises when any historical event affects the future economic path – it is frequently used to describe the consequences for an economy of unemployment. In a sustained period of high unemployment, people become accustomed to a lower standard of living. Over time they may be less willing to take actions to pursue their previous living standards. Furthermore, being unemployed becomes less unusual and less unacceptable to others in society. Once the economy recovers from its period of high unemployment, some people may not be willing to re-enter the labour market. One outcome of this is that the economy's NAIRU will tend to rise.

However, there are some benefits to an economy arising from unemployment. Real wage rates can fall, increasing the competitiveness of businesses because they may be able to reduce their unit labour costs. Rates of inflation may be lower as a consequence. Businesses that are seeking to expand will have a larger potential workforce available and these workers may already be trained, reducing start-up

costs and the viability of the new business activity. There may also be environmental benefits arising from high unemployment because growth rates will slow and the extent of negative externalities, such as air and water pollution, may be reduced.

Review questions

Total: 32 marks

- 1 Distinguish between voluntary and involuntary unemployment. **(5 marks)**
- 2 Draw a diagram to illustrate how real wage or classical unemployment can occur in a labour market. **(5 marks)**
- 3 Which of the following is *not* a cause of real wage unemployment in a labour market?
 - A The trade union premium
 - B The minimum wage rate
 - C Wage stickiness
 - D Hysteresis**(1 mark)**
- 4 Which of the following types of unemployment form a part of the natural rate of unemployment: (i) structural unemployment, (ii) cyclical unemployment, (iii) frictional unemployment?
 - A (i) only
 - B (i) and (iii)
 - C (ii) and (iii)
 - D (i) and (ii)**(1 mark)**
- 5 Explain how the immobility of labour contributes to the natural rate of unemployment. **(5 marks)**
- 6 What is meant by the acronym NAIRU? **(2 marks)**
- 7 Which type of unemployment exists when the labour market is in equilibrium?
 - A The natural rate of unemployment
 - B Classical unemployment
 - C Real wage unemployment
 - D Long-term unemployment**(1 mark)**
- 8 What is meant by the term 'hysteresis'? **(2 marks)**
- 9 Explain one reason why unemployment might have positive consequences for an individual. **(4 marks)**
- 10 Explain why a government's finances may be weakened by a rise in the level of unemployment. **(6 marks)**

Inflation & deflation

Key concepts from Year 1

This chapter builds on Chapters 33 and 40 of the Year 1 companion textbook.

- In Chapter 33 we looked at the operation of index numbers and the benefits they offer to economists. We saw how weighted index numbers are used in the UK to measure the rate of inflation – for example, in constructing the Consumer Prices Index or CPI.
- In Chapter 40 we distinguished between inflation, disinflation and deflation and explored the factors (both supply side and demand side) that can cause changes in the rate at which prices rise or fall. We considered causes arising in the domestic economy as well as global factors, such as changes in commodity prices and events in other economies.

This chapter introduces the quantity theory of money. It also explores the Fisher equation of exchange and explains how it relates to the rates of inflation and deflation. We link this theory to the monetarist interpretation of the relationship between the money supply and the price level. We examine the effects of expectations on changes in an economy's price level. The chapter also considers the effects of inflation and deflation on individuals as well as the performance of the economy. Finally, we investigate how changes in the global prices of commodities can affect domestic inflation.

The quantity theory of money and Fisher's equation of exchange

The notion that there is a relationship between the quantity of money in an economy and the rate at which prices change has existed since at least the sixteenth century. Economists including Karl Marx and David Hume wrote about it. The theory states that there is a direct relationship between the quantity of money circulating in an economy and the general level of prices. The theory expresses the belief that if the quantity of money in an economy rises by 50%, the outcome will be an increase in prices of 50%.

Irving Fisher was an American economist and statistician who made a number of important contributions to the field of economics. One contribution, which bears his

Key terms

Inflation is the rate of increase of the general price level and the corresponding fall in the value of money.

Disinflation occurs when prices are rising, but at a decreasing rate.

Deflation is the rate of decrease of the general price level and the corresponding rise in the value of money.

The **quantity theory of money** states that increases in the money supply within an economy will result in increases in the general level of prices.

Fisher's equation of exchange is an equation, or arguably an identity, which sets out the relationship between money, the frequency with which it is spent and the value of transactions in an economy.

name, was the equation of exchange, developed in the 1920s. This equation expresses the relationship between the quantity of money and value of transactions:

$$MV = PT$$

where M is the supply of money, V is the velocity of circulation that measures the average frequency with which each unit of money is spent, P is the average price level, and T is the volume of transactions that take place in the economy.

Some economists have argued that this is not really an equation, but an identity in that it is two ways of saying the same thing. They believe it is an identity because both elements measure the value of transactions in an economy over a period of time. As a consequence, the formula is sometimes presented as shown below:

$$MV \equiv PT$$

The symbol \equiv denotes that this is an identity, meaning that it is true by definition.

Fisher's equation suggests that the value of money (or the price level) can be determined by the amount of money in the economy. Thus, an increase in the money supply may result in it losing value.

Intermediate transactions and the equation of exchange

One problem with the equation of exchange as proposed by Fisher is that the use of T (the number of transactions) as part of the relationship entails the double counting of intermediate transactions. For example, it includes transactions relating to the purchase of car components as well as the transaction in which the car was purchased by the consumer. As we saw in Chapter 34 of the Year 1 textbook, national income measures the value of final output that occurs within an economy. Thus, to be accurate, the exchange equation should only include transactions relating to final goods and services produced within an economy or to nominal national income. The revised equation is therefore as follows:

$$MV = PQ$$

where Q is the number of transactions relating to the purchase of final goods and services.

Monetarism and the quantity theory of money

Milton Friedman, Professor of Economics at the University of Chicago, restated the quantity theory of money in 1956. He contended that although 'the quantity theory is in the first instance a theory of demand for money', the supply of money in the economy does influence the price level. Friedman stated that the supply of money was unstable due to actions of organisations such as central banks. An increase in the money supply would result in some households having larger quantities of cash. Friedman argued that they would seek to reduce their holdings of cash by, in part, increasing spending on goods and services. Thus he asserted that an increase in the money supply will lead to a rise in the price level or income or both, assuming that the demand for money remains constant: 'Inflation is always and everywhere a monetary phenomenon'.

Friedman believed that changes in the supply of money can influence the level of real national income, but only in the short term. His work, and the work of others, gave rise to the monetarist school of economics.

Key term

Monetarism is a school of thought that supports the close control of the money supply based on the belief that increases in this supply result in inflation.

In terms of the quantity theory, most monetarists believe that the velocity of circulation might fluctuate in the short term but that it only changes slowly in the long term. Thus, if the velocity of circulation is assumed to be relatively constant in the long term, then the money supply should be permitted to grow roughly in line with the normal rate of economic growth for the economy. A faster rate of increase in the money supply will otherwise simply translate into a rise in prices.

Thus, the monetarist group of economists believes that changes in the money supply affect the level of prices in the economy in the long run. As a consequence, these economists argue that governments should seek to control the money supply rather than use elements of monetary policy, such as interest rates, to achieve macroeconomic objectives.

Monetarism gained popularity during the 1970s and early 1980s and influenced government economic policies, especially in the UK and USA. Since that time it has declined in popularity and money supply targets were abandoned in the UK in 1984. The link between the money supply and inflation was not clear, and it proved difficult to control the supply of money effectively in a modern economy.

Expectations and the price level

Expectations play a significant role in economic theory, including those relating to inflation and deflation. Decisions by consumers, firms and governments can all affect future rates of inflation or deflation and many decisions will be affected by expectations.

- **Wage settlements:** Employees and trade unions will push for percentage wage increases to at least match their expectations of future inflation, thereby protecting real wages and future living standards. Firms may be more willing to agree wage claims if they believe that prices will rise in the future, offering the opportunity to pass on higher costs to consumers.
- **Pricing decisions:** One influence on the prices that are set by firms will be expected rates of inflation. If prices are expected to rise by, say, 2.5%, firms will feel more confident in raising prices to at least this extent without risking the loss of too many sales. This helps to protect profit margins.

Expectations of inflation are influenced by past and current rates of inflation, rates of price increase experienced by major trading partners as well as forecasts for the economy's macroeconomic performance. If future growth rates are expected to be high and rising, for example, this is likely to lead to expectations of upward pressure on prices.

The Bank of England has recognised the importance of expectations in decision making by individuals and firms and is trying to influence these by a policy known as 'forward guidance'. Since August 2013 the Bank has offered this guidance on its monetary policy (for example, the likely future trend of interest rates) to influence business and consumer expectations and decisions.

There are two different types of model that use expectations in relation to changes in the price level as well as other economic indicators.

Adaptive expectations

The theory of adaptive expectations places importance on past events in predicting future outcomes. It is frequently used in models of inflation. Such models state that if

Key term

Adaptive expectations are theories that state that firms and households use past information as the best indicator of future events.

inflation has increased in the past year, households and firms will expect a higher rate of inflation the following year. This may result in decisions relating to wage claims and the setting of prices that reflect that higher expected rate of inflation. These decisions result in actions that can generate the expected rate of inflation.

For example, if inflation in the past year had been forecast to be 2% but actually reached 5%, this may prompt expectations that it will continue at this higher rate in the following year. Consequently, decisions relating to wage claims and settlements, and prices to be charged in the future, will reflect these changed expectations. The outcome is likely to be a higher rate of inflation – perhaps around 5%.

This theory is often criticised as simplistic. Most decision makers in an economy do not believe that the future will merely reflect what has happened in the recent past. In response to this criticism, more complex models of adaptive expectations have been developed. These incorporate weights to the model to reflect different rates of inflation in previous years and the extent to which they differed from forecast rates.

We shall consider the role of expectations further in the next chapter when we look at the Phillips curve.

Rational expectations

In economics, rational expectations theory assumes that decision makers base their judgements on all the information that is available to them. This will include historical data but also other factors, such as any relevant current and forecast data, which might provide understanding of future economic performance.

Rational expectations theory assumes that individuals and firms can predict future economic conditions with some accuracy and take decisions based on their predictions. It may be that the relationships between causes and outcomes are not understood fully, but their own experience and analysis of a range of data allow decision makers in an economy to develop expectations that are rational in terms of the information available to them. Such rational decision making is often proved to be correct, although random and unexpected factors can result in errors. Rational expectations models do tend to assume that decisions makers have access to large amounts of information to assist them in making good quality decisions.

Key term

Rational expectations theories are based on the assumption that decision makers use all available information – past, current and forecasted – before reaching judgements.

REALWORLD ECONOMICS 42.1

UK inflation expectations decline

The public's expectations for inflation in the UK over next 12 months fell back slightly in October, according to a survey. The monthly YouGov/Citi survey found inflation expectations for the next year slipped to 1.4% from 1.5% in September 2015, and have been below the Bank of England's 2% target for 13 months now, which is the longest period in over a decade.

Expectations for UK inflation over the next five to 10 years were stable at 2.7%.

Previously published official figures showed the UK economy to be experiencing deflation: consumer prices fell by 0.1% over the year to September 2015.

YouGov surveyed 2031 people in October 2015.

Source: Adapted from Reuters

Exercises

Total: 14 marks

- 1 Explain why the public's expectations regarding inflation might affect the future rate of inflation in the UK. (7 marks)
- 2 Why might people expect inflation of 1.4% over the next year when the current rate is below zero? (7 marks)

The consequences of inflation and deflation

The consequences for individuals

Inflation

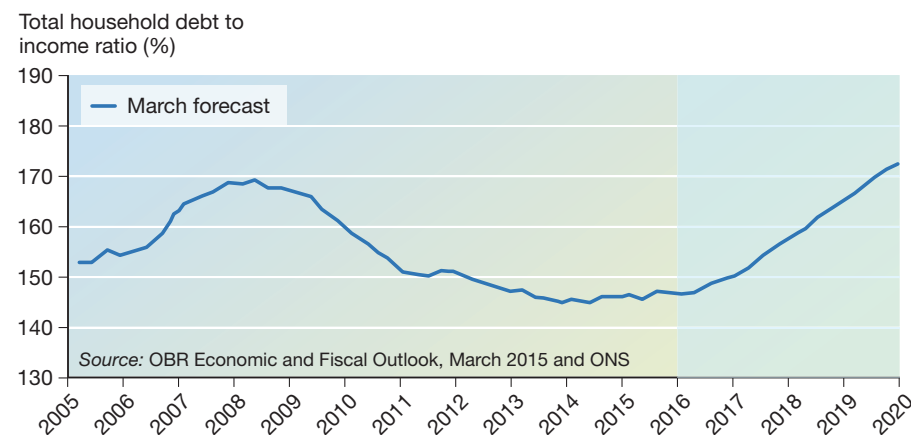
The impact of inflation on individuals depends upon a range of factors including the rate of inflation that is experienced and their ability to negotiate increased incomes.

- **A loss of purchasing power and redistribution of income:** Inflation reduces the purchasing power of money and therefore the value of incomes and savings. It can do huge damage to the standards of living of weaker, low-income groups in society because they are often unable to command a rate of increase in income to compensate for the increase in prices. This can lead to a redistribution of income, as those who are able to negotiate compensatory price rises – often those on higher incomes – are more likely to be protected against the worst effects of inflation. Lower rates of inflation can be less troubling for individuals, especially if they are compensated by higher incomes. For much of the past ten years the UK’s inflation rate has been around 2%, while average incomes have risen slightly more quickly. However, even in such periods, individuals whose incomes do not rise along with prices, such as some pensioners, do suffer a steady decline in living standards.
- **Saving and investment:** Inflation can reduce the level of saving in an economy, especially if real interest rates (i.e. actual interest rates minus the rate of inflation) become negative. This means that the value of savings erodes over time and may reduce the funds available to banks to lend out for purposes such as investment.

Deflation

A fall in prices could be expected to appeal to many individuals, especially those who receive fixed incomes. The outcome is an increase in the purchasing power of money, offering individuals the prospect of a higher living standard.

Figure 42.1 Actual and forecast debt to income ratio for UK households, 2005–20



However, there are some distinct disadvantages associated with deflation. Many individuals have borrowed money to finance major purchases such as cars and property. Figure 42.1 shows that since 2014 borrowing by UK households has risen as a proportion of income and is expected to reach 175% by 2020. This means that a typical household earning £50,000 a year will have debt (including mortgages) amounting to £87,500.

Deflation increases the real value of debts and, if incomes are falling, makes them

more difficult to repay. The UK has recently experienced very mild deflation; for example, prices fell by 0.1% in September 2015. If the rate of deflation becomes lower (i.e. prices fall more quickly), the impact on individuals in the UK could be severe given the high and increasing level of debts. This could reduce income available for other purposes and reduce living standards.

The effects of deflation on individuals become more pronounced if their incomes are reduced in line with prices. Some firms may do this to protect their profit margins.

The consequences for the economy

Inflation

The consequences of inflation depend upon the rate that is experienced by an economy. The UK government targets a rate of 2%, believing that this will encourage spending and investment by providing a steady increase in nominal incomes and profits. As annual rates of inflation rise above this figure, the effects outlined below will begin to appear.

- **The balance of payments:** Inflation can pose a significant threat to the competitiveness of an economy, especially if the rate is higher than the rates experienced in other countries. This can make the price of an economy's exports higher and, if demand is strongly price elastic, the volume sold may decline sharply. At the same time domestic products may struggle to compete in terms of price with imports, and sales of the latter may rise. This combination is likely to weaken the balance of payments on current account, *ceteris paribus*.
- **Employment:** The rate and level of employment in an economy may decline in these circumstances. If employees are able to negotiate compensatory wage rises, then unit labour costs will increase. Firms may opt to replace labour with capital, reducing employment opportunities. Falling sales resulting from a lack of price competitiveness will increase the negative impact of inflation on employment.
- **Investment:** Firms may reduce the level of investment they undertake if interest rates rise along with the rate of inflation. The declining competitiveness of the economy will dent business confidence and reduce forecast returns from investment. It will also create uncertainty. If investment does fall (and exports may be falling too), there could be a negative multiplier effect impacting on the level of aggregate demand.
- **Economic growth:** Rates of real economic growth may decline during periods of inflation. Aggregate demand may decline, reducing the level of real income at least in the short term.

Deflation

A period of deflation can have serious consequences for an economy. The prime negative consequence will be a fall in the level of aggregate demand. Consumption expenditure is likely to fall as consumers stop spending while they wait for prices to fall and for products to become cheaper. The increasing value of debts for businesses and households will result in a further decline in consumption and investment expenditure. Finally, governments have substantial debts, and deflation alongside falling revenue from taxes is likely to add to these. Thus government expenditure may fall too. The impact on aggregate demand will be magnified by the multiplier effect.

Key term

A **commodity** is a basic product for which there is global demand and which is often used in the manufacturing process. Examples include oil, copper, cotton and rice.

World commodity prices and domestic inflation

Many countries import large quantities of commodities from producers around the world. The economies of these countries can import inflation or deflation if the prices of these globally traded products are changing sharply. For those countries, such as the UK, that produce relatively few commodities themselves, importing inflation or deflation in this way is highly possible.

Author tip

This section refers to the calculation of the rate of inflation using price indices and also makes reference to commodity price indices. If you are unsure how weighted indices of this type are calculated, you should re-read Chapter 33 of the Year 1 textbook.

Commodity prices have the potential to have a significant impact on the rates of inflation experienced by consumers in most countries. One group of commodities is foodstuffs – wheat, rice, tea and coffee, for example. These are consumed by people in all countries and form an important part of consumption expenditure, especially in poorer countries. This is reflected in their contribution to the calculation of the rate of inflation.

Other commodities – cotton and oil are good examples – are used in a wide range of products and, in the case of oil, also contribute to transport costs. Thus, they play an important direct and indirect part in calculating the extent of price changes.

There are two major factors that influence the extent to which commodity prices can have an impact on an economy's domestic rate of inflation.

1 The contribution made by commodity imports to the calculation of inflation in an economy

Imports of commodities can have a substantial effect on the rate of inflation calculated using indices such as the Consumer Prices Index in the UK. Commodities represent a wide range of imports including important items such as food and minerals. Many are used in the production of other products, notably oil, so have a 'knock-on' effect on the price of a range of products.

In the long term, price changes for materials and fuels have been two of the main causes of the change in prices in the UK and play a significant part in the calculation of the CPI. For example, in 2015 foods and non-alcoholic beverages accounted for 110 of the 1000 weights used to calculate the CPI and fuels accounted for 42. Materials are not accorded a separate weight because they are used in a wide range of products.

According to the Office for National Statistics (ONS), the overall price of materials and fuels bought by UK manufacturers fell 12.4% in the year to July 2015, compared to a fall of 13.1% in the year to June 2015. At the same time the rate of inflation in the UK was hovering around zero.

2 The extent of the change in the price of commodities

At times, global commodity prices can be volatile. For instance, between October 1973 and March 1974 the price of oil quadrupled on world markets as a result of decisions regarding output taken by the oil-producing countries. The impact was

seen in terms of high rates of inflation in most major economies as rising oil prices pushed up prices of a wide range of goods and services. Annual inflation in the UK peaked at 25% later in 1974.

The current situation is very different. Commodity prices are falling sharply and have been for a year or so. Figure 42.2 shows the commodity price indices of the International Monetary Fund (IMF). These indicate the changes in global prices of energy-related commodities, such as oil and coal, and non-energy commodities, such as wheat, iron ore and cotton. Prices of commodities have fallen back to a level last seen in 2005.

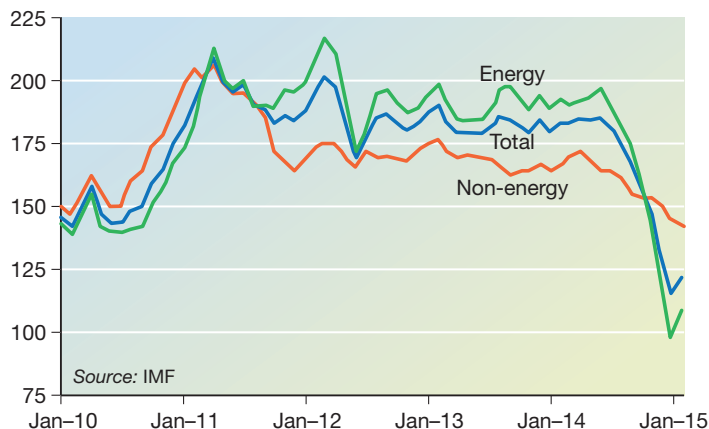


Figure 42.2 The IMF's commodity price indices, 2010–15 (2005 = 100)

REALWORLD ECONOMICS 42.2

Global commodity prices slump to an 11-year low as deflation fears grow

Global commodity prices have slumped amid a glut of supply and low demand. As a result, some economists are foreseeing the prospect of deflation in some developed economies.

A basket of commodities measured by the Bloomberg Commodity Index has fallen to an 11-year low, and the index is down 42% since its peak in 2008.

Although commodities have had a hard time since the financial crisis of 2008, falls in the price of everything from gold and oil to copper and nickel have accelerated. Tin has fallen 25% so far during 2015, Nickel is down 54% from its

February 2011 high, and the price of natural gas has slumped 40% over the same period.

There is a glut of industrial commodities on the world market and, with economic growth flagging, demand is simply not strong enough. Perhaps most telling is the price of copper, which is trading around its lowest level since the financial crisis. Known as Dr Copper due to its use as a barometer of global economic health, the industrial metal fell to \$5240 a metric tonne in July 2015, heading towards prices last seen in the summer of 2008.

Looking ahead, the ongoing weakness in oil prices is likely to

spell deflation for economies in Europe and North America. Oil plays a large part in dictating inflation levels in many economies. The last round of inflation data (showing the UK had zero consumer price inflation) was calculated before the latest fall in oil prices took place.

Source: Adapted from City AM, 21.7.15

Exercises Total: 15 marks

- 1 Explain why oil prices are an important determinant of the rate of inflation in many economies. (6 marks)
- 2 Why are economists worried about the prospect of deflation? (9 marks)

These very substantial falls in commodity prices are contributing to low or negative rates of inflation in many countries across the globe. The World Bank expects low commodity prices to continue.

Review questions

Total: 34 marks

- 1 What is the difference between inflation, disinflation and deflation? (6 marks)
- 2 What is meant by the term 'the quantity theory of money'? (2 marks)
- 3 Which of the following is the equation of exchange as devised by Irving Fisher?
 - A $MT = PV$
 - B $MV = PQ$
 - C $MV = PT$
 - D $MV = P$(1 mark)
- 4 Explain why the concept of intermediate transactions created problems for the equation of exchange, as formulated by Irving Fisher. (5 marks)
- 5 Which of the following best describes the view of most monetarists regarding the velocity of circulation of money?
 - A It is always constant
 - B It is highly volatile in the short and long term
 - C It is relatively constant in the short term, but fluctuates in the long term
 - D It fluctuates in the short term, but is relatively constant in the long term(1 mark)
- 6 Explain the difference between adaptive expectations theory and rational expectations theory. (5 marks)
- 7 Which of the following is the most likely consequence of a period of inflation for individuals?
 - A A reduction in living standards for some groups of people
 - B An increase in the real value of debt
 - C A redistribution of income from those on higher incomes to those on lower incomes
 - D An increase in the standard of living of all people within an economy(1 mark)
- 8 Which of the following is *not* likely to be a consequence of a period of deflation?
 - A A fall in the level of consumption
 - B A fall in the real value of debt
 - C A fall in government spending
 - D A fall in investment(1 mark)
- 9 Why do governments worry about the effects of periods of deflation? (6 marks)
- 10 Explain why changes in global commodity prices can affect an economy's domestic rate of inflation. (6 marks)

Possible conflicts between macroeconomic policy objectives

Key concepts from Year 1

In Chapter 42 of the Year 1 companion textbook we introduced the concepts of negative and positive output gaps and explained the effects of these on key macroeconomic variables such as inflation and unemployment. We analysed the possible causes of conflicts in policy objectives (such as reducing unemployment at the cost of increasing inflationary pressures) in both the short and long run. We also considered how economic policies might be used to try to reconcile such conflicts.

In this chapter we introduce the Phillips curve, which examines the relationship between inflation and unemployment rates. We consider how the shape of the Phillips curve might differ in the short and long terms and the implications of both short-run and long-run curves for economic policies.

The origins of the Phillips curve

Professor A.W. Phillips published a research paper in 1958 entitled *The relation between Unemployment and the Rate of Change of Money Wage Rates, 1861–1957*. In this he examined the relationship between money wage rates (as a proxy for inflation) and the rate of unemployment in the UK over a period of nearly one century. He was building on earlier research by Irving Fisher.

When Phillips plotted the data representing the relationship between money wage rates and unemployment in the UK over the period of nearly 100 years, he discovered that:

- an inverse relationship existed – a rise in one variable resulted in a reduction in the other;
- the relationship had been remarkably stable over that lengthy period of time.

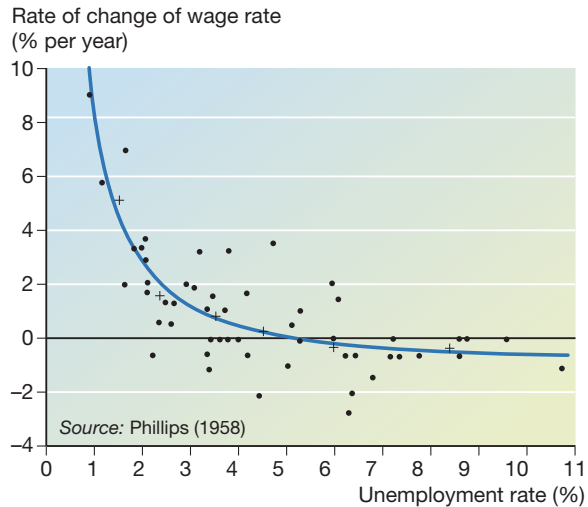
Phillips's research revealed that when unemployment was high, money wages increased slowly or even fell, but when unemployment was low, wages rose rapidly. His findings for the period 1861–1913 are shown in Figure 43.1. It is apparent that most of the data for the period in question suggest an inverse and relatively stable relationship between the rate of unemployment and the rate of change of money wages. Phillips was able to draw a line of best fit based on this data that is shown in Figure 43.1. It is this line that became the Phillips curve. He was further able to use this curve to predict the relationship between unemployment and the rate of change of money wages over the period between 1914 and 1957.

Key terms

The **Phillips curve** illustrates the relationship between the rates of inflation and unemployment that may exist within an economy.

The **rate of unemployment** is the percentage of economically active people within a country who are unemployed.

Figure 43.1 *The original Phillips curve*



One noteworthy finding from Phillips's original work is that an unemployment rate of 5.5% was shown to be sufficient to create a situation in which the rate of change of money wages was zero.

Developing the Phillips curve

The research by A.W. Phillips was based on changes in money wages rather than in prices because this was the only data available to him for the long period he analysed. However, he argued that money wages are a major component of the total costs faced by firms. If a firm's total costs rise, it is likely to seek to maintain its profit margins by increasing its prices. Therefore, Phillips contended, the rate of change of money wages is a good indicator of the rate of change in prices (i.e. the rate of inflation).

A number of economists took forward the ideas set out by Phillips in 1958. In 1960 Paul Samuelson and Robert Solow undertook further research into inflation and unemployment in the American economy. They converted the original Phillips curve into a relationship between price inflation and unemployment by deducting the estimated rate of growth of productivity from growth of money wages to provide an estimated change in prices. Their research provided remarkably similar results. It is notable that the rate of unemployment that generates zero price inflation is the same as for Phillips's original UK-based research (5.5%). Furthermore, it did not appear possible to reduce the rate of unemployment below about 1%.

Figure 43.2 shows the curve suggested by the work of Samuelson and Solow. It clarified the tradeoff that appeared to exist between inflation and unemployment: a particular rate of inflation could be achieved if decision makers were prepared to accept a certain level of unemployment.

Phillips's findings, and the subsequent developments, were important for Keynesian economists who used this work to support their views that changes in aggregate demand would effect both unemployment and prices within an economy. The aggregate supply curve could therefore not be vertical, even in the long run. The Phillips curve suggested that it was upward sloping to the right and that increases in aggregate demand could increase output, thereby reducing unemployment but at the cost of increasing inflationary pressures.

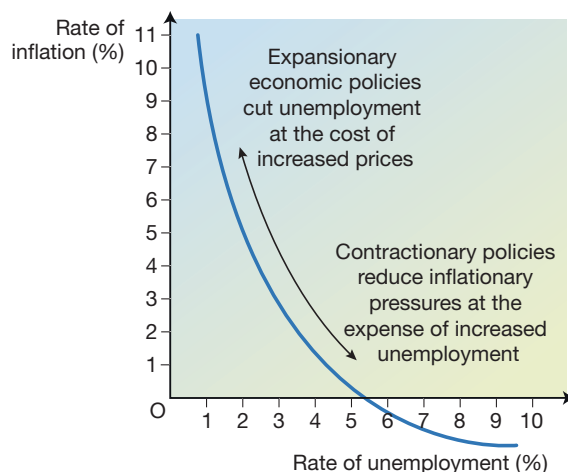


Figure 43.2 *The Phillips curve following development work by economists such as Samuelson and Solow*

The Phillips curve helped economists to identify causes of inflation: unemployment can be reduced by expanding aggregate demand, and rising aggregate demand will increase inflationary pressures in the economy as resources become relatively scarce and more costly. When unemployment is low, workers have more bargaining power to negotiate wage rises, resulting in higher prices and thus inflation.

Governments throughout the world began to base decisions relating to economic policies on the belief that there was a tradeoff between unemployment and inflation. It was possible, for example, to implement economic policies to reduce the rate of unemployment, but only at the cost of rising rates of inflation. In the 1960s and 1970s, many governments used the ideas behind the Phillips curve to inform their economic decision making.

The breakdown of the Phillips curve relationship

By the middle of the 1970s, the tradeoff between inflation and unemployment suggested by the Phillips curve appeared to have broken down. High rates of inflation and unemployment occurred simultaneously, far from the levels predicted by the Phillips curve. For example, between 1974 and 1975 the rate of unemployment in the UK rose from 2.0% to 3.2%; over the same period inflation increased from 17% to over 23%.

These developments led to much debate about the Phillips curve between economists. Some denied that it had ever existed. Others argued that the relationship remained but operated at higher rates (i.e. the Phillips curve had moved upwards, allowing higher rates of unemployment and inflation to coexist).

The short-run and long-run Phillips curve

Two economists, Edmund Phelps and Milton Friedman, challenged the Keynesian view of a downward-sloping Phillips curve. Working independently, they argued that employers and workers would behave rationally and would have access to relevant information. Thus they would only consider real wages (i.e. nominal wages adjusted for inflation). Friedman and Phelps believed that real wages would adjust to clear the labour market so that the demand for labour equalled the supply of labour. The resulting level of unemployment was termed the 'natural rate of unemployment'.

They argued that, in the short run, a downward-sloping Phillips curve existed describing the tradeoff between unemployment and inflation but that, in the long term, no such relationship existed. They concluded that the long-run Phillips curve was vertical.

The short-run Phillips curve

The short-run Phillips curve shows that a tradeoff does exist between unemployment and inflation. For example, policies implemented by a government to reduce unemployment result in rising prices. This reflects the principle that lies behind the short-run aggregate supply curve (SRAS) that we discussed in our Year 1 companion textbook. The SRAS shows that, in the short run, an economy can only increase output (or real GDP) at the expense of rising prices, or reduce prices by lowering output and raising unemployment. The short-run Phillips curve shows the same relationship and the tradeoff that occurs when the level of aggregate demand in an economy changes.

Figure 43.3 The impact of a change in aggregate demand shown using a short-run aggregate supply curve and a short-run Phillips curve

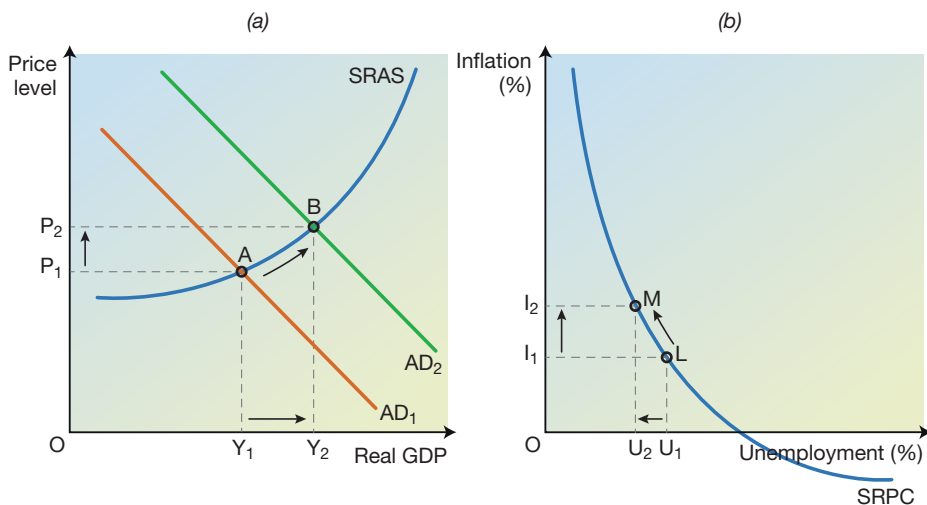


Figure 43.3(a) illustrates the effect on the price level and real GDP of an increase in aggregate demand. This could be the result of a government’s economic policy or a demand-side shock. The increase in aggregate demand leads to a movement up the SRAS from A to B. The impact is a higher price level at P_2 , reflecting inflation and a rise in real GDP to Y_2 . This is likely to result in a lower level of unemployment.

Figure 43.3(b) illustrates the same tradeoff. A rise in aggregate demand could provoke a movement up the short-run Phillips curve (SRPC) from Point L to M. The consequence is that the rate of inflation rises from I_1 to I_2 and that unemployment falls to U_2 – falling unemployment is a normal consequence of a rise in real output.

Key terms

Money illusion takes place when individuals and firms do not distinguish between nominal and real values of money when taking decisions.

The **natural rate of unemployment** is the level of unemployment that exists when the labour market is in equilibrium.

NAIRU stands for the non-accelerating inflation rate of unemployment and is the rate of unemployment that is associated with stable rates of inflation.

The long-run Phillips curve

Free market economists (including many neoclassical economists) believe that in the long run the Phillips curve is vertical. They argue that expectations play a key role in this and that the simple version of the Phillips curve focused on the current rate of inflation, ignoring expectations of future rates of inflation.

The forces that operate to make the long-run Phillips curve (LRPC) vertical are illustrated in Figure 43.4. The short-run Phillips curve intersects the LRPC at the expected rate of inflation – the SRPC is sometimes referred to as the expectations-augmented Phillips curve. So, if an economy is operating on $SRPC_2$ in Figure 43.4, the expected rate of inflation will be 2%. A change in expectations regarding inflation leads to a movement in the SRPC.

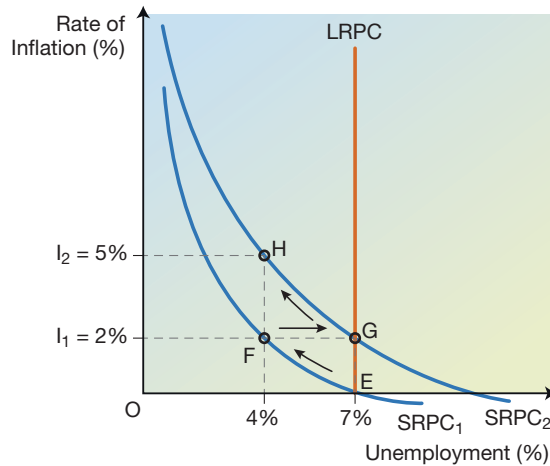


Figure 43.4 The operation of the long-run Phillips curve

Initially the economy is at Point E in Figure 43.4. The level of unemployment is at its natural rate, and in this example we assume that this is 7%; inflation is expected to be zero. The government decides to implement expansionary fiscal or monetary policies to increase aggregate demand with the intention of reducing unemployment below 7%. The expansionary policy results in the economy moving along the relevant short-run Phillips curve ($SRPC_1$) to Point F, where unemployment is below the natural rate at, say, 4%. Labour becomes increasingly scarce and trade unions and others representing workers have more bargaining power. Nominal wages and prices begin to increase. The rate of inflation rises to, say, 2% as the economy moves to Point F. At this point, money illusion may occur: more workers enter the labour market in the belief that real wages have risen; firms may believe prices are rising faster than costs. Remember, at the outset inflation was expected to be zero.

However, Point F is not a long-term position. The money illusion is temporary and in the longer term the level of unemployment rises once more to its natural rate, which we have assumed to be 7%. It is at this point that expectations play a key role in the creation of inflation. In modern economies much more economic information is available to workers and other economic decision makers. Workers, for example, now expect the rate of inflation to be, say, 2%, as in our example in Figure 43.4. They build this into their negotiations and this rate of inflation becomes embedded in the economy as rises in wages are passed on in higher prices. The economy has moved to Point G, with the natural rate of unemployment restored but accompanied by a higher rate of inflation.

Author tip

The long-run Phillips curve, as drawn in Figure 43.4, is sometimes referred to as the L-shaped Phillips curve.

This analysis has some important implications. It suggests that in the long run the Phillips curve is vertical. Further attempts to reduce unemployment below the natural rate through increasing aggregate demand will simply result in ever-increasing rates of inflation and a movement up the LRPC. Thus, there is no tradeoff between inflation and unemployment in the long run.

NAIRU

NAIRU is an acronym that stands for the **non-accelerating inflation rate of unemployment**. It refers to a rate of unemployment that is associated with stable rates of inflation. It is thus slightly different from the natural rate of unemployment, which occurs when the labour market is in equilibrium.

Most economists now accept the view, originally expressed by Phelps and Friedman, that there is a rate of unemployment that is compatible with a stable rate of inflation. This has developed into the NAIRU. The NAIRU is preferred to the natural rate of unemployment because it does not sound as if it maximises economic welfare.

The NAIRU exists when inflationary pressures are not increasing. It is therefore the rate of unemployment at which inflation will be stable and at which prices will rise at the same rate over time. It is sometimes thought to be the rate of unemployment at which inflation is zero. However, this is not the case – it can apply at times when inflation is not zero but is stable.

Estimates by Oxford Economics of the NAIRU for the UK place it at 5.75%, having risen from around 5% following the financial crisis of 2008 and the subsequent recession. In part this rise has been due to the shedding of jobs in the public sector and a mismatch between the skills of redundant public-sector employees and the needs of firms in the private sector.

The implications of the short-run Phillips curve

The short-run Phillips curve shows that it is possible for governments to implement policies in the short run that can reduce the level of unemployment or the rate of inflation in an economy. Since the shape of the short-run curve shows that a tradeoff exists, there will be a cost resulting from implementing expansionary or contractionary fiscal or monetary policies. Furthermore, the effects on employment levels will only apply in the short term. In the longer term the natural rate of unemployment will be restored.

This analysis suggests that fiscal and monetary policy can be used in the short term to achieve particular goals such as a reduction in the level of unemployment. In the longer term, such a situation might be maintained by use of other policies, notably supply-side policies. We will consider these in more detail in the following section.

The implications of the long-run Phillips curve

A key implication of the long-run Phillips curve is that a government cannot implement expansionary economic policies, intended to reduce unemployment below NAIRU, without incurring rising prices. An expansionary policy intended to decrease unemployment that is implemented over a sustained period of time is likely to result in rising rates of inflation as expectations of inflation amongst workers and firms steadily increase. For example, in Figure 43.4 expansionary economic policies that

reduce the unemployment rate to 4% initially increase the rate of inflation to 2%; expectations of inflation rise to 2% and the economy moves onto $SRPC_2$. If the fiscal or monetary stimulus is maintained over time, the economy will move to Point H where inflation has risen to a rate of 5% (a worrying figure above the government's 2% target). This could become embedded in the economy if expectations rise to this level and the economy could move to a new and higher SRPC. In this case, unemployment would be rising while inflation is at a relatively high level. This is an unpleasant combination for any government to handle.

The importance of supply-side policies

The analysis of the effects of fiscal and monetary policies on the LRPC suggests that the use of supply-side policies are likely to prove more effective in achieving the government's macroeconomic objectives in the long run. Supply-side policies include improvements to education and training to enhance skills, promoting innovation and entrepreneurship, or policies allowing inward migration. The use of supply-side policies has the potential to bring long-term benefits with less likelihood of conflict between macroeconomic objectives.

Key term

Supply-side policies are intended to increase aggregate supply by improving the effectiveness of markets.

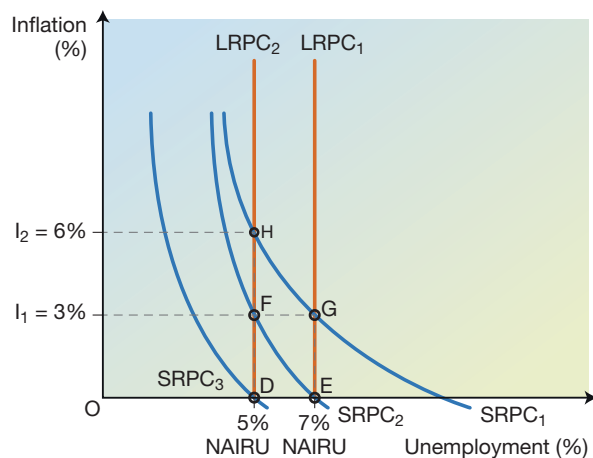


Figure 43.5 A shift in the LRPC

In Figure 43.5 the effects of a shift in the LRPC as a result of the successful implementation of supply-side policies can be seen. The movement from $LRPC_1$ to $LRPC_2$ means that the economy has expanded and that NAIURU has fallen significantly from 7% to 5%. This has implications for the rate of inflation as well. For example, the shift to $LRPC_2$ means that the economy can achieve a rate of unemployment of 5%, at which inflation is not rising; this compares favourably to the previous 7%. If the rate of inflation is stable at 3%, then the economy's short-run Phillips curve will move from $SRPC_1$ to $SRPC_2$ as the economy moves from Point G to Point F in Figure 43.5.

The value of the Phillips curve for shaping economic policy

The Phillips curve is an imperfect model, even if we use the vertical long-run version. Many economists do not accept the notion of a 'natural rate of unemployment'. The notion of real wages adjusting to clear the labour market is perhaps unrealistic given the presence of trade unions and monopsonistic employers in labour markets. Real wages have proved to be sticky, preventing labour markets clearing smoothly.

The NAIRU is a concept that many economists accept, but the complexity and dynamism of modern economies mean that it is likely to change significantly and regularly. Some economists have contended that there is only a limited tendency for an economy to return to the NAIRU. Furthermore, the NAIRU can change significantly as an economy changes – for example, as it moves through the stages of an economic cycle. In such circumstances, factors such as hysteresis can play a role in changing the NAIRU.

So, what is the value of the Phillips curve to those responsible for deciding economic policy? Economists, perhaps predictably, have differing views on its ability to explain the relationship between inflation and unemployment and thus its value to economic decision makers.

The relationship between unemployment and inflation in the UK is complex, as shown by the data in Figure 43.6. It is a complex one and there is insufficient evidence of a tradeoff between inflation and unemployment. There are periods when both are falling – as occurred between 2012 and 2015 – while both rose in 2008. On the other hand, in 2009 the two variables did show some evidence of a tradeoff taking place.

In the early part of the period covered by the graph in Figure 43.6, it is apparent that the rate of unemployment is falling while inflation is relatively stable. This would suggest that the NAIRU was falling over that period. This may have been due to UK labour markets becoming more flexible as trade union power was reduced and workers were able to move more easily between jobs. The UK introduced inflation

REALWORLD ECONOMICS 43.1

America's senior bankers hold different views on the relevance of the Phillips curve

The Federal Reserve System is a network of twelve regional banks that comprise the USA's central bank – the equivalent of the Bank of England. One governor of the bank, Daniel Tarullo, in an interview with CNBC, stated that the so-called Phillips curve has not been working for the last 10 years.

The Phillips curve is the relationship between unemployment and inflation. In the crudest definition, the lower that unemployment goes, the higher inflation goes. However, this has not been the case for the US economy. While unemployment has dropped from a peak of 10% in October 2009 to just 5.1% in September

2015, inflation has not accelerated, with prices up just 0.3% in the 12 months ending August 2015.

Tarullo's remarks match those of another member of the Fed's Board of Governors, Governor Lael Brainard. 'To be clear, I do not view the improvement in the labour market as a sufficient statistic for judging the outlook for inflation,' she said on Monday. 'A variety of estimates would suggest that the classic Phillips curve influence of resource utilisation on inflation is, at best, very weak at the moment.'

Federal Reserve Chairwoman Janet Yellen sounds a different tone. She acknowledged in a recent speech 'that significant

uncertainty attaches to Phillips curve predictions, and the validity of forecasts from this model must be continuously evaluated in response to incoming data'.

Source: Adapted from MarketWatch, 14.10.15

Exercises **Total: 14 marks**

- 1 How might using predictions based on the Phillips curve help America's central bank to take decisions regarding interest rates? **(8 marks)**
- 2 Does a period over which unemployment falls but inflation remains stable mean that the Phillips curve is 'not working'? **(6 marks)**

targeting in 1992 (one of the first countries to do so) and this has helped to make expectations of inflation lower and more realistic. The UK is an ‘open’ economy, reliant on trade, and low global rates of inflation have helped to hold down the price of imported materials, components and goods and services.

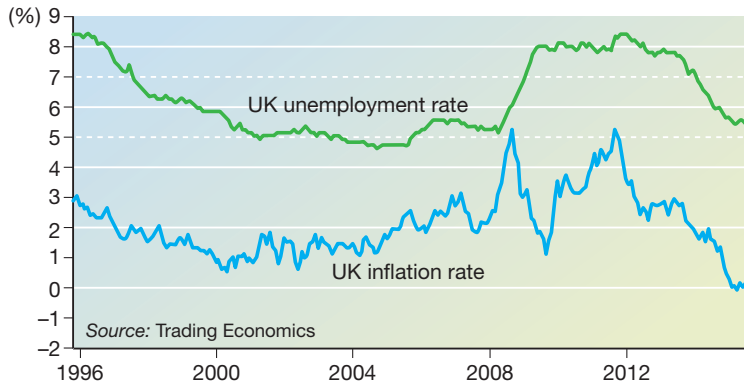


Figure 43.6 UK unemployment and consumer price inflation rates, 1996–2015

It could possibly be argued that the UK had a NAIRU of around 5% from 2000 to the onset of the financial crisis in 2008. Unemployment was relatively stable at a time when inflation remained between 1% and 2%. However, this relationship becomes less clear after that period. The evidence is not sufficiently consistent or clear-cut to provide much assistance to those taking decisions on economic policy.

The fact that the relationship seems to have varied so much over this 20-year period illustrates the problem facing those responsible for economic policies: the relationship between unemployment and inflation is far from straightforward and is subject to the influence of many variables. As a consequence, it is difficult to predict the effects of policy decisions.

Since 1958 the Phillips curve has been transformed into its expectations-augmented version and has been interpreted as vertical in the long term. The concept of the NAIRU has been added. Despite these developments, some economists believe it has no relevance; others think that it remains the key to relating unemployment to inflation in mainstream macroeconomic analysis. The debate is likely to continue.

Review questions

Total: 26 marks

- 1 Which of the following relationships was shown by AW Phillips's curve in 1958?
 - A An inverse relationship between money wages and unemployment
 - B An inverse relationship between the rate of change of money wages and the rate of unemployment
 - C A direct relationship between money wages and unemployment
 - D An inverse relationship between the rate of change of real wages and the rate of unemployment

(1 mark)

- 2 Describe how the economists Paul Samuelson and Robert Solow developed Phillips's original work.

(4 marks)

- 3** Why did Keynesian economists believe that the Phillips curve supported their view of the importance of using changes in aggregate demand to implement economic policy?
- A It suggested that the long-run aggregate supply curve was not vertical
 - B It showed a positive correlation between the rates of unemployment and inflation
 - C It showed that the rate of inflation was zero when unemployment reached 5.5%
 - D It showed that it was very difficult to reduce unemployment below about 1%
- (1 mark)**
- 4** In the 1970s the UK experienced high rates of unemployment at a time when the rate of inflation was rising. Explain why this led some economists to cast doubt on the validity of the Phillips curve.
- (4 marks)**
- 5** What name was given by Milton Friedman and other economists to the level of unemployment that exists when the labour market is in equilibrium?
- A NAIRU
 - B Full employment
 - C Involuntary unemployment
 - D Natural rate of unemployment
- (1 mark)**
- 6** Explain the difference between the short-run and long-run Phillips curves as developed by Edmund Phelps and Milton Friedman.
- (8 marks)**
- 7** A rise in the expected rate of inflation will have which of the following effects on the relevant short-run Phillips curve?
- A A movement to the left along the existing SRPC
 - B A movement to the right along the existing SRPC
 - C An upward shift to the right of the SRPC
 - D A downward shift to the left of the SRPC
- (1 mark)**
- 8** Which of the following would be most *unlikely* to cause a change in an economy's level of NAIRU?
- A Improvements in education and training
 - B Hysteresis
 - C Increased flexibility in labour markets
 - D Increases in the level of indirect taxes such as VAT
- (1 mark)**
- 9** Explain one reason why some economists do not believe in the existence of a 'natural rate of unemployment'.
- (5 marks)**

Topic 11 Exam-style questions

A-LEVEL PAPER 2

SECTION A Context: Economic growth, unemployment and deflation

Extract A **Slowdown in the rate of economic growth**

The latest figures for economic growth show that the UK economy is growing more slowly than forecast and more slowly than in the second quarter of 2015. GDP growth has slowed over recent months and entered the final quarter of the year at a rate of just 0.3% – far below the 0.7% enjoyed in the second quarter. However, the rate of economic growth in the UK is higher than that in many other developed economies, offering benefits to individuals and businesses.

Output in the services sector continues to expand quickly, but construction and manufacturing output produced very disappointing data. British manufacturing, which accounts for 10% of GDP, is experiencing a recession once more. Part of the slowing rate of growth was due to the biggest fall in construction output in three years: a drop of 2.2%. Output in the manufacturing sector declined by 0.3%. In contrast, the service sector, the biggest element of the economy, grew by 0.7%.

Economists have argued that the slowing of the rate of economic growth in the UK is likely to rule out the prospect of a rise in interest rates, from their current record low of 0.5%, in the near future. This would be a relief to many households that have high borrowing and has implications for the level of consumption in the UK economy. The pound weakened against the euro and the dollar after the growth figures were announced.

Sources: Adapted from various media, December 2015

Extract B **UK unemployment data**

The rate of unemployment in the UK was 5.4% in the three months to August 2015 – its lowest since 2008 – according to figures released by the Office for National Statistics. The unemployment rate reached 8.5% in 2011. The number of people out of work was 1.77 million between June and August 2015, down 79,000 from the previous quarter.

At the same time the number of people employed in the UK economy increased by 140,000, giving an employment rate of nearly 74% – a record figure. Approximately 75% of the people who have found a job in the UK over the year to August 2015 are from elsewhere in the European Union. The rise in employment in the UK is strongly due to migration: employment amongst EU migrants leapt by almost 20% compared to a 0.4% rise in the number of UK citizens now in work.

Over the three months to August 2015, workers' earnings were 3% higher than in the same period of 2014 and were rising less quickly than expected. The growth in wages in the UK remains slower than before the financial crisis in 2008 but is starting to rise more quickly.

Extract C **Flirting with deflation**

The UK is experiencing deflation once again as consumer prices fell in September 2015 for the first time since April 2015.

Inflation was -0.1% in September, matching the lowest level seen in the UK economy since 1960. Economists had expected prices to remain unchanged at zero. Aggressive clothing discounts and falling petrol prices were to blame for September's numbers, said the Office for National Statistics.

Clothing and footwear prices rose by 2.8% between August and September 2015 compared with a rise of 4.0% during the same period in 2014. Food prices, which have been falling as result of the supermarket price wars, rose last month by 0.02% . However, in more encouraging signs for the economy, core inflation, which strips out volatile elements such as energy and food, remained unchanged from August at 1.1% . Economists now expect inflation to start bouncing back at the end of the year.

Figure A UK CPI inflation, 2010–15



Source: Adapted from the *Daily Telegraph*, 13.10.15

The UK has been flirting with deflation since the start of 2015, with CPI hitting 0% four times over the last nine months. Average inflation in 2015 is set to register at 0%, one of the lowest years of consumer price rises in the postwar era, according to analysts at Barclays Bank.

Questions

Total: 40 marks

- Use the data in Extract B to calculate the number of economically active people in the UK in the three months to August 2015. (2 marks)
- Use the information in Extract B to explain the factors that led the UK to experience a record rate of employment in 2015. (4 marks)
- Analyse the likely consequences of a period of deflation in 2015 for the performance of the UK economy. (9 marks)
- Using the data in the extracts and your economic knowledge, evaluate the view that the benefits of economic growth always outweigh the costs. (25 marks)

SECTION B Essays

Total: 40 marks

The Indian economy is achieving consistently high rates of long-run economic growth. In 2015 the rate of economic growth reached 7.5% and the rate of unemployment fell to 4.9% , its lowest on record. During the same year the rate of inflation rose from 3.5% to over 5% at a time when global inflation rates were very low.

- Explain, using relevant diagrams, the difference between short-run and long-run economic growth. (15 marks)
- Discuss the view that it is inevitable that the rate of inflation will increase as an economy nears full employment. (25 marks)



Topic 12

Financial markets & monetary policy

The structure of financial markets & financial assets

Key concepts from Year 1

There are no relevant concepts from Year 1 as this is an entirely new topic.

This chapter describes the characteristics and functions of money and two ways in which the supply of money can be categorised. We examine the differences between the money markets, the capital markets and the foreign exchange market as well as the roles played by financial markets in the wider economy. We also consider the distinction between debt and equity, and the inverse relationship between the price of bonds and market interest rates.

Money and the money supply

Money is anything that is generally acceptable in payment for goods and services or for settling debts. A key feature of any form of money is that the people and organisations using it have confidence in its value as a means of making payments. Since the 1970s electronic credit systems have been developed, allowing the use of credit and debit cards for payment and creating flexible forms of money that are not cash. Since 2009, cryptocurrencies have been available. These are digital currencies created electronically on computers through the use of software that solves mathematical problems. Bitcoin is perhaps the best-known example of this most modern type of money.

The functions of money

It is generally recognised that money has four functions, though other functions do exist (Figure 44.1).

1 A medium of exchange

Arguably, the most important function of money is to act as a medium of exchange to

Key terms

Money is a medium of exchange – something that is acceptable in payment for goods and services or to settle a debt.

The **double coincidence of wants** exists when two parties to a transaction have products of approximately equal value and both want what the other possesses.

Hyperinflation is a rate of increase at which the authorities have lost control of prices. It is sometimes thought to be a figure in excess of 1000% per annum.

Barter is the direct exchange of goods or services between parties to a transaction without the use of money.

Liquidity measures the ease with which assets can be turned into cash.

enable transactions to take place. Without money, all transactions would have to be conducted through the process of barter, which entails swapping one good or service for another. The difficulty with a barter system is that the ‘double coincidence of wants’ has to exist. This means that two parties to a transaction have to hold products of roughly equal value and, critically, to want each other’s product. The likelihood of both parties to a transaction wanting each other’s products is comparatively small and the absence of money to act as a medium of exchange is likely to result in far fewer transactions taking place in an economy. Economies are unlikely to grow in such circumstances.

Money eliminates the need for a double coincidence of wants for every transaction. It acts as a medium of exchange that is acceptable to all parties to a transaction because it can be exchanged for any goods or services that are desired. Thus, its existence and use enables many more transactions to take place and for the level of economic activity in an economy to increase.

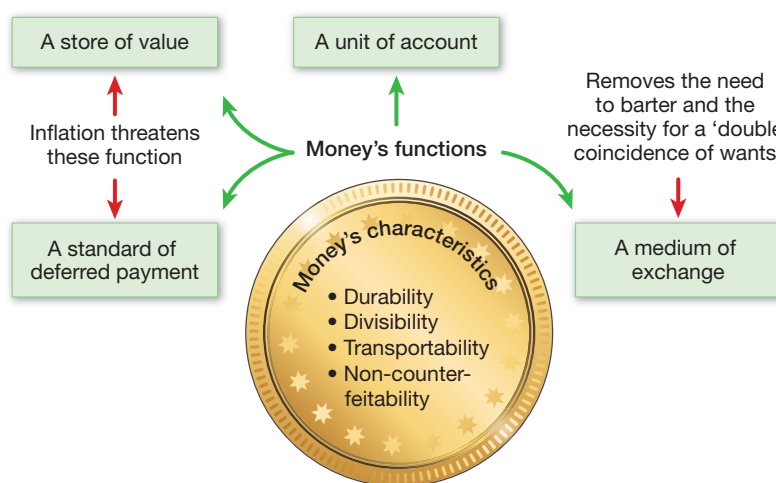


Figure 44.1 *The characteristics and functions of money*

2 Unit of account

Money also serves as a unit of account. This means that it provides a common and agreed measure of the value of goods and services, allowing parties to transactions to assess the value of the diverse goods and services that are available to them. It also allows those engaged in transactions to compare the values of goods and services and to judge the opportunity cost of a specific transaction. Knowledge of the monetary value of a good allows suppliers and buyers to make decisions about how much of the good to supply or purchase.

3 Store of value

If money is to be effective in acting as a medium of exchange, it must maintain its value over time. In order to do this, it must act as a store of value. The users of money require that it be stored (for example, as savings) over some period of time and still have value when exchanged for goods and services. If money were to lose value quickly over time, it would not be adopted as a medium of exchange. At times, money in modern economies has lost value very quickly due to hyperinflation, with the result that it has not been viable as a medium of exchange.

As a store of value, money is not unique. Many items can be used as stores of value, such as houses, antiques and precious metals. Other means of storing value have developed because money is sometimes subject to the effects of inflation. However, money does have an advantage in that it is highly liquid and easily exchanged for goods and services. Other stores of value may be less easily converted into a form that can be exchanged in transactions. This is termed ‘illiquidity’.

4 Standard of deferred payment

In modern economies, many contracts are agreed that involve payment at a later date. It is common for governments, firms and households to engage in transactions that entail payment at some specified future date or dates. Examples include the purchase of products (such as houses and cars) that require loans and a stream of future payments, or a firm signing a contract for the delivery of materials for a period of time with payments made regularly throughout the period. The parties receiving payment in the future (the creditors) would only agree to such a transaction if money were an acceptable form of payment in the future.

Inflation can pose a threat to money being used as a standard of deferred payment because, if inflation is persistent, it can reduce the value of money over the longer term. If inflation is anticipated, creditors can protect themselves against its effects by charging higher interest rates for any loans.

The characteristics of money

If money is to be effective within an economy and to fulfil the functions discussed above, it has to have a number of characteristics. Money has four primary characteristics.

- **Durability:** This means that money should not degrade over a relatively short period of time. This explains why metal coins have been used as money for thousands of years and why the Bank of England is to issue its first plastic banknotes in 2016. These notes will last up to three times as long as paper banknotes and will be more difficult to counterfeit. Durability is essential if money is to act as a medium of exchange and a store of value. A form of money will work as a medium of exchange if it stores value from one transaction to the next. Thus it must be durable.
- **Divisibility:** In order to be used as a medium of exchange in relatively trivial transactions, it must be possible to divide money into fairly small units of account. If it is not divisible, then transactions of limited value will have to use a different medium of exchange. The most effective forms of money are capable of being divided into very small units of value to facilitate all types of transactions.
- **Transportability:** If money is to allow transactions to take place in any location, it must be small and light enough to be easily transported. Some modern forms of money do not have a physical existence (money available on a credit card, for instance), meaning they are extremely portable.
- **Non-counterfeitability:** If money is to perform its functions successfully, it must not be easy to copy. If money can be copied, it will become less scarce and will lose value, making it less effective as a store. Furthermore, many economists would argue that it is vital for governments to control the supply of money in the economy. Thus it is important to avoid the counterfeiting of notes and coins.

**REALWORLD
ECONOMICS** 44.1

Greeks turn to barter

The barter economy is growing in Greece after a bank shutdown has left many struggling to maintain businesses with limited cash. A rising number of Greeks in rural areas are resorting to swapping goods and services as they struggle to keep businesses going under tight government capital controls, according to reports.

Christos Stamatis started the website Mermix three months ago to link farmers wanting to share heavy machinery in exchange for cash or other goods. 'A barter



People queue to withdraw money from an ATM outside a branch of Greece's Alpha Bank in Athens, 28 June 2015

economy is something that we shouldn't aspire to and should be a thing of the past – the last time we had it on a large scale was when we were under occupation,' he told *Reuters*, referring to the Nazi occupation of Greece in the Second World War.

The Greek government shut the banks on 28 June in the heat of the Greek debt crisis, fearing a run on banks. After they reopened on 20 July, strict controls on withdrawals were put in place to stop cash from trickling out of the system. Greeks are prevented from taking more than €60 a day from their accounts, and services in branches are limited to basic processes, such as making domestic payments.

In rural areas, some farmers have admitted to going to the bank every day but still not having enough money to pay employees and suppliers. One farmer said that when he rented a field recently,

he agreed to pay with part of his clover production. Another farmer, Mimis Tsakanikas, said: 'I've begun bartering in some forms – it existed in the past but now it is growing... Times have become really tough, and friends and relatives help each other out'.

It is difficult to quantify the scale of the barter economy growing in rural Greece, but online forums and anecdotal evidence suggest that it is surging.

Source: *The Independent*, 30 July 2015

Exercises (12 marks)

- Which of the functions of money was affected by the decision of the Greek government to close the country's banks for several weeks? Explain your reasoning. (5 marks)
- Why might the increasing reliance on barter prove harmful to the Greek economy? (7 marks)

It is possible to argue that money should have other characteristics, such as being in limited quantity to preserve its value and having a long history of acceptance.

The supply of money

The money supply is the total stock of money that is circulating in an economy. This is a general description and there are a number of more technical explanations of what comprises the money supply in the UK. However, the measurement of the money supply in the UK, or any other economy, is complicated by difficulties in defining what exactly constitutes money.

We can divide the supply of money into two categories: narrow money and broad money.

Narrow money

Narrow money includes all physical money like coins and currency along with deposits on demand and liquid assets held by the central bank. Deposits on demand refer to money held in accounts that is available immediately (i.e. 'on demand')

to account holders. Narrow money is that part of the wider money supply that is available for transactions.

Table 44.1 *Narrow and broad measures of the money supply*

Category	Example of a measure	What it includes
Narrow money	M0	<ul style="list-style-type: none"> ■ Notes and coin in circulation with the public ■ Money in the tills of UK banks and building societies ■ Deposits with the Bank of England from high street banks
Broad money	M4	<ul style="list-style-type: none"> ■ Private sector (i.e. the UK private sector other than monetary financial institutions holdings of: <ul style="list-style-type: none"> – sterling notes and coin – sterling bank deposits, including certificates of deposit – a range of bonds due to mature within five years plus estimated holdings of some financial securities

The composition of M0, which is one measure of narrow money, is shown in Table 44.1. M0 is believed to have a correlation with high street spending, and thus can provide an indicator of inflation. Changes in the size of M0 can have a significant impact on consumption and hence on the performance of the economy. Despite this, the data needed to calculate M0 have not been collected by the Bank of England since 2006.

There are other ‘narrow’ measures of the money supply, such as M1. This includes slightly more forms of money than M0.

Author tip

The UK has used many different methods of measuring the money supply since the 1970s and the definitions of these measures is highly technical. You only need to have a general understanding of the distinction between narrow and broad measures of the money supply.

Broad money

Broad money is a much more inclusive measure of the money supply including notes and coins as well as financial assets which are less liquid, such as savings accounts and institutions’ bank accounts. The Bank of England has collected data relating to M4 as a measure of the broad supply of money since 1982.

Many economists regard M4 as an important indicator of the level of economic activity in the economy. Figure 44.2 shows the rate of growth of the broad money supply in the UK (as measured by M4). It is apparent that the rate of growth of money supply dropped dramatically, and became negative, with the onset of the financial crisis and subsequent recession. The negative growth in M4 has a clear

Key terms

The money supply is the total stock of money that circulates within an economy.

Narrow money includes all physical money like coins and currency along with deposits on demand and liquid assets held by the central bank.

Broad money is a much more inclusive measure of the money supply including notes and coins but also less liquid financial assets, such as savings accounts and institutions’ bank accounts.

correlation with the UK's severe recession in 2008–09. However, economists may argue about the direction of causality (i.e. which caused which).

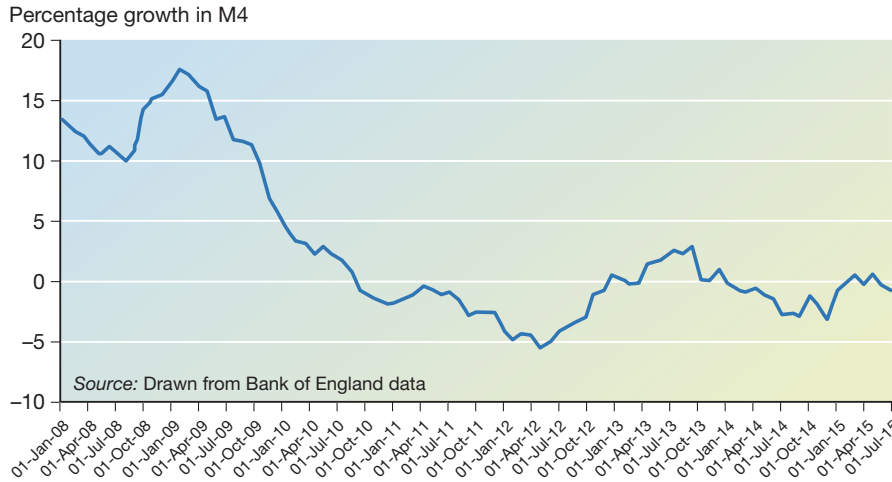


Figure 44.2 Annual growth rate for M4 in the UK, 2008–15 (monthly data, seasonally adjusted)

Its rate of growth provides some indication of the underlying rate of inflation experienced by the UK over this period. However, it does not reflect the rise in inflation to over 5% in the latter part of 2011, although to some extent this was due to cost-push factors because some import prices rose along with an increase in the rate of VAT from 17.5% to 20% in January 2011.

The UK's financial markets

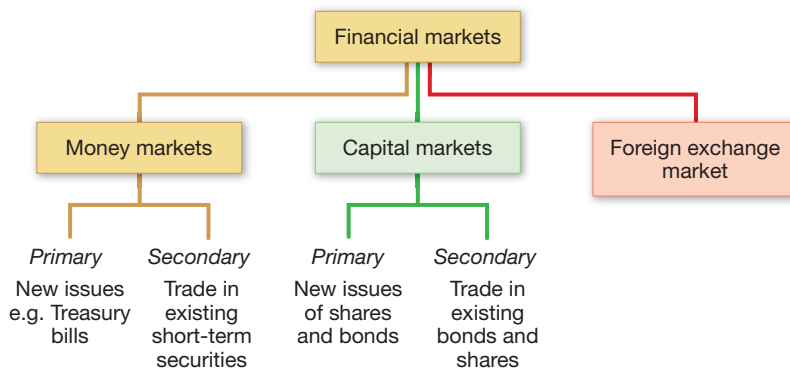


Figure 44.3 Some of the UK's financial markets

The UK and other economies each possess a number of financial markets (Figure 44.3). These are important as sources of finance and foreign currency for banks and other

The **money market** comprises financial institutions and other organisations that wish either to borrow or to lend on a short-term basis.

Capital markets are those markets in which financial securities such as shares and bonds are issued to raise medium- to long-term finance.

Financial securities are financial assets such as shares that can be bought and sold on markets.

financial institutions, insurance companies, individuals and companies, as well as local and central government. The UK is a key global financial centre and has larger and more active financial markets than most other countries.

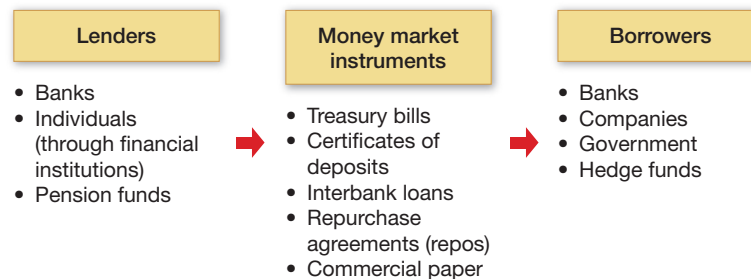
Financial markets in the UK exist to facilitate saving, provide short- and long-term finance to businesses, individuals and governments, facilitate the exchange of goods and services by making payments, and to provide forward markets (where deals can be arranged in advance) for commodities and currencies.

We shall look at three of the financial markets that operate in the UK: the money market; the capital market; and the foreign exchange market.

1 The money market

The fundamental purpose of the UK's money market is to provide a means for lenders and borrowers to satisfy their short-term financial needs. Funds can be borrowed for a period of a few days to just under a year. Although we talk about the London money market, it is really a series of linked markets trading in different short-term financial products – known as ‘money market instruments’ (Figure 44.4) or more generally as securities. Money markets provide liquidity to the UK and global financial system by meeting the needs of governments and commercial organisations for cash.

Figure 44.4 *The components of the UK's money market*



The money markets in London and in other major economies are typically seen as safe places to invest money due to the highly liquid nature of the securities and their short maturities. Since they represent a relatively safe investment and can easily be turned into cash, money market instruments or securities offer significantly lower returns than are available in other financial markets.

Participants in the money market

The money market brings together those who have surplus funds that they are willing to lend short term and those who need finance. Those lending are likely to be organisations of one form or another because of the sums involved. Individuals normally lend onto the money markets through a bank or other financial institution.

Key terms

Treasury bills are short-term government securities, maturing after 91 days and sold at a discount on their face value.

Hedge funds are privately owned investment companies that provide services to wealthy individuals or to professional investors such as pension funds.

Commercial paper is a short-term debt instrument issued by a company.

The **Treasury** is the government's economic and finance ministry, maintaining control over public spending and overseeing the UK's economic policy.

An individual wishing to purchase Treasury bills would have to invest a minimum of £500,000.

The other half of the market comprises large companies, specialist financial businesses such as hedge funds and the government. They are seeking to borrow short term in order to meet their needs for cash to allow them to carry out their operations without interruption.

Money market instruments

The instruments or securities that are traded on the money market are sold by organisations that are normally reliable and secure. All mature (i.e. are repaid) within one year and can be converted easily to cash by being sold on the secondary money markets. In other words, they are very liquid assets.

The main instruments traded on the money market are listed below.

- **Treasury bills:** These are issued by the Treasury each week on behalf of the UK government to ensure that the UK authorities have sufficient holdings of cash to meet expenditure requirements. Treasury bills are normally issued for three months (91 days). Interest is not paid on Treasury bills; instead, they are sold below face value and the buyers receive the full value of the bill when they mature. Treasury bills are sometime said to be ‘zero-coupon’ as there is no interest paid. The difference between the selling price and the face value of the Treasury bill represents the investor’s return. Treasury bills are judged to be very low risk because they are a loan to the UK government, which is highly unlikely to default on repayment.

Number crunching

Assume a Treasury bill is issued with a face value of £500,000. It is sold for £490,000. What rate of interest (or yield) does the buyer of this Treasury bill receive when they are paid the full value at maturity?

- **Commercial paper:** It is common for companies to issue short-term bills, known as ‘commercial paper’, to ensure that they have sufficient cash to meet day-to-day expenses. They are slightly more risky than Treasury bills, and are normally issued at a discount on face value, in a similar way to Treasury bills. The buyer receives the full value of the security when it matures. Commercial paper can be issued for as long as nine months, but a period of one to two months is more common.
- **Certificates of deposit (CDs):** An organisation deposits funds with a bank and in return receives a certificate that a deposit has been made. The certificate of deposit can then be sold in the money market whenever the organisation needs cash. Most CDs are issued for periods of less than six months. Interest can be at a fixed or variable rate, although some may also be issued at a discount to face value in the same way as Treasury bills.
- **Interbank loans:** Banks in the UK grant loans to one another to ensure that they have sufficient funds to meet customers’ demands for cash at any time. Banks in the UK are required to hold an adequate amount of cash and other liquid assets to ensure that they can maintain the confidence of their customers. A bank that is short of liquid assets can borrow, sometimes just overnight, from another

bank that has surplus cash. The rate charged for these loans is called the London Interbank Offered Rate or LIBOR.

- **Repurchase agreements (repos):** These are a form of overnight borrowing backed by government securities. The seller of the government security (for example, a Treasury bill) agrees to buy it back at a specified price and time.

2 The capital market

Capital markets are those markets in which financial securities such as shares and bonds are issued to raise medium- to long-term finance. There are a number of different capital markets in the UK. The London Stock Exchange operates important capital markets in the UK, providing services to the UK and other governments as well as to businesses from across the globe. In order to suit the needs of different organisations, the Stock Exchange offers three separate but complementary capital markets: the Main Market, AIM and Professional Securities Market (PSM). The Main Market offers access to a potentially enormous supply of capital, while the AIM (previously called the Alternative Investment Market) is designed to help smaller growing businesses raise capital. The PSM is used to raise specialist forms of capital.

Firms issue bonds and shares to raise capital, though only a minority of UK companies do so in public capital markets. Bonds are also issued by other organisations such as the UK government when in need of medium- to long-term funds. In capital markets, securities such as shares and bonds are issued in what is known as the ‘primary market’ and traded subsequently in the ‘secondary market’.

Participants in the capital market

Capital markets bring together those in need of capital with those who wish to invest over the medium to long term. Investors include individuals, businesses, organisations such as pension funds, and insurance companies as well as local and national governments.

The capital markets in London are used by governments and businesses from over 70 countries. Approximately 3200 companies use the services of the London Stock Exchange for the purposes of raising capital.

Capital market instruments

The capital market is used to trade a range of financial instruments or securities. Some of these are highly complex and used for specialised purposes. We shall consider two categories of security that can be issued and traded on capital markets.

- **Bonds:** These are simply debts that can be traded. Governments, companies and other organisations issue bonds to raise capital. These are not repaid for at least

Key terms

Bonds are financial securities issued by governments and companies with the aim of raising capital for repayment over the medium to long term.

Shares are equal units of ownership of a company offering financial and other benefits.

The **coupon rate** is the interest rate attached to the bond.

Maturity is the date at which a bond (or Treasury bill) is due for repayment.

one year and often much longer. The bonds confer on the holder (i.e. the person or organisation that has lent the money) the right to repayment at an agreed time. Once issued, bonds (which include the right to the repayments) can be traded on capital markets such as those operated by the London Stock Exchange. Most bonds are repaid at their face value on a particular date, with a fixed rate of interest paid at regular intervals until that time. The coupon rate is the interest rate attached to the bond, while the yield is the coupon rate divided by the market price of the bond. There are numerous types of bond, including zero-coupon bonds that are sold at a discount on their face values to make up for the lack of interest payments. The price of bonds on markets can fall and rise and is subject to a range of influences, not least the rate of interest.

- **Shares:** Shares are sold on capital markets by public companies. These companies are generally large and have the initials 'plc' after their names. Shares are sold to raise capital for a variety of purposes including expansion. Capital raised by selling shares is sometimes referred to as equity. Those buying shares have purchased a portion of the company concerned and may receive a part of the company's profits (a payment called dividends). The payment of dividends for many types of shares is not guaranteed and depends on the level of profits made by the company concerned. The price of shares can rise or fall on capital markets and at times may be highly volatile, making their purchase potentially risky. If individuals purchase a sufficient number of a company's shares, they can gain control of the organisation.

The relationship between market interest rates and bond prices

An inverse relationship exists between market interest rates and bond prices. Thus a rise in market interest rates can be expected to reduce bond prices. This is best explained by use of a simple example.

Assume that a government has issued a bond with a face value of £100,000. It is paying a fixed interest payment of £5000 per year – this is known as the 'coupon'. The annual yield (or coupon rate) on this bond is calculated by use of the following formula:

$$\text{Annual yield} = \frac{\text{Annual return or interest}}{\text{Market price}} \times 100$$

$$\text{Therefore, the yield in our example} = \frac{£5000}{£100,000} \times 100 = 5\%$$

This bond would have been issued at a time when market interest rates were around 5% to make it a relatively attractive investment. The risk entailed in lending to the government is generally regarded as low, which might be reflected in slightly lower interest rates.

If we assume that the market rate of interest suddenly increases to 8%, new bonds issued by the government will have to provide coupon rates that offer a yield of approximately 8% if they are to sell. Bonds with lower coupon rates will be unappealing to investors and their prices on the market will fall. The £100,000 bond in our example with a 5% yield will experience a fall in price. It could be expected to fall in price to around £62,500, at which time the annual payment of £5000 will represent a yield of 8% – equivalent to the market rate of interest.

Similarly, when interest rates fall, bond prices rise and their yields fall to be consistent with current rates.

Author tip

You need to be able to calculate the yield on a government bond. Practise this, starting with the number crunching exercise provided here.

Number crunching

The UK government issues a bond valued at £1 million. The bond is due to mature in 2025 and the holder receives an annual coupon payment of £20,000. After two years, the market interest rate falls to 1%.

- (a) What is the yield on this bond at the time it was first sold?
- (b) What will happen to the price of the bond when the market rate of interest falls to 1%?

3 The foreign exchange market

The foreign exchange market is one in which one currency is traded for another. It is sometimes called the 'forex market'. The scale of transactions on foreign exchange markets is enormous: average global turnover is in excess of £3500 billion per day. The exchange rates of many currencies (the price of one currency in terms of another) are determined by the forces of demand and supply in foreign exchange markets, and on occasions governments buy and sell currencies to influence their rates. A typical transaction involves the purchase of an amount of one currency by paying with another currency. Thus all transactions involve the sale and purchase of currencies.

Participants in the foreign exchange market

Individuals, businesses and governments all participate in the foreign exchange market, though it is usual for them to use other organisations to act as intermediaries to buy and sell currencies on their behalf. The biggest users of the foreign exchange market are banks, as they operate globally and advance loans in many currencies. Most trade on the foreign exchange market is between either companies and banks or institutional investors and banks. Banks therefore act as the intermediaries in most transactions.

Currency trading takes place continuously, except at weekends. All the major banks that make up the foreign exchange market have offices across the world, in London, New York and Tokyo. Thus as one city ceases business, currency trading continues in another.

Individuals buy currencies for purposes such as tourism, although wealthy individuals

Key terms

Yield refers to the interest or dividends received from a security, usually expressed as a percentage based on the security's market price.

Foreign exchange markets are those in which traders buy and sell currencies.

The **exchange rate** is the price of one currency expressed in terms of another, for example £1 = \$1.55.

may speculate on the foreign exchange market. Businesses trade in currencies to enable them to purchase goods and services from overseas; foreign suppliers often want to be paid in their domestic currency to eliminate any risk to revenues from changing exchange rates. Other businesses (for example, hedge funds) may speculate in the foreign exchange market in the hope of purchasing currencies that rise in value. There is much speculative trading in this market: some estimates suggest that over 70% of foreign exchange transactions may be speculative.

Central banks also participate in the foreign exchange market. They accumulate reserves in the form of diverse foreign currencies. Some governments use this foreign exchange, for example, to buy their own currency in the foreign exchange market if it falls in value.

The difference between debt and equity

Businesses have three major options when seeking to raise capital for medium- to long-term investment purposes. They can sell shares, sell bonds – those issued by businesses are called corporate bonds – or borrow from a bank by arranging a medium- to long-term loan.

Businesses that participate in the capital market with the intention of raising funds have to make the fundamental decision as to whether to raise finance through borrowing (i.e. incurring debt) or through the sale of shares (i.e. through equity). Both approaches have advantages and disadvantages. Raising capital through debt (whether by selling corporate bonds or arranging bank loans) ensures that the existing owners retain control, but it commits the business to regular interest payments that may impact upon profits. Raising too much capital in the form of debt can make businesses vulnerable to rises in interest rates.

Raising capital by the sale of shares removes the need to make regular interest payments. Payments to shareholders, in the form of dividends, may only be made if sufficient profits are generated. However, the sale of large numbers of shares can result in the existing owners of a company losing control, as most shares confer voting powers on their owners.

The role of financial markets in the wider economy

Financial markets perform essential roles in the economy. They provide short- and long-term finance as well as foreign currencies for governments, organisations and individuals who require them. Financial markets also provide investment opportunities, of differing timescales, allowing interested parties access to a range of financial securities. In this way they provide a means to channel funds from those who wish to invest to those who require funds for a variety of purposes. By bringing together ‘buyers’ and ‘sellers’ as efficiently as possible, these markets play an essential role in enabling firms to raise funds. This, in turn, is vital to allow firms to increase productive capacity and efficiency and for governments to develop the economy’s infrastructure as well as to provide essential services such as education and health.

Money markets ensure that governments and companies have access to sufficient cash so that they can conduct their day-to-day operations successfully. The activities of central banks in money markets can also be used to manage the economy’s money supply. For example, the Bank of England may buy Treasury bills from commercial banks (such as HSBC or Lloyds Banking Group) so that the banks can boost their

Key terms

Debt is an amount of money borrowed by one person or organisation from another and repaid to an agreed schedule, usually with interest added.

Equity refers to the value of that part of a business’s capital that is generated through the sale of shares.

cash reserves, enabling them to increase their loans and potentially the money supply. The securities are sold back at a later date. When the central bank wants to tighten money supply, it sells the Treasury bills first, and buys them back later. This is called a reverse repo – an agreement to lend securities rather than funds. This reduces the cash reserves of the commercial banks.

Capital markets play an important role in financing the activities of non-financial companies in the UK. These markets provide a viable alternative to bank loans and retained profits as sources of medium- to long-term finance. Since the start of the financial crisis in 2008, the sale of bonds and shares has allowed some large companies to overcome the decline in bank lending. The capital markets have therefore helped to stimulate investment and economic growth in the UK economy and have aided its recovery from the recession.

Even though only a minority of UK companies use these markets to raise capital through the sale of bonds or shares, those that do account for a relatively large share of economic activity, including domestic employment and investment. Thus the successful and effective operation of capital markets in the UK has helped to channel funds to productive uses and will have boosted rates of economic growth. Furthermore, the government uses the capital market to sell bonds to finance its budget deficit and the national debt and to ensure that it can raise sufficient funds for its long-term expenditure plans.

Foreign exchange markets are essential to businesses engaging in international trade. For example, those businesses that purchase supplies from overseas rely on being able to exchange sterling for other currencies. For a relatively ‘open’ economy such as the UK, having an efficient market for foreign exchange is essential to promote production, employment and trade.

Finally, the UK’s financial markets provide employment directly and contribute significantly to the UK’s GDP. In 2014, the ONS estimated that financial and insurance services contributed £126.9 billion to the UK’s GDP – about 8% of its total.

Review questions

Total: 27 marks

- 1 Explain how a barter system is dependent upon the double coincidence of wants. (6 marks)

- 2 Which of the following is a characteristic of money?
 - A A medium of exchange
 - B A measure of value
 - C A standard of deferred payment
 - D A long history of acceptance(1 mark)

- 3 Explain the difference between the terms ‘narrow money’ and ‘broad money’. (5 marks)

- 4 In which of the markets listed below are Treasury bills sold?
 - A Money market
 - B Capital market
 - C Foreign exchange market
 - D AIM(1 mark)

- 5** Which of the following types of financial security would be sold on capital markets?
- A Treasury bills
 - B Company shares
 - C Certificates of deposit
 - D Commercial paper
- (1 mark)**
- 6** Which of the following terms describes the time at which a bond is repaid?
- A Coupon
 - B Yield
 - C Maturity
 - D Equity
- (1 mark)**
- 7** Explain the difference between primary and secondary financial markets.
- (4 marks)**
- 8** Explain why businesses may become involved in transactions in the foreign exchange market.
- (6 marks)**
- 9** Assume that a government has issued a bond with a face value of £500,000. It is paying a fixed interest payment of £20,000 per year. Which of the following is the annual yield (or coupon rate) on this bond?
- A 4.0%
 - B 0.4%
 - C 2.0%
 - D 4.4%
- (1 mark)**
- 10** Assume the Treasury issues a bond with a face value of £1 million paying a coupon of £15,000 a year. Market interest rates rise from 1.5% to 2.0%. Which of the following is most likely to represent the market price of this bond following the rise in interest rates?
- A £1,250,000
 - B £1,333,330
 - C £800,000
 - D £750,000
- (1 mark)**

Commercial banks & investment banks

Key concepts from Year 1

There are no relevant concepts from Year 1 as this is an entirely new topic.

In this chapter we explain the difference between a commercial bank and an investment bank. Thereafter we focus on commercial banks in detail, examining their main functions, the structure of their balance sheets and how commercial banks create credit. We also consider their main objectives – liquidity, profitability and security – and how these three objectives may conflict.

Commercial banks and investment banks

Since 1975 the size of the UK banking system has grown rapidly, with the value of the total assets of UK banks rising from around 100% to around 450% of the economy's GDP. A prominent feature of the UK banking industry is the high proportion of banks that are foreign. There are banks operating in the UK from at least 56 different countries.

Commercial banks

A commercial bank is a financial institution whose main business is taking deposits and making loans, though it may provide other financial services. Commercial banks offer services such as providing current and savings accounts, providing loans to individuals and businesses including mortgages, as well as other financial products such as insurance, pensions and financial advice. Traditional commercial banks operate a network of physical branches in convenient locations, although increasingly many commercial banking services are provided online.

HSBC and Lloyds Banking Group are examples of major UK commercial banks. Table 45.1 shows the ten largest banks in the UK, ranked according to the value of their assets. The UK's commercial banking industry is concentrated. This means

Key terms

A **commercial bank** is a financial institution whose main business is taking deposits and making loans, though it may provide other financial services such as insurance.

An **investment bank** specialises in complex financial activities, such as assisting governments and businesses to raise capital or advising businesses on mergers with other businesses.

Assets are items that are owned by a business. The assets of commercial banks include property, loans granted to customers, and investments in stocks, shares and Treasury bills.

that the industry comprises relatively few commercial banks, some of which are large. There are four very large commercial banks operating in the UK: HSBC, Barclays, Royal Bank of Scotland and Lloyds Banking Group. Between them, these four banks held a market share of 75% of current bank accounts of individuals in the UK in 2014.

Rank	Name of bank	Value of assets in 2014	Value of assets in 2013
1	HSBC	1732.99	1757.45
2	Barclays	1357.91	1343.63
3	Royal Bank of Scotland Group	1050.76	1027.88
4	Lloyds Banking Group	854.90	847.03
5	Standard Chartered	477.58	443.67
6	National Westminster	308.94	353.46
7	Santander UK	275.98	270.29
8	Nationwide	189.93	190.72
9	The Co-operative	40.39	45.05
10	Clydesdale and Yorkshire	38.60	38.50

Source: Banks around the World website

Table 45.1 UK commercial banks ranked by value of assets (£ billion), 2013 and 2014

Some new commercial banks have entered the UK market in recent years, a trend encouraged by the government to stimulate greater competition. These new entrants are termed ‘challenger banks’ and include the Metro Bank and Virgin Money. The challenger banks are very small in comparison with the UK’s so-called ‘big four’ banks.

Author tip

Although you need to be familiar with the functions, objectives and balance sheets of commercial banks, this is not the case with investment banks. You only need to have a broad understanding of what investment banks do and how they differ from commercial banks.

Investment banks

Investment banks primarily provide financial services to companies and governments. The banks’ services include raising capital through underwriting or assisting in the issue of new shares. They provide services to businesses that are involved in mergers and takeovers. They also offer financial advice and trade in a range of financial securities. Unlike commercial banks, investment banks do not accept deposits.

There are a number of UK specialist investment banks including Rothschild and Brewin Dolphin, but many of the UK’s commercial banks also have investment banking divisions. These include Barclays, HSBC and the Royal Bank of Scotland. London is a global centre for investment banking and many of the world’s largest international banks have operations in the city including JP Morgan, Deutsche Bank, Credit Suisse and Nomura.

The UK banking industry and systemic risk

Investment banking is much more risky than commercial banking. Some investment bank activities involve trading in complex financial securities that offer high returns but can also result in large losses being incurred.

It is partly because of the degree of risk attached to investment banking and the enormous cost to governments of offering financial support to banks if they are in danger of failing, that the UK and many other governments are passing laws to separate the investment and commercial or retail activities of banks.

Governments do not want to allow commercial banks to fail because there would be enormous negative implications for businesses, households, the banking system and the economy. If one bank failed, it could involve the loss of savings, pensions and other investments held by the bank. However, even more worrying is the potential for it to result in a loss of confidence in the banking system and the collapse of other banks and financial institutions as people and firms rush to withdraw funds. This threat means that banks represent systemic risk – the collapse of one could bring down the whole banking system and pose a threat to the economy.

The banking industry in the UK is particularly large in relation to the size of the economy. It is estimated that the UK authorities spent £850 billion supporting the economy's banks following the 2008 financial crisis. The possible failure of a commercial bank in the UK represents a systemic risk because it may result in severe economic instability or even pose a threat to the entire banking system and hence to the economy. The UK's banks were described at the time of the financial crisis in 2008 as being 'too big to fail'. This meant that the government could not allow them to fail because of their potential to cause extreme long-term damage to the economy.

Furthermore, many European governments are committed to protecting individuals' savings. In the case of the UK, savings are protected up to a maximum of £75,000 per person per bank. Thus, the cost of a failing commercial bank to the UK government is potentially huge. However, the consequences of a failed investment bank are not so great. Hence there have been moves to separate them, as described in Real World Economics 4.5.1.

Author tip

There are other institutions apart from commercial banks, investment banks and the central bank (in the UK this is the Bank of England, discussed in the next chapter) that operate in financial markets. These include insurance companies, investment trusts and building societies. While you should know that these financial institutions exist, you do *not* need to know about their functions or activities.

The main functions of a commercial bank

The main functions of commercial banks can be placed into four categories.

Key terms

Mergers occur when two or more firms join together to form a new, larger business.

Systemic risk exists, in finance, when there is a risk to the whole market, system or even economy, as opposed to a risk linked to one organisation.

A **takeover** is the purchase of a controlling interest in one business by another.

Underwriting is a process whereby investment banks support governments and companies in raising capital by issuing bonds and shares.

1 Accepting deposits

Individuals, businesses and other organisations all deposit funds with commercial banks. Money is deposited with banks for a number of reasons including security, the prospect of earning a return or for reasons of convenience. Commercial banks offer different types of bank accounts, in which money can be deposited, to meet these diverse needs. These can be separated into two categories.

- **Current accounts:** These are basic bank accounts held by individuals and firms, into which money is deposited regularly. Money can be withdrawn from these accounts without notice. Holders of these accounts are usually given debit cards and chequebooks to enable them to spend money in forms other than cash. Thus banks provide their customers with payment services, enabling them to spend money as necessary.
- **Savings accounts:** Firms and individuals use savings accounts to store surplus funds, and to earn interest until the money is needed. At the time of writing the rate of interest is very low on this type of account – often just 1% or 2% per annum.

The commercial banks in the UK have huge sums of money deposited with them at any time. In September 2015 the Bank of England reported that banks and other financial institutions (such as building societies) held deposits totalling £246.2 billion. Commercial banks have to use these deposits in a range of profitable activities to enable them to pay interest on the deposits.

Key term

Bank deposits are sums of money paid into accounts held with financial institutions such as commercial banks.

REALWORLD ECONOMICS 45.1

Investment banking activities to be operated as ‘if they are entirely unrelated companies’

The UK’s biggest banks will have to run their commercial banking operations as independent banks, almost entirely separate from their activities as investment banks. The Bank of England stressed that there will be a clear division between the two activities and that commercial banking activities will be ring-fenced.

As a result, the UK authorities hope the commercial banks will be able to continue running the high street operations with no difficulties even if their investment banking divisions get into trouble. Basic services such as payments

and access to bank accounts should be able to continue even if the parent group collapses.

The ring-fenced commercial banking divisions of the banks must hold bigger capital reserves to protect themselves against an economic downturn, and have their own independent IT, human resources, processing and risk teams.

The ring-fence rules apply to HSBC, Barclays, Royal Bank of Scotland, Lloyds Banking Group, Santander UK and the Co-operative Bank, as they all have core deposits of more than £25 billion. Challenger banks that

reach this size by 2019 also need to prepare for the changes.

From 2019, each UK commercial bank must treat its investment banking operation as if it is an entirely unrelated company.

Source: Adapted from the Daily Telegraph, 15.10.15

Exercises (13 marks)

- 1 Why does the Bank of England want commercial banks to operate their investment banking activities separately? (6 marks)
- 2 Explain why commercial banks are required to hold increased reserves of capital. (7 marks)

2 Advancing loans

Banks provide different types of loans to meet the needs of their customers. The main types of loans are as follows:

- **Overdrafts:** These are flexible short-term loans. They allow customers to overspend their current accounts up to an agreed limit.
- **Credit cards:** These are a popular form of loan for many consumers as they can be very flexible. From the point of view of commercial banks, they are very profitable as high interest rates are commonly charged.
- **Bank loans:** This type of loan is short to medium term and repaid over an agreed time period. These loans are granted to businesses as well as individuals.
- **Mortgages:** These are long-term loans provided for the purpose of buying property. They are repaid over periods of up to 40 years.

Key term

Mortgages are long-term loans provided for the purpose of buying property.

The amount that commercial banks can lend is based on the amount of deposits they attract, and the interest received on loans is a major source of income. Lending is a major activity for commercial banks. For example, in September 2015, banks in the UK had outstanding loans to businesses totalling £36.5 billion.

Deposits and loans: the core of the commercial banks' business model

Commercial banks provide credit to the economy in the form of loans or advances. This is a vital function: they take deposits from customers and make them available to others in the form of loans. Those borrowing from banks, especially businesses, frequently require large medium- to long-term loans. However, customers frequently make relatively small deposits and may wish to have swift access to their funds. The major distinction between depositors and borrowers is maturity. Banks have to turn large numbers of small deposits with a short maturity into fewer large loans available over longer terms.

Commercial banks have to manage this 'maturity transformation' to ensure that they have sufficient liquid assets to meet the demands of customers to withdraw funds that they have deposited on a day-to-day basis. We explore this issue further later in this chapter, when we analyse the structure of banks' balance sheets.

3 Providing a range of other financial services

Increasingly, commercial banks provide a range of financial services to supplement the basic banking activities of accepting deposits and advancing loans. These activities may include:

- providing diverse insurance products for individuals and businesses, including life assurance as well as property and vehicle insurance;
- offering expert financial advice in relation to savings, buying shares and pensions.

4 Investing funds

Commercial banks invest some of the funds they hold in a range of securities to generate high returns whenever possible. This might include corporate and government bonds and shares as well as other, more complex, financial securities. In 2013, 15% of the assets held by UK-owned banks were in the form of financial securities.

The structure of a commercial bank's balance sheet

A balance sheet is a financial statement recording the assets and liabilities of a commercial bank (or any other business) on a particular day, normally at the end of an accounting period. The balance sheet only represents a picture of a bank's assets and liabilities at a moment in time: it is commonly described as a 'snapshot' of the financial position of an organisation. Thus balance sheets always carry a date on which the valuation of assets and liabilities took place.

By recording assets and liabilities, the balance sheet sets out the ways in which the bank has raised its capital and the uses to which this capital has been put. A commercial bank's balance sheet shows a bank's sources of funds – these are its liabilities and capital. The other side of the balance sheet records the uses to which the bank has put these funds – these are its assets. Since a balance sheet must 'balance', these two must be equal. This relationship is shown in Figure 45.1.

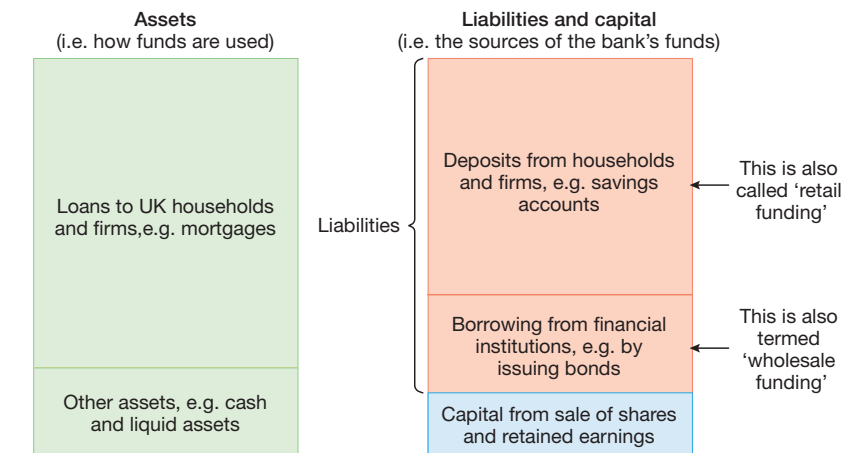


Figure 45.1 The basic structure of a commercial bank's balance sheet

A commercial bank's liabilities and capital

The balance sheet sets out the liabilities that have been incurred by a commercial bank at a particular point in time. This will include the following components:

- **Retail funding:** This is the money that is deposited with banks by firms and households in their current or savings accounts. Some of these funds will be in accounts that allow immediate withdrawal and therefore may only be available to the bank for a very short period of time.
- **Wholesale funding:** Banks also borrow funds from institutions such as pension funds and insurance companies that have large amounts of their customers' money to invest. It is common for banks to issue bonds on the capital markets to

Key terms

A **balance sheet** is a financial statement recording the assets (possessions) and liabilities (debts) of a business on a particular day at the end of an accounting period.

A **liability** is a debt that must eventually be paid. Thus it represents a claim on assets.

Liquidity measures the ease with which assets can be turned into cash.

Capital is the finance raised through issuing shares and retaining earnings from previous trading periods.

attract this source of finance. In addition, a commercial bank may borrow from other banks, making the banking system highly interdependent.

A bank also uses capital as a source of finance. This is finance that is raised by the sale of shares on the Stock Exchange and profits that have been retained (and not paid out to shareholders as dividends) in previous financial years.

Capital is often considered to be a bank's 'own funds' as opposed to those that it has borrowed in the form of customer deposits, for example. Capital is important because it shows a bank's ability to withstand losses, as this source of funds does not have to be repaid. We shall see later that having sufficient capital is vital for a bank's long-term financial security.

A commercial bank's assets

A bank's assets are the items that it owns. These can be physical assets – for example, commercial banks own valuable properties in many high street locations. Other assets can be intangible (i.e. not have a physical existence) – the loans granted to households and businesses and other banks are probably the best example of this type of asset.

The importance of liquid assets

A bank's balance sheet must contain sufficient liquid assets to ensure it can meet day-to-day needs for cash. Liquid assets are those that can be easily turned into cash, which is the most liquid of all assets. A commercial bank is likely to hold a range of liquid assets including:

- cash in its tills;
- deposits at the Bank of England (these are really the commercial banks' current accounts);
- short-term securities such as Treasury bills.

Since the 1980s, banks in the UK have invested much more heavily in purchasing securities such as shares and bonds. However, an even more noticeable trend is the increase in holding derivatives. 'Derivatives' is a broad term covering a range of complex financial securities that can be very profitable but also risky.

UK banks are obliged to hold sufficient cash reserves to meet the liquidity coverage ratio (LCR). This is essential to protect a bank against a situation in which a large number of its customers withdraw cash from their accounts at short notice. If a bank has insufficient funds to meet this demand for cash, it may be forced to sell assets at unfavourable prices. If there were thought to be a risk of a bank having insufficient cash, this could lead to a loss of confidence by its customers and a 'bank run'. This would result in large queues outside banks as customers sought to withdraw cash.

How banks create credit

The Bank of England recognises that most of the money in the UK economy comprises bank deposits. In an article published in 2014 entitled 'Money in the modern economy: an introduction', it noted that '97% of the money held by the public is in the form of deposits with banks, rather than currency'. These bank deposits are money rather than simply liabilities of the bank owed to households and firms. The reason for this is that firms and households receive incomes in the form of credits in their current

Key term

The **liquidity coverage ratio (LCR)** requires banks to hold sufficient high-quality liquid assets to exceed the net cash outflows of the next 30 days.

accounts and spend much of this money by transferring it to the accounts of other firms and households using methods such as direct debits and debit or credit cards. Thus the Bank of England argues that bank deposits are performing more of money's functions: they have become a medium of exchange as well as a store of value.

In the Bank of England's quarterly bulletin, some of the Bank's employees engaged in monetary analysis argue that the principal way in which money is created is by commercial banks taking decisions to grant loans to their customers. Whenever a bank makes a loan, it creates an equivalent deposit in the borrower's bank account, thereby creating new money:

'Commercial banks create money, in the form of bank deposits, by making new loans. When a bank makes a loan, for example to someone taking out a mortgage to buy a house, it does not typically do so by giving them thousands of pounds worth of banknotes. Instead, it credits their bank account with a bank deposit of the size of the mortgage. **At that moment, new money is created.** For this reason, some economists have referred to bank deposits as 'fountain pen money', created at the stroke of bankers' pens when they approve loans.'

Bank of England (2014) 'Money creation in the modern economy'

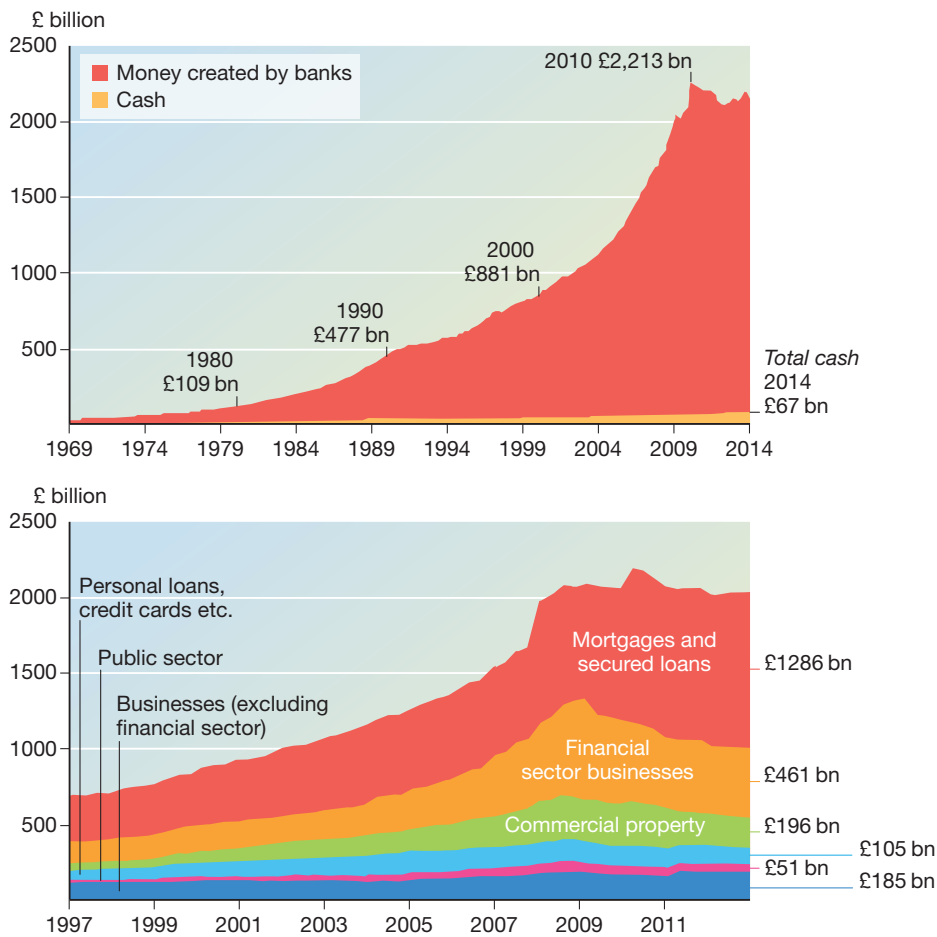


Figure 45.2 The extent to which UK banks have created credit since 1969 and the uses to which it has been put

Source: Positive Money website

The decision to offer a loan to a household or firm creates an equivalent deposit in the borrower's bank account. The bank's balance sheet will balance following the decision to advance a loan because the loan will appear at the same time as an asset held by the bank.

Figure 45.2 shows the extent to which banks have created credit in the UK. The vast majority of the money in circulation is in the form of bank deposits and this form of money is becoming comparatively more important. It is also interesting to note that an increasing proportion of the credit created has been lent to businesses in the financial sector. Some of these loans, which peaked before the financial crisis started in 2007, are in the form of high-risk securities such as derivatives.

The credit multiplier

Key terms

The **credit multiplier** measures the multiple by which the expansion in the money supply is greater than the increase in the monetary base – high-powered money.

Base money is the banknotes in circulation along with the balances or reserves held by commercial banks and building societies at the Bank of England. This is also referred to as central bank money.

In the past the monetary authorities in the UK have required commercial banks to hold a proportion of their assets as reserves in the form of high-powered money. This requirement set a minimum percentage of customers' deposits that every commercial bank must hold as reserves. This percentage is also called the 'reserve assets ratio'. Most other major economies, with the exception of Canada, still operate reserve asset ratios, though these are not always used actively as part of the monetary policies of these countries.

The operation of a reserve assets ratio gives rise to the existence of the credit multiplier, also known as the 'money multiplier'. Any increase in the commercial banks' holdings of base money in the form of reserves under this system can bring the credit multiplier into operation. This multiplier measures the multiple by which the expansion in the money supply is greater than the increase in the monetary base: if the multiplier is 8, then a £1 increase in the monetary base will cause an £8 increase in the money supply.

Changes in the reserve ratio that is required can be used as a tool of monetary policy. By altering the requirement to hold reserves, this action by the central bank can have a multiplied impact on the commercial banks' lending and hence its ability to create credit.

The limits on the ability of commercial banks to create credit

We have seen that commercial banks have the ability to create credit by advancing loans to customers. However, despite the absence of a reserve assets ratio in the UK, they cannot do so endlessly. This is the case because they are dependent on a demand for credit from firms and households. A key factor that will determine the demand for credit is its price (i.e. the rate of interest charged on loans). The Bank of England is able to influence the rate of interest charged on new loans, not least through decisions on bank rate taken each month by the Monetary Policy Committee.

The Bank of England does not determine the quantity of loans (and hence deposits) that commercial banks can create by saying that there must be a relationship between liquid assets held by banks (often termed 'reserves') and the amount of loans the bank may provide. There are no reserve requirements in the UK currently and this relationship therefore has no relevance for banks in the economy. Rather than controlling the quantity of reserves held by the commercial banks, the Bank of England implements its monetary policy through interest rates.

The objectives of a commercial bank

Commercial banks generally operate with a range of objectives. For example, they may want to increase their share of the market or develop a respected brand to enable them to sell related financial services. However, they are likely to have three important objectives, as discussed below.

1 Maintaining sufficient liquidity

Liquidity measures the ease with which assets can be turned into cash. Banks hold some liquid assets, most obviously the cash in their branches, and some less liquid ones, for example long-term loans such as mortgages.

We saw earlier in this chapter that commercial banks in the UK are required to manage their business so as to meet the requirements of the liquidity coverage ratio (LCR). This requires that commercial banks hold sufficient high-quality liquid assets by January 2019 to exceed the net cash outflows over the next 30 days. The objective of the LCR is to ensure banks are resilient against the risk of liquidity crises.

High-quality liquid assets that can be held by commercial banks include the following:

- coins and banknotes;
- reserves held at the Bank of England;
- liquid securities such as Treasury bills;
- corporate bonds not issued by other banks.

Commercial banks in the UK seek to ensure that they hold sufficient assets in a liquid form and the requirements of the LCR represent a minimum.

2 Profitability

Most commercial banks in the UK are public limited companies. This means that they are owned by their shareholders and that their shares are bought and sold on the London Stock Exchange. Many of the individuals, firms and institutions that purchase shares in the UK's commercial banks will do so for financial reasons.

For commercial banks to maximise profitability, they should aim to use their funds in ways that generate the highest possible returns. This helps to explain why banks have invested increasing proportions of their funds in financial securities. Short-term loans, frequently not secured, can also generate high returns for banks.

Discussion point

Do you think that profitability is the most important objective for a commercial bank in the UK?

3 Security

Security in the context of the objectives of commercial banks refers to the extent to which they manage the risk associated with their lending activities. Banks can take

Key terms

Objectives are medium- to long-term goals that are agreed to coordinate the activities of an organisation.

Profitability is a measure of a business's performance that

compares profits to another factor such as earnings.

Dividends are a share in the profits of a company that are paid to some groups of shareholders.

Secured loans allow individuals and firms to borrow a specific sum of money that is secured against a property or other valuable asset, which is sold in the

event of the borrower defaulting.

A **creditor** is an individual or organisation that is owed money by another person or organisation.

a number of steps to reduce the risk associated with their lending activities, such as possessing sufficient liquid assets and having enough capital.

Capital is not borrowed money; it takes the form of funds raised by banks selling shares or profits retained from earlier trading periods. Capital can help a commercial bank to mitigate losses incurred in its trading – for example, if the holders of unsecured loans default on repayment. If a commercial bank incurs losses and does not hold sufficient capital, it may not be able to repay its creditors. Capital offers other advantages in that, unlike other sources of funds such as loans, it does not ever have to be repaid. Furthermore, payments to shareholders depend on the profits generated by the business. If the bank is not profitable, it does not have to pay dividends to its shareholders, whereas it would have to pay interest on loans. The Lloyds Banking Group in the UK did not pay dividends to its shareholders following the financial crisis from 2009 until 2015. However, over time capital is a more expensive source of funds for commercial banks than liabilities such as debt.

Commercial banks are subject to the risk of two types of loss:

- **Expected losses:** These are losses that can be forecast based on past history and the types of loan a bank is making. If it makes a higher proportion of high-risk loans, then its expected losses will increase proportionately.
- **Unexpected losses:** The past is not always a good indicator of future events. For example, economies may be subject to significant negative demand- or supply-side shocks. These can result in an unexpectedly high proportion of the bank's customers defaulting on the repayment of loans. The financial crisis is a major and recent example of a factor causing unexpected losses.

Commercial banks have to make sure that they have sufficient capital to meet both expected and unexpected losses. The UK's commercial banks are subject to regulation regarding the amount of capital they must hold. These laws apply to all commercial banks in the European Union and meet international standards that were agreed by the Basel Committee on Banking Supervision in 2009. We will consider the issue of regulation of banks more fully in Chapter 47.

REALWORLD ECONOMICS 45.2

Proposal to make creditors accept the risk of commercial banks failing

The Financial Stability Board, which advises the G20 group of countries, said that the world's 30 biggest banks should raise up to \$1.2 trillion in loss-bearing debt on top of their capital. By compelling banks to issue debt that can be drawn on in times of trouble, the goal is to make creditors who lend the debt liable for a bank's failure. This would protect the country's taxpayers from having

to bail out banks, as in 2008.

Under the proposals, a bank would fund at least 18% of its assets using this form of debt as well as capital by 2022. Mark Carney, the Governor of the Bank of England, who also chairs the FSB, said that if this proposal were adopted by regulators, the plan would 'support the removal of the implicit public subsidy' enjoyed by

banks that are 'too big to fail'.

Source: Adapted from *The Economist*, 14.11.15

Exercise

- 1 Explain the possible reasons why the adoption of this proposal might please governments and taxpayers, but not banks and their shareholders. (7 marks)

The potential for conflict between the objectives of commercial banks

It is perhaps unsurprising that the objectives of commercial banks have the potential to conflict with one another.

Liquidity versus profitability

Liquid assets tend to be less profitable than illiquid ones. Banks may earn little or no return on cash. Other liquid assets, such as Treasury bills, earn a low rate of return because they represent a very low degree of risk and can easily be converted into cash on the UK's money markets. On the other hand, less liquid assets can be expected to generate much higher returns.

Thus banks need to hold sufficient liquid assets to ensure that they can satisfy the demands of their customers for cash whenever necessary, but not so much as to reduce their profitability. Profits are important for shareholders, and those managing the UK's commercial banks have to balance the needs of these two groups of stakeholders.

2 Profitability versus security

The international agreements reached at Basel require the UK's commercial banks to hold sufficient levels of capital (funds from the sale of shares and retained profits) to mitigate expected and unexpected losses. However, over time this source of funds is more costly than capital raised in the form of loans – whether from customers or the capital markets. Thus, without regulation and the need for security, banks would prefer to hold lower amounts of capital because this enhances their profitability.

Review questions

Total: 30 marks

- 1 Briefly explain the difference between a commercial bank and an investment bank. **(5 marks)**
- 2 Define the term 'systemic risk' in relation to finance. **(3 marks)**
- 3 Explain one reason why the UK government cannot allow a major commercial bank to fail. **(4 marks)**
- 4 Which of the following is not a function of a commercial bank?
 - A Accepting deposits from customers
 - B Granting loans to households and businesses
 - C Investing funds
 - D Offering advice in relation to mergers and takeovers **(1 mark)**
- 5 Which of the following would not appear as an asset on the balance sheet of a commercial bank?
 - A Deposits from customers
 - B Loans to customers
 - C Cash in the bank's tills
 - D Holdings of Treasury bills **(1 mark)**

- 6** Explain, using examples, the difference between a commercial bank's liability and its capital. **(6 marks)**
- 7** Which of the following is an example of retail funding provided to commercial banks?
- A A bond issued by the commercial bank
 - B Money deposited in a customer's savings account
 - C Money lent to the bank by a pension fund
 - D Money lent to the bank by another commercial bank
- (1 mark)**
- 8** Briefly explain how commercial banks can create money. **(4 marks)**
- 9** Which of the following is most likely to be considered an illiquid asset by a commercial bank?
- A Cash and notes held in the bank's tills
 - B Its holdings of Treasury bills
 - C Loans advanced to customers in the form of mortgages
 - D The money in the commercial bank's accounts at the Bank of England
- (1 mark)**
- 10** Describe one objective of a commercial bank. **(4 marks)**

Central banks & monetary policy

Key concepts from Year 1

The Year 1 companion textbook covered monetary policy in Chapter 43.

- We investigated the objectives of the UK authorities in implementing monetary policy and the different forms that monetary policy may take. These forms included changes to the rate of interest, controlling the money supply and influencing the exchange rate.
- We examined the role of the Monetary Policy Committee in determining and implementing the UK's monetary policy and the factors that influence its decisions regarding bank rate.
- Finally, the chapter considered the role that the exchange rate can play in monetary policy by influencing the level of aggregate demand.

In this chapter we describe the main functions of a central bank, with particular reference to the Bank of England. We examine the monetary policy transmission mechanism in the UK and the relationship between changes in interest rates and exchange rates. We also consider the ways in which the Bank of England can influence the growth of the money supply in the UK.

The main functions of a central bank

The UK's central bank is the Bank of England and we shall use this as an example of the functions carried out by a central bank. The Bank of England was established in 1694 primarily to act as a bank for the government. Over the three centuries since its foundation, the role of the Bank of England has developed and expanded.

The Bank of England's mission statement (i.e. its overall aim) is clearly focused on its management of the UK's financial system:

'The Bank of England's mission is to promote the good of the people of the United Kingdom by maintaining monetary and financial stability.'

Bank of England website

Key terms

A **central bank** is responsible for managing the monetary system in an economy.

Monetary policy refers to actions taken by the government (or the central

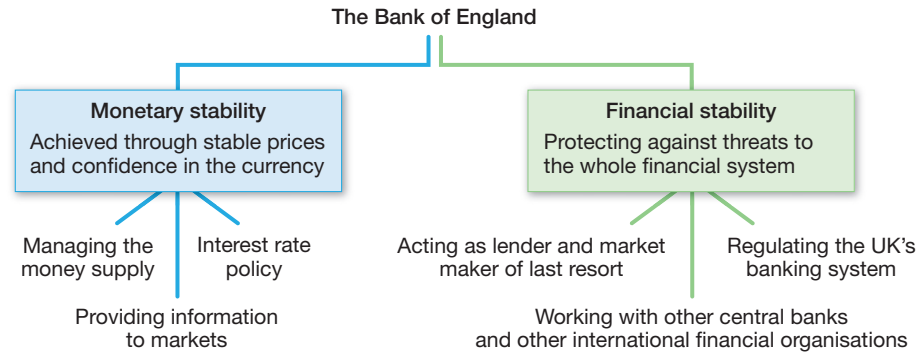
bank acting on its behalf) to manipulate interest rates, the supply of money and the exchange rate to achieve its macroeconomic objectives.

The **Monetary Policy Committee (MPC)** is a group of nine experts in monetary issues who meet monthly to make a decision on the UK's bank rate.

The **International Monetary Fund** is an organisation with 188 member countries that aims to encourage global monetary cooperation, secure financial stability and promote international trade.

In order to fulfil this mission of maintaining monetary and financial stability, the Bank carries out a diverse range of functions, which are summarised in Figure 46.1.

Figure 46.1 The functions of the Bank of England



We consider the major functions of the Bank of England in the following section.

1 Managing the UK's money supply

The Bank of England designs and issues the UK's coins and its bank notes. These are both designed to be durable, and the banknotes contain security features that are easy to check and difficult to counterfeit. The Bank has announced that it is to issue its first plastic banknotes in 2016, which will remain in circulation for a longer period and will contain more advanced security features.

The Bank controls the amount of notes that are in circulation at any time and this can affect the money supply. However, notes are only a tiny proportion of the UK's money supply – approximately 3%. The Bank can influence the UK's money supply in other, more important, ways, as we shall see later in this chapter.

2 Setting the rate of interest

Since 1997 the Bank of England has been given responsibility for setting the UK's bank rate. This is an important rate of interest that influences rates charged throughout the UK economy. It also influences the level of activity in the economy.

The nine members of the Monetary Policy Committee (MPC) of the Bank of England meet each month to set a rate of interest that will meet the government's 2% inflation target. Since it takes up to two years for a change in bank rate to have its full effect on the UK's rate of inflation, the MPC is setting interest rates based on its forecast for inflation in two years' time. Thus, the MPC's decisions on bank rate in summer 2016 will have been based on forecasts for inflation in 2018.

3 Providing information to markets

Since August 2013, the Bank of England has provided information to markets in the form of forward guidance. The aim of forward guidance is to enhance the effectiveness of the bank's monetary policy by providing financial market participants (such as firms and households) with more information about the future path of monetary policy.

4 Acting as 'lender of last resort'

This is an important part of the Bank's aim of maintaining financial stability in the UK.

In acting as ‘lender of last resort’, the Bank provides funds if commercial banks suffer a lack of liquidity. This is an important function because it helps maintain confidence in the UK’s banking system. The Bank rarely acts as lender of last resort. It did so in 2007 when it acted to save Northern Rock. This was the first time it has acted in this role for over 30 years. The lender of last resort function comprises a tailor-made rescue package that has to be agreed by the Bank of England in conjunction with the Treasury and relevant regulators before use.

During the financial crisis that began in the UK in 2007, this function of the Bank expanded as it became the ‘market maker of last resort’. This meant that it provided liquidity to important markets, such as the capital markets in which mortgages were provided, to help to ensure financial stability. The Bank has operated, since January 2009, an Asset Purchase Facility (APF) financed by the issue of Treasury bills with the intention of improving liquidity in the capital markets.

5 Regulating the UK’s banking system

Since the financial crisis, the authorities in the UK have moved to strengthen the regulation of banking in the UK. A new regulatory structure was introduced into the UK with effect from April 2013. This established the Prudential Regulation Authority (PRA) as a part of the Bank of England. The PRA is responsible for the regulation and supervision of around 1700 banks, insurance companies and other financial institutions. The PRA operates with the general objective of promoting the safety and soundness of the organisations that it regulates.

6 Acting as banker to other banks and the government

The Bank of England provides banking facilities to the UK’s commercial banks. Commercial banks, along with all other banks offering credit in the UK, must hold an account at the Bank of England. We shall see later in this chapter that the balances held in these accounts can influence the ability and willingness of these banks to create credit.

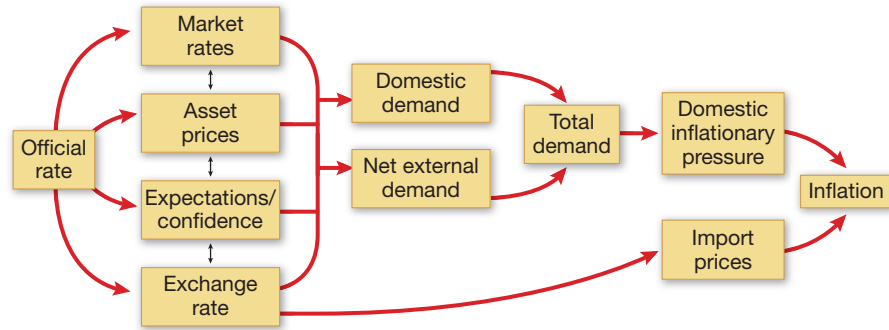
The Bank is the UK government’s banker. The government keeps its accounts with the Bank, including its main account known as the consolidated fund account. The Bank also manages the country’s reserves of gold and foreign currencies.

The monetary policy transmission mechanism

One important way in which monetary authorities in the UK and elsewhere implement monetary policy is by changing interest rates. The link between a change in the bank rate in the UK and subsequent changes in inflation is an example of the monetary policy transmission mechanism in operation. This is illustrated in Figure 46.2, which shows how a change in bank rate feeds through the economy to influence the rate of inflation.

Decisions by the Bank of England to alter interest rates have ‘knock-on’ effects on other interest rates in the economy. Simultaneously, announcements by the Bank of England, possibly as part of its forward guidance, affect confidence and expectations of households and firms. In turn households and firms take decisions that impact on the level of aggregate demand in the economy. Monetary policy operates principally through its effect on aggregate demand by determining the money value of goods and services.

Figure 46.2 A summary of the transmission mechanism of monetary policy



Source: Bank of England

The stages in the transmission mechanism

There are three stages in the transmission mechanism.

Stage 1: The effects on financial and asset markets

The Bank of England implements its monetary policy by lending to the money market at the rate chosen by its MPC. This rate is transmitted across all the different elements of the UK's money markets. This change almost immediately affects rates charged on short-term loans such as overdrafts for households and firms, and commercial banks will alter their base rates that determine interest rates on all their loans.

Changes in interest rates will affect the market value of financial assets such as shares and bonds. We saw in Chapter 44 that there is an inverse relationship between interest rates and the price of bonds. Thus a rise in interest rates will reduce the price of bonds. It would also normally have the same impact on the price of shares, as returns on low-risk savings accounts become more attractive, potentially reducing demand for shares.

Changes in interest rates can also affect the exchange rate of the economy's currency. A rise in the rate of interest will result in an increase in the exchange rate of the pound, and a fall in interest rates in the UK will reduce the pound's value against other currencies, *ceteris paribus*. Following an interest rate rise, money will flow from low interest rate yielding countries into higher interest rate yielding countries as investors look to make the highest return. Thus an increase in interest rates in the UK is likely to lead to an inflow of funds into the UK. The outcome will be an increase in demand for sterling on foreign exchange markets and a higher price or exchange rate for the currency.

Finally, changes in interest rates affect expectations and confidence of firms and households with regard to future transactions in financial markets and in other parts of the economy. The precise effects on expectations and confidence can vary, but an unexpected rise could be likely to lower expectations and dent confidence.

Stage 2: The effects on aggregate demand

A rise in interest rates could be expected to reduce aggregate demand. It will reduce the level of consumption by households as borrowing becomes more expensive and the returns on saving increase. At the same time, investment by firms will fall as the cost of borrowed funds increases, making fewer projects viable. A fall in interest rates would have the opposite effects.

There would be a ‘wealth effect’ arising from a change in interest rates. A fall in interest rates tends to increase the value of financial assets such as bonds and shares, but also other assets such as property. Cheaper borrowing can increase the demand for, and price of, houses, for example. A positive wealth effect such as this can lead to a rise in consumption.

The change in the exchange rate that follows from interest rate changes will also affect aggregate demand. Assume that the MPC decides to increase interest rates at one of its monthly meetings. Although the precise effects of an interest rate change on exchange rates are difficult to predict, it is likely that a rise in interest rates will result in an increase in the exchange rate of the pound sterling. A higher interest rate in the UK will make sterling financial assets more attractive to overseas investors. This will increase the demand for pounds and hence its price – the exchange rate. The rise in the exchange rate will increase the price of exports and reduce sales after a time lag, depending on their price elasticity of demand. Simultaneously, the price of imports will fall, fuelling demand, though once again the extent of this will be dependent upon price elasticity of demand for the products concerned.

The combined impact of the fall in exports and the rise in imports will be to reduce the value of ‘net exports’. Remember, net exports are the $(X - M)$ element of aggregate demand. Thus a rise in interest rates will reduce net exports and contribute to the overall fall in the level of aggregate demand in the economy.

Stage 3: The effects on inflation

The spending decisions of households and firms, outlined above, will have an impact on aggregate demand in an economy. Rising interest rates will tend to reduce it; falling rates will have the opposite effect. Rising aggregate demand will increase inflationary pressures as it increases the level of real GDP, at least in the short term. Furthermore, the change in interest rates will affect the rate of inflation in the UK directly via import prices.

However, these effects will not be felt immediately. We discuss the duration of time lags in relation to real GDP and the rate of inflation in the section below. Time lags occur because the effects of changes in interest rates are not felt straightaway. A fall in interest rates may not reduce consumers’ mortgage repayments for several months, and firms may reduce stocks of goods initially before increasing production. The impact via net exports may also be delayed because firms may be supplying goods and services under contracts that extend several months into the future. The price effects and subsequent impact on demand may actually take months to emerge.

The timing of the effects of monetary policy

At its simplest, monetary policy works by altering the level of aggregate demand in an economy relative to its ability to supply goods and services. If monetary policy makes goods and services scarcer relative to aggregate demand, it is likely to increase inflationary pressures.

The Bank of England uses a macroeconomic model to simulate and measure the size and timing of the impact of changes in bank rate on the economy. The Bank’s research shows the effects on GDP and inflation of an *unexpected* 1% rise in bank rate lasting for *one year*.

- **Effects on real GDP:** This interest rate increase would be likely to reduce GDP quite quickly, reaching a maximum fall of between 0.2% and 0.35% of GDP after approximately 15 months. GDP would return to its original level after about three years, *ceteris paribus*.
- **Effects on inflation:** The effects on inflation take longer to appear. Little changes in the first year after the interest rate rise. In the second year, however, the Bank of England predicts that the rate of inflation will fall sharply by between 0.2 and 0.4 percentage points.

Differing views of the effectiveness of the monetary transmission mechanism

It will not be a surprise that classical and Keynesian economists have different views on the way in which the effects of interest rate changes are transmitted to the economy. One key point of dispute between the two schools of thought is the extent to which planned levels of consumption and investment in the economy are responsive to changes in interest rates. Classical economists argue that both consumption and investment are interest rate elastic. Thus they are responsive to a change in interest rates. This means that the effects on aggregate demand of a change in interest rates will be significant.

In contrast, Keynesian economists argue that consumption is highly dependent on the level of income and not the cost of borrowing. Some Keynesians also argue that factors other than interest rates are the main determinants of investment. This means that the level of aggregate demand will not change significantly following a change in interest rates. This, of course, means that the impact of interest rate changes on the economy, at least through routes other than the exchange rate, are muted. In part, this explains why some Keynesian economists favour the use of fiscal policy to manage the economy.

The Bank of England and the supply of money

The Bank of England can influence the growth of the money supply in a number of ways.

1 The price of money – interest rates

In Chapter 45 we discussed how the ability of banks to create credit by advancing loans to firms and households is one of the major ways in which the UK's supply of money can be increased. The commercial banks do not have an infinite ability to create credit or money by making loans. They are dependent upon there being a demand for new loans that they may be able to meet. The Bank of England can influence the demand for loans by influencing the rate of interest that is charged for such loans. A change in bank rate will have a 'knock-on' effect on other rates, such as those for mortgages as well as business and personal loans. This, in turn, will result in a change in the demand for credit, affecting the ability of the commercial banks to create money.

2 Funding for Lending

The Bank of England launched its Funding for Lending Scheme (FLS) in 2012 in conjunction with the Treasury. The FLS offers incentives to commercial banks and

building societies to boost their lending to UK firms and households with the aim of increasing lending by up to £70 billion. In 2014 the FLS was redesigned to focus more on the needs of small businesses for loan capital. In November 2015, the FLS was extended until 31 January 2018, allowing banks additional time to draw on funding available for loans to small and medium-sized businesses.

Commercial banks and building societies swap financial assets they hold, such as loans, with the Bank of England. In return it provides them with Treasury bills, for up to a four-year period. The commercial banks are then able to use the Treasury bills as security to borrow cash in financial markets. Since the commercial banks have the guarantee of government funds behind them, they can borrow funds at very cheap rates.

This mechanism has potentially added significantly to the growth of the money supply in the UK. By the end of June 2015, commercial banks and building societies in the UK had accessed funds from the Bank of England totalling £61,434 million. However, this may not have had as much of an impact on the growth of the money supply in the UK as the total figure would suggest because much of the FLS funding replaced more traditional bank lending. Real World Economics 46.1 suggests that much of it may have simply been a replacement.

3 Open market operations

The Bank of England can use open market operations to alter the level of liquidity available to commercial banks. It can conduct open market operations in two ways.

Buying or selling government bonds

The Bank of England can conduct open market operations by buying or selling government bonds on the financial markets. If it buys bonds back from commercial banks, it can increase their liquidity and increase the amount of central bank money they hold in their accounts at the Bank of England. This increases their liquidity and enables them to increase lending. Buying bonds in this way can also help to lower interest rates, thereby potentially increasing the demand for loans as well as the commercial banks' ability to supply them. Thus the money supply is increased with the aim of stimulating economic activity. Selling bonds can have the opposite effect.

Entering into a repo

A repo is a sale and repurchase transaction. This entails a commercial bank or building society selling an asset (for example, a Treasury bill) to the Bank of England and agreeing to repurchase it. Repos tend to be short-term transactions, so the duration of the deal can be as little as 24 hours. In return, the Bank of England adds liquid

Key terms

Open market operations are actions undertaken by central banks to provide or withdraw liquidity from one or more commercial banks.

Central bank money comprises the banknotes in circulation along with the balances or reserves held by commercial banks and building societies at the Bank of England.

Government bonds are debt issued by the Bank of England on behalf of the UK government. These normally receive a fixed interest payment and have a lifespan in excess of one year.

A repo is a sale and repurchase transaction of a financial asset between one or more commercial banks and the Bank of England.

funds to the commercial bank's account with it. This can encourage the commercial bank to increase lending while remaining prudent. The Bank of England can also influence general interest rates by the interest it charges on its repo (called the 'repo rate'). In this way it can influence the demand for loans as well as the potential of commercial banks to supply them. The outcome is that it indirectly influences the money supply.

When the Bank of England wants to tighten money supply, it sells the financial security first and buys it back later. This is called a reverse repo and is an agreement to lend financial securities rather than liquid funds.

4 Quantitative easing

Key term

Quantitative easing (QE)

is a form of monetary policy where a central bank creates new money electronically to buy financial assets.

The Bank of England defines quantitative easing (QE) as 'an unconventional form of monetary policy where a central bank creates new money electronically to buy financial assets, like government bonds'. QE was first used by the central bank in Japan in the 1990s in an attempt to end a period of deflation. Economists disagree about the effectiveness of the policy.

The objective of quantitative easing in the UK is to create a level of aggregate demand in the economy consistent with meeting the Bank's 2% inflation target in the medium term. It operates by boosting the money supply through large-scale asset purchases from financial institutions. This increases lending along with consumption and investment and hence the level of economic activity. The operation of QE is illustrated in Figure 46.3.

REALWORLD ECONOMICS 46.1

Funding for Lending Scheme fails to help small firms

The Funding for Lending Scheme established by the Treasury and the Bank of England to encourage banks to make loans to small companies has been branded a failure after new statistics yesterday showed that lending to the sector continued to fall throughout last year.

Figures released by the Bank of England showed net lending under the Funding for Lending Scheme slumped by £810 million in the final quarter of 2014. This followed falls in the previous three quarters and took the total decline over 2014 to just under £2 billion.

The FLS channels cheap funding from the Bank of England to commercial banks on the condition

that they pass on the money to borrowers. It was launched in the summer of 2012 to boost mortgage and business loans but was refocused in 2014 with the specific aim of encouraging new lending to small companies.

However, the Bank's data show that despite participating lenders drawing down £15.6 billion in cheap loans in 2014, they collectively carried on reducing their overall levels of business lending. As well as the £2 billion decrease in net lending to small businesses, net lending to large firms declined by £14 billion.

John Allan of the Federation of Small Businesses observed that small companies were bypassing

banks when looking for funds, despite the FLS: 'Evidence shows that rather than turn to their banks, they are increasingly using their own resources to meet their financing needs, and paying down their debts rather than increasing them,' he explained.

Source: *Independent*, 27.2.15

Exercises **Total: 11 marks**

- 1 Why did the government establish the Funding for Lending Scheme in 2012? (5 marks)
- 2 Explain the apparent effect of the Funding for Lending Scheme on the money supply in the UK. (6 marks)

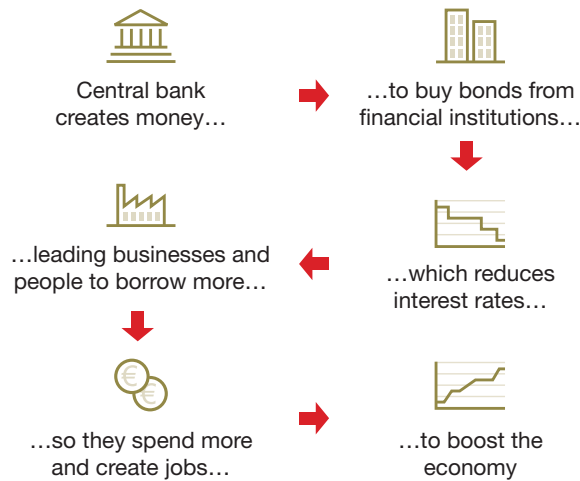


Figure 46.3 How quantitative easing works

The Bank of England first used QE in the UK between March and November 2009 at a time when the UK economy was in recession and it urgently needed to stimulate the economy. Decisions regarding the use of QE are made by the MPC at the Bank. During 2009 the Bank purchased £200 billion of financial assets in an attempt to boost the economy. At this time interest rates were at 0.5%, so less conventional methods of monetary policy were needed. More recently the MPC has decided on further purchases: £75 billion in October 2011 and £100 billion between February and July 2012. Thus, the total assets purchases under QE amount to £375 billion.

The Bank of England considers the use of QE to have been effective and to have had a significant impact on the rate of growth (and inflation) in the UK economy.

Review questions

Total: 26 marks

- 1 What is the difference between a central bank and a commercial bank? **(5 marks)**
- 2 What is the Bank of England's mission? **(2 marks)**
- 3 Which of the following is not a function of the Bank of England?
 - A Setting interest rates
 - B Acting as lender of last resort
 - C Managing the consolidated fund
 - D Issuing corporate bonds**(1 mark)**
- 4 Define the term 'monetary policy transmission mechanism'. **(3 marks)**
- 5 Why might a rise in interest rates affect the value of bonds and company shares? **(6 marks)**
- 6 Which of the following would be most likely to result from a fall in interest rates?
 - A A fall in the volume of UK exports
 - B A fall in the volume of UK imports
 - C A fall in the price of bonds on capital markets
 - D A fall in the price of Treasury bills**(1 mark)**

- 7** Which of the following offers financial incentives in the form of cheap borrowing to commercial banks and building societies intended to boost their lending to UK firms and households?
- A The Funding for Lending Scheme
 - B Quantitative easing
 - C Open market operations
 - D The monetary transmission mechanism
- (1 mark)*
- 8** Define the term 'central bank money'.
- (3 marks)*
- 9** Which of the following is most likely to result in a reduction in the money supply, ceteris paribus?
- A A repo transaction involving the purchase of Treasury bills by the Bank of England
 - B The sale of government bonds through open market operations
 - C The use of quantitative easing to buy bonds
 - D A purchase of government bonds through open market operations
- (1 mark)*
- 10** State the objective of the use of quantitative easing by the authorities in the UK.
- (3 marks)*

The regulation of the financial system

Key concepts from Year 1

There are no relevant concepts from Year 1 as this is an entirely new topic.

This chapter examines the roles played by four organisations in the regulation of the UK's financial system: the Bank of England, the Prudential Regulation Authority, the Financial Policy Committee and the Financial Conduct Authority. It investigates the reasons why a bank may fail, focusing in particular on the risks associated with borrowing short term and lending long term, and how liquidity and capital ratios can affect the stability of a financial institution. It also explains the meaning and implications for the financial sector of the term 'moral hazard'. Finally, it considers the nature of systemic risk and how problems in the UK's financial markets can impact upon the real economy.

Organisations that regulate the financial system in the UK

The financial crisis suffered by the UK during 2007–08 made the UK authorities aware of the need for a review of the regulation of the UK's financial system. This was conducted by Lord Turner, who was Chairman of the Financial Services Authority until its demise in 2013. He led the Turner Review (2009) into the crisis and made a range of recommendations to strengthen banking regulation and to separate the activities of investment banks from the deposit taking and lending activities of commercial banks.

The Financial Services Act of 2012 established a new regulatory structure to oversee the financial system and the provision of financial services within the UK. The new system is designed to overcome some of the weaknesses that contributed to the financial crisis in the UK. It became effective from April 2013.

A number of organisations play important roles in regulating the UK's financial system.

1 The Bank of England

The Bank of England's mission is to maintain financial stability in the UK, as we saw in Chapter 46. It carries out a number of roles, as stated on its website, including the following:

- reinforcing trust and confidence in the money that is used in the UK;
- acting as lender of last resort during periods of financial difficulty;

- promoting the safety and soundness of individual financial institutions (through the Prudential Regulation Authority);
- removing or reducing risks to the whole UK financial system (using the Financial Policy Committee);
- ensuring the stability of payment systems in the UK;
- supporting failing UK financial institutions (known as resolution).

Key term

Resolution is the process through which the Bank of England intervenes to manage the failure of a financial institution.

The responsibilities of the Bank of England in terms of resolution mean that if a financial institution in the UK fails, it will take action to minimise any adverse effects on the wider economy – sometimes termed the ‘real economy’.

Payment systems in the UK enable the lending and repayment of money, allow businesses to receive payments for goods and services, and facilitate the payment of salaries and benefits to the general public. The Bank of England’s role in maintaining the smooth functioning of these systems is essential to effective working of the wider economy.

The Financial Services Act of 2012 established The Prudential Regulation Authority as a part of the Bank of England as well as the Financial Policy Committee.

2 The Prudential Regulation Authority

The Prudential Regulation Authority (PRA) was established as a part of the Bank of England with the remit of ensuring the prudential regulation and supervision of around 1700 banks, building societies, insurance companies and other institutions providing financial services.

The Board of the PRA is its decision-making body. It is chaired by the Governor of the Bank of England (a post held by Mark Carney since July 2013) and includes five other members of the Bank’s staff. However, the majority of those sitting on the Board are independent of the Bank and all appointed because of their expertise and experience in financial services.

The PRA was established with the general objective of promoting ‘the safety and soundness’ of those firms it regulates. It has two other objectives:

- to facilitate competition in the financial services industry;
- to assist in securing an appropriate degree of protection for the policyholders (customers) of the UK’s insurance industry.

The PRA does not expect that its regulation and supervision will prevent all failures amongst financial firms. However, it will aim to make sure that the failure of any financial firm has a minimal effect on the supply of financial services to the wider economy.

Key terms

The **Prudential Regulation Authority** is the organisation responsible for the regulation and supervision of financial institutions in the UK.

Systemic risk exists, in finance, when there is a risk to the whole market, system or even economy, as opposed to a risk linked to one organisation.

Macroprudential regulation operates to mitigate systemic risk (i.e. the risk to the entire financial system).

3 The Financial Policy Committee

The Financial Policy Committee (FPC) was created on an interim basis in 2011 before being established in a more permanent form from April 2013. The FPC has two objectives:

- its primary objective is to identify, monitor and take action to remove or reduce systemic risks, which is intended to protect and enhance the resilience of the UK's financial system;
- its secondary objective is to support the UK government's economic policy.

The FPC is modelled on the Monetary Policy Committee and examines the macroeconomic and financial matters that have the potential to damage the long-term growth prospects for the UK economy (Figure 47.1). It sets policy to meet the Bank of England's statutory objective of financial stability. The FPC is designed to operate to mitigate systemic risk (i.e. risk to the entire financial system). Thus as the macroprudential regulator, the FPC focuses on financial stability. It can impose broad policy changes on the Prudential Regulatory Authority, which will regulate individual banks and other financial institutions. The creation of the FPC with the task of macroprudential regulation arose following the financial crisis of 2007–08 which threatened the entire UK financial system.

The FPC has considerable powers to ensure financial stability. For example, it can require banks to hold additional capital if it thinks the economy is overheating and inflationary pressures are developing. Alternatively, it can place limits on the size of mortgages that financial institutions are able to offer in order to help control the rate of increase of house prices. It can also limit overall bank borrowing or control specific kinds of lending.

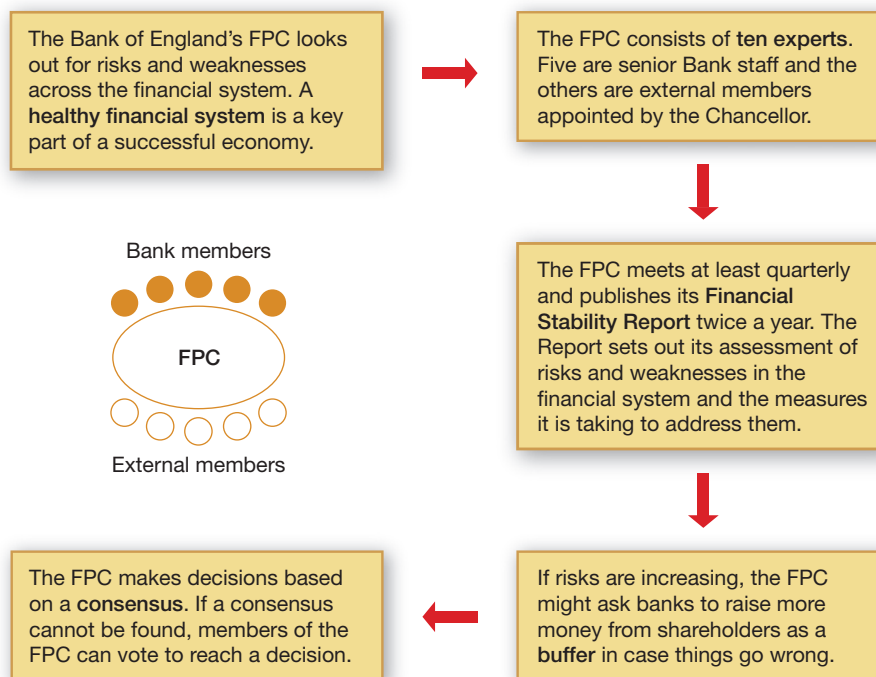


Figure 47.1 The Financial Policy Committee

Source: Bank of England

4 The Financial Conduct Authority

The Financial Conduct Authority (FCA) regulates firms providing financial services to consumers and maintains the integrity of the UK's financial markets. Its funding derives from fees charged to members of the financial services industry. The FCA website sets out its remit clearly: 'We aim to make sure that financial markets work well so that consumers get a fair deal. This means ensuring that: the financial industry is run with integrity; firms provide consumers with appropriate products and services; consumers can trust that firms have their best interests at heart. To do this we regulate the conduct of more than 70,000 businesses – and for many of these we also consider whether they meet prudential standards that reduce the potential harm to the industry and consumers if they fail'.

The FCA's role has several elements:

- **Authorisation:** Firms wishing to provide financial services in the UK have to be authorised by the FCA. If they do not meet the FCA's standards, they are not permitted to enter the financial services market.
- **Supervision:** The FCA supervises approximately 73,000 firms providing financial services in the UK with the aim of ensuring that their business models operate in the best interests of consumers. It also confirms that the businesses are financially sound. The PRA has prudential responsibility for all deposit takers (i.e. banks and building societies), insurance companies and large investment firms. Meanwhile, the FCA is the prudential supervisor for most other firms in the financial services industry, such as financial advisers as well as mortgage and insurance brokers.
- **Enforcement:** If the FCA uncovers poor business practice, it uses its power to make sure that firms providing financial services operate in the interests of consumers and do not damage confidence in the UK's financial markets.

Why a bank might fail

A number of factors can contribute to the potential failure of a commercial bank. Although these are presented as separate causes, they are interrelated.

1 Borrowing short and lending long

Commercial banks in the UK have always faced an unavoidable risk. This arises because they borrow short term and lend long term. They borrow short term because commercial banks promise their lenders (depositors and short-term money markets) that they can have their money back whenever they like (i.e. at short notice).

The risk is generated when they lend these funds long term, meaning that borrowers do not have to repay their loans for a considerable period of time. This can lead to problems if those who have lent short term want their money back from the bank, or if the long-term borrowers become unable to pay back their loans. In either event, banks have to make sure that they have sufficient cash to cover any shortfall. If they do not have sufficient cash, a crisis can result.

The banking system depends on trust, particularly amongst those (depositors) who have lent short term to the bank. If they think the bank has sufficient cash, and that it has not engaged in risky trading activities, then large numbers of depositors are unlikely to demand their cash back. However, if depositors fear that the bank has lent money in a risky way, then the trust is lost and they will demand their cash back.

Commercial banks are less likely to fail if they do two things:

- hold plenty of cash reserves to cover difficult periods;
- do not lend long term in risky ways.

2 Trading with insufficient capital and liquidity

We saw in Chapter 45 that having sufficient capital is essential for a commercial bank's long-term survival. Capital is important because it shows a bank's ability to withstand losses, as this source of funds does not have to be repaid. Capital can help a commercial bank to mitigate losses incurred in its trading – for example, if the holders of unsecured loans default on repayment. If a commercial bank incurs losses and does not hold sufficient capital, it may not be able to repay its creditors.

Equally, a lack of liquidity can threaten the survival of commercial banks, especially as they tend to be engaged in borrowing short term and lending long term as outlined above. If a bank loses the confidence of its depositors, it can lead to a crisis and a 'run on the bank', as happened to Northern Rock in 2007.

We consider these issues further when we discuss liquidity ratios and capital ratios later in this chapter.

3 Engaging in risky investment banking activities

In Chapter 45 we explained that some commercial banks operate risky investment banking divisions. These divisions can incur substantial losses, which can place the institution as well as the entire financial system at risk. The latter is termed systemic risk.

Higher returns on investments can only be achieved by taking higher risks. In the recent past, some UK banks created new financial securities as they sought to increase returns on their assets at a time of relatively low interest rates. As well as creating riskier financial products, they increasingly lent to one another. Thus many banks were holding the same risky financial products, increasing the risks to the banks as well as increasing systemic risk.

Many of these risky financial products were created through securitisation, which first became popular during the 1990s. Securitisation is the process of creating relatively liquid assets, that can be traded on financial markets, by packaging together assets such as mortgages or other types of loans, which can generate an income. These are called asset-backed securities.

However, these financial assets can be highly risky, in part because of their complexity. Some of the assets that were used to back the financial securities could not be relied upon to generate a steady stream of income in the future. For example, in the run-up to the financial crisis, some banks provided mortgages to those on low incomes. These borrowers were unable to repay the loans (which included high rates of interest) and assets based on these mortgages were described as being 'toxic'.

The financial markets in which these securitised products were traded failed, partly because of asymmetric information. It was common for the seller of the financial products (which had packaged up the security) to have more information than the buyer. The buyer was thus unable to assess the degree of risk accurately. If banks purchase this type of financial product, they increase their risk of failure because they are unable to repay depositors their money on demand if these loans become toxic.

Key terms

Capital is the finance raised through issuing shares and retaining earnings from previous trading periods.

Liquidity measures the ease with which assets can be turned into cash.

Key terms

Securitisation is the creation and sale of financial products, such as bonds, backed by the income generated by an asset, such as mortgages or car loans.

Asymmetric information exists whenever either the seller or buyer of a product has more information than the other party to the transaction.

Liquidity and capital ratios

Holding sufficient liquid assets and an appropriate amount of capital are both important indications of the vulnerability of banks to periods of economic difficulty.

Measuring liquidity

Liquidity measures the ease with which assets can be turned into cash. Banks hold some liquid assets, most obviously the cash in their branches, and some less liquid ones, such as long-term loans like mortgages.

We saw in Chapter 45 that a bank's balance sheet must contain sufficient liquid assets to ensure it can meet day-to-day needs for cash. Liquid assets are those that can be easily turned into cash, which is the most liquid of all assets. Liquidity ratios were agreed upon by the members of the Basel Committee on Banking Supervision in 2011 and will be implemented by March 2019. This committee comprised representatives from 27 countries including the UK, the USA, France, Germany, China, India, Russia and Brazil. It is commonly referred to as Basel III, as there were two earlier meetings. It introduced two liquidity ratios, to which commercial banks in the UK are subject, and also developed capital ratios, as we shall see later in this chapter.

The liquidity coverage ratio (LCR)

Commercial banks in the UK are obliged to hold sufficient cash reserves to meet the liquidity coverage ratio (LCR). The LCR requires that commercial banks hold sufficient high-quality liquid assets to exceed the net cash outflows over the next 30 days. This ratio is to be implemented fully by January 2019. This is essential to protect a bank against a situation in which a large number of its customers withdraw cash from their accounts at short notice. The objective of the LCR is to ensure that banks are resilient against the risk of liquidity crises. It does this by ensuring that banks have an adequate stock of available high-quality liquid assets that can be converted easily and immediately into cash to meet their liquidity needs for a 30 calendar day period of liquidity stress.

The net stable funding ratio (NSFR)

The net stable funding ratio (NSFR) is designed to address long-term structural liquidity problems by establishing a minimum amount of stable funding, based on the liquidity of bank assets over a one-year time period. It is defined as the net amount of stable funding available to banks over one year divided by the net amount of stable funding they require. The NSFR is intended to reduce the likelihood that disruptions to a bank's regular sources of funding will weaken its liquidity position in a way that could increase the risk of its failure. It is expected to become a minimum standard for UK commercial banks in January 2018.

Assessing capital reserves

Most monetary authorities, including the Prudential Regulation Authority in the UK, measure the reserves held by commercial banks to assess their resilience against shocks. This is sometimes termed 'capital adequacy'. Many of the measures of capital adequacy that are used were agreed between the monetary authorities of 27 nations as part of the Basel III accord in 2010–11. These were established in the light of the weaknesses in the commercial banks' holdings of capital during the financial crisis.

**REALWORLD
ECONOMICS** 47.1

Near-failure of RBS due to seven factors including poor management

Poor management decisions and the Labour government's light-touch regulatory regime were key factors in the near-collapse of Royal Bank of Scotland (RBS), a long-awaited report has revealed.

The report highlighted deficiencies in the management, governance and culture at RBS



London headquarters of the Royal Bank of Scotland – formerly offices for the Dutch bank ABN AMRO

and stated that the deal which effectively broke the bank – the £50 billion takeover of Dutch bank ABN AMRO – was carried out following inadequate research.

The report also highlighted the regulatory shortcomings in the lead-up to the bank's collapse, saying that the management of RBS was not challenged over its risky strategic decisions. It identified six key factors in the failure of RBS – most significantly, its weak capital position and overreliance on risky short-term funding in money markets.

In terms of the takeover of ABN Amro, the report concluded that RBS proceeded without appropriate heed to the risks involved and without sufficient research of the Dutch bank that it purchased in April 2007. Its research was said to amount to 'two lever-arch folders and a CD'.

The seventh key factor in explaining the bank's demise was the management, led by Chief Executive Sir Fred Goodwin. According to the report: 'The multiple poor decisions that RBS made suggest that there are likely to have been underlying deficiencies in RBS management, governance and culture which made it prone to make poor decisions'.

Source: Adapted from the *Independent*, 12.12.11

Exercises

Total: 14 marks

- 1 Do you think that strict regulation or good quality management is the most important factor in ensuring the survival of commercial banks? (7 marks)
- 2 Explain why a 'weak capital position' might put a commercial bank at risk of failure. (7 marks)

The Basel III agreement resulted in a number of different capital ratios. The most important of these relate to 'core tier one' capital. This is regarded as the 'best' form of bank capital and of most value to financial regulators when assessing a commercial bank's resilience to crises. It comprises the bank's capital raised from the sale of certain types of shares (often termed 'common equity') as well as its disclosed reserves of retained profit.

The adequacy of a commercial bank's capital is measured against its risk-weighted assets or RWA. Its total assets (loans, mortgages and financial securities) are recalculated to reflect the degrees of risk associated with them. The capital ratios required by the PRA in the UK express core tier one capital as a percentage of risk-weighted assets. This means that commercial banks that invest in risky assets such as derivatives have to hold a larger amount of core tier one capital to compensate for the risk involved.

The PRA has set out its requirements for UK banks in terms of core tier one capital ratios, as shown in Table 47.1.

Table 47.1 The PRA's core tier one capital ratios

Requirement	Core tier one ratio (%)
Baseline minimum requirement	4.5
Capital conservation buffer	2.5
Global systemically important bank	2.0
Counter-cyclical capital buffer	0.5*

Source: Bank of England

*(at the time of writing, April 2016)

In addition to the 'baseline minimum requirement', commercial banks are required to hold further core tier one capital for other reasons.

- **Capital conservation buffer:** This is a requirement for banks to hold further core tier one capital, equivalent to 2.5% of risk-weighted assets. The Basel agreement stated that this was necessary 'to withstand future periods of stress'. This, together with the baseline minimum requirement, means that UK commercial banks have to have a core tier one capital ratio of 7% from 2019 onwards.
- **Global systemically important bank (G-SIB):** The Financial Stability Board (FSB) has identified a list of the banks that fall into this category, which includes Barclays, HSBC, Lloyds, RBS and Standard Chartered from the UK. A G-SIB is a financial institution whose distress or failure, because of its size, complexity and systemic interconnectedness, would cause significant disruption to the global financial system and economic activity. G-SIBs in the UK are required to hold additional core tier one capital to guard against the international risk that can result from their failure.
- **A counter-cyclical buffer:** This permits national regulators, such as PRA in the UK, to require up to an additional 2.5% of core tier one capital during periods of high credit growth. This is most likely to be imposed during the recovery and peak stages of the economic cycle. Currently, the RPA has set this figure at zero for commercial banks in the UK.

A further tier one capital requirement, the systemic risk buffer (SRB), will be phased in by 2019. This is designed to protect the real economy against systemic risk resulting from the failure of a major commercial bank. In the United Kingdom, the SRB will be set by the PRA, applying a methodology to be determined by the Financial Policy Committee.

Author tip

You will not be required to calculate any liquidity or capital ratios for banks. You should understand their nature and how they help to maintain the financial stability of banks and other financial institutions.

Key terms

Core tier one capital

comprises a bank's common equity and its disclosed retained profit.

Retained profit, or retained

earnings, is the amount of profit kept by the bank and not paid out to shareholders in the form of dividends.

A global systemically

important bank (G-SIB) is a financial institution whose distress or failure would cause significant disruption to the international financial system and global economic activity.

Stress testing the UK banking system

In March 2013, the Financial Policy Committee (FPC) recommended that, ‘looking to 2014 and beyond’, the Bank and the Prudential Regulation Authority (PRA) should develop proposals for regular stress testing of the UK banking system. These tests would assess the system’s capital adequacy.

A stress test examines the potential impact of an imaginary adverse scenario on the entire banking system as well as on the individual institutions of which it comprises. These stress tests inform regulators of the extent of banks’ resilience to adverse shocks. They are designed to ensure that they are adequately capitalised and that they do not represent a risk to the wider or real economy if such shocks occur.

The PRA will operate a stress test of UK banks every year. The nature of the adverse scenario will vary annually, but it will be designed to reflect regulators’ assessment of the state of the financial cycle (i.e. they will be more severe if bank credit has expanded rapidly).

Moral hazard and systemic risk

The concept of moral hazard

In economics, moral hazard exists when one person or organisation takes greater risks because third parties carry the burden of those risks. It therefore exists when one party to an agreement takes actions which may affect the other party negatively. Moral hazard is more likely to arise when parties to an agreement have incomplete information about one another or where asymmetric information exists.

Moral hazard can occur in a range of situations. We consider two examples below.

- **Healthcare:** Consumers may engage in activities that pose a risk to their health, such as smoking tobacco, if the state provides free healthcare. The government or health authority bears the risk in this example.
- **The payment of sales teams:** If a salesperson is paid a salary rather than commission on each sale made, the risk exists that they may not try as hard to achieve sales as they may have done otherwise. The risk in this case is borne by the employer.

Key term

Moral hazard exists when one person or organisation takes greater risks because third parties carry the burden of those risks.

Moral hazard and the financial crisis

A better-known example of moral hazard is that of banks in the UK taking excessive commercial risks because they believe that the Bank of England or government will support them if they encounter financial difficulties. The financial crisis of 2007–08 is frequently quoted as an example of this. A major cause of the crisis was excessive risk taking by some commercial and investment banks in the expectation of financial support from governments if faced with the possibility of failure. They were protected by their belief that they were ‘too big to fail’.

The problem of banks taking excessive risks has resulted in much discussion and some important decisions being taken by regulators. The notion of being ‘too big to fail’ has received a lot of attention and has prompted a number of policy changes. Alan Greenspan, a former chairman of the Federal Reserve Bank (the USA’s central bank), has argued that if banks are ‘too big to fail’, then they are simply too big and should be broken up. In the UK the Bank of England has encouraged the development of ‘challenger banks’ – small new banks intended to provide competition to existing commercial banks.

The financial authorities in a number of countries have sought to separate commercial banks from riskier investment banks. We looked at this in detail in Chapter 45. Finally, some economists have proposed imposing taxes on large banks, with the level of the tax rising along with the banks' size. The intention is to discourage them from becoming very large.

Moral hazard does not just exist in relation to banks. Very large insurance companies can also behave in risky ways in the belief that they will not be allowed to fail. Those companies with close links to the financial system will have to meet stringent capital requirements to protect against this risk.

The implications of systemic risk in financial markets

We saw in Chapter 45 that the activities of financial institutions in financial markets can pose systemic risk. There is a potential for a loss of confidence in the banking system following the failure of a single bank or other financial institution. The failure of one financial institution could lead to the collapse of other financial institutions as households and firms rush to withdraw funds. This threat means that banks and other financial institutions (insurance companies and building societies, for example) represent systemic risk – the collapse of one could bring down the whole banking system. Thus, they pose a threat to the real economy and not just to businesses trading on financial markets.

Key term

The **real economy** refers to that part of the economy concerned with producing goods and services.

Figure 47.2 Financial institutions and the real economy

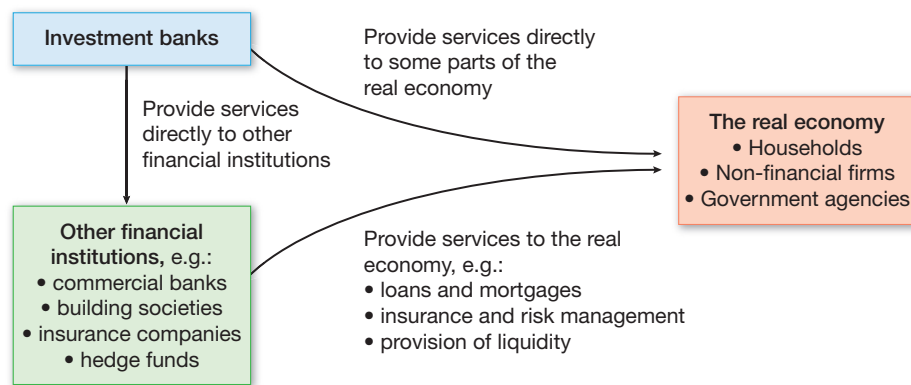


Figure 47.2 shows that all financial institutions have close links with the real economy and provide credit and capital to firms, households and government agencies such as the National Health Service.

The emergence of problems in financial markets has the potential to threaten the wellbeing of the real economy. These threats come in a number of forms.

- **A shortage of liquidity:** The actual or threatened failure of one or more financial institutions is likely to result in insufficient liquidity being available to households and firms. Payment systems may not operate effectively. Firms may have to reduce production or, in more extreme circumstances, to cease trading. This has the potential to damage consumption, economic growth and employment levels.
- **A lack of capital:** An important part of the activities of many financial institutions is the provision of capital for investment purposes and for infrastructure projects. A shortage of capital will increase its cost and reduce investment. This will have a negative impact on current aggregate demand as well as future aggregate supply.

REALWORLD ECONOMICS 47.2

The results of the PRA's stress testing in 2015

The seven largest commercial banks in the UK have passed the 2015 stress test – the second one to be imposed on the banks since the financial crisis. Officials are testing the seven biggest banks in the country to see how they would cope with a fictionalised economic crash that starts in China, before spreading to other emerging markets and to the rest of the world.

The test was based on the capital position of the banks at the end of 2014. Barclays,

Lloyds, HSBC, Santander and Nationwide all passed comfortably. The Royal Bank of Scotland would have struggled under the circumstances imposed by the test. However, regulators were satisfied that the bank had strengthened its capital position sufficiently since December 2014 and would pass now. Standard Chartered, a UK bank that does most of its business overseas, especially in emerging markets such as India, is taking action

to further strengthen its capital position and the PRA felt that no additional action was necessary.

Source: Adapted from a variety of media sources

Exercises Total: 12 marks

- 1 Why did the PRA only test the UK's largest seven commercial banks? (5 marks)
- 2 Why might the UK monetary authorities have opted to publicise the results of their stress tests as widely as possible? (7 marks)

- **Government financial support:** If governments are required to provide 'bail outs' to support failing financial institutions, this increases their expenditure and can result in heavy budget deficits because tax revenues will probably fall simultaneously. This is likely to mean that governments find it difficult to maintain expenditure in other areas of the economy. The overall impact could be a further fall in aggregate demand and the level of real GDP of the economy. The UK has experienced many of these consequences since the financial crisis in 2007–08.

Review questions

Total: 25 marks

- 1 Which organisation in the UK has a primary objective to identify, monitor and take action to remove or reduce systemic risks?
 - A The Bank of England
 - B The Prudential Regulation Authority
 - C The Financial Policy Committee
 - D The Financial Conduct Authority(1 mark)
- 2 Which organisation in the UK regulates firms providing financial services to consumers and maintains the integrity of the UK's financial markets?
 - A The Bank of England
 - B The Prudential Regulation Authority
 - C The Financial Policy Committee
 - D The Financial Conduct Authority(1 mark)
- 3 Explain the relationship between systemic risk and macro-prudential regulation. (4 marks)

- 4** Which term describes the process through which the Bank of England intervenes to manage the failure of a financial institution?
A Resolution
B Macroprudential regulation
C Securitisation
D Moral hazard *(1 mark)*
- 5** Explain the difference between liquidity and capital. *(4 marks)*
- 6** Which of the following are components of core tier one capital:
(i) common equity; (ii) secured mortgages; (iii) disclosed retained earnings;
(iv) unsecured loans?
A (i) and (ii)
B (i) and (iii)
C (i) and (iii)
D (ii) and (iv) *(1 mark)*
- 7** At the time of writing, which of the following does not count currently towards the core tier one capital requirements of commercial banks?
A Baseline minimum requirement
B Capital conservation buffer
C Globally systematic important bank (G-SIB) buffer
D Counter-cyclical buffer *(1 mark)*
- 8** What is meant by a globally systematic important bank (G-SIB)? *(5 marks)*
- 9** Why does the PRA 'stress test' the UK's major commercial banks each year? *(5 marks)*
- 10** Define the term 'moral hazard'. *(2 marks)*

Topic 12 Exam-style questions

A-LEVEL PAPER 2

SECTION A Context: Financial markets and monetary policy

Extract A Bond prices

Not everyone understands bond yields and many would assume that a rise in yields is good news. However, bond yields move in the reverse direction to bond prices and so, if bond yields rise, it follows that their prices have fallen. Naturally, this is bad news for anyone who owns the bonds.

In fact, the yields on many bonds issued by governments have risen sharply recently. Since the middle of April 2015 the value of German bonds has fallen by 15%, with similar British and American government bond prices falling by 5% and 7% respectively over the same period.

UK government bonds are available with a range of coupons, reflecting the market interest rates when issued. Investors are paid twice yearly: an investor who holds £100,000 of 4% Treasury bonds 2016 will receive two coupon payments of £2000 each on 7 March and 7 September.

Source: Adapted from the *Daily Telegraph*, 11.5.15

Extract B Influences on the MPC

Expectations of the first change in interest rates since 2009 receded following a report from the Bank of England. The Bank warned that global growth had slowed, as had inflationary pressures. Economists think that this indicates that rates will not rise until late 2016 and possibly later.

In December 2015, the Monetary Policy Committee held UK interest rates at 0.5%.



Figure A UK CPI inflation rate, 2010–15

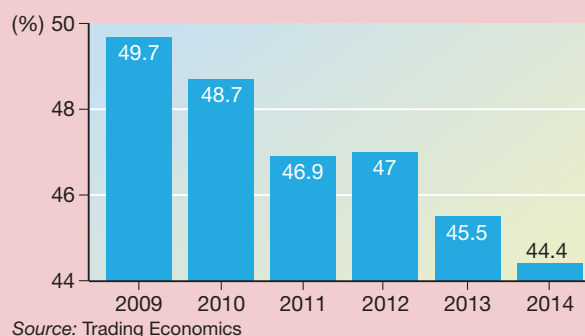


Figure B UK government spending as a percentage of GDP, 2009–14

The UK government has confirmed its commitment to reducing its spending. It is forecast to continue to fall steadily from its 2014 level to about 37% of GDP by 2019–20.

Extract C **More capital required**

Commercial banks in the UK including HSBC and Barclays have almost completed their task of building up their core tier one capital resources against a future financial crisis, according to Bank of England Governor Mark Carney. Mark Carney announced details of new capital requirements for extra capital for a counter-cyclical capital buffer (CCB) of 1% of risk-weighted assets. This is equivalent to £10 billion across the UK banking system. However, this is a relatively small sum compared to the many billions of capital already raised by the UK's commercial banks.

The Bank of England and its regulatory arm, the Prudential Regulation Authority (PRA), have also published new plans to separate the investment and high street banking operations of the UK's commercial banks. It is estimated that the UK's commercial banks will have to hold up to £3.3 billion of additional capital between them when they are forced to ring fence their high street banking from riskier investment banking activities in 2019. However, there has been some reduction in the degree to which the two operations are to be separate. Senior bank executives were having 'weekly behind-the-scenes talks with the Bank of England' on the possibility of modifying the proposed new rules.

Sources: Various media

Questions

Total: 40 marks

- 1 Assume that the value of the £100,000 bond referred to in Extract A fell by 7% to £93,000. Calculate the yield on this bond at this new price. (2 marks)
- 2 Use the information in Extract A to explain why there is an inverse relationship between market interest rates and bond prices. (4 marks)
- 3 Use the information in Extract B to analyse why an immediate rise in bank rate in the UK appeared to be unlikely in the late autumn of 2015. (9 marks)
- 4 Using the data in the extracts and your economic knowledge, evaluate the view that the requirement for UK commercial banks to hold high levels of core tier one capital is the reason why a UK bank is unlikely to fail over the next few years. (25 marks)

SECTION B ESSAYS

Total: 40 marks

The Governor of the Bank of England, Mark Carney, has said that the UK's commercial banks have nearly completed the task of building up financial reserves to protect against another financial crisis. UK banks should hold core tier one capital of 13.5% of risk-weighted assets by 2019, up from the current figure of 13.0%.

The Governor is also believed to be in favour of holding bank rate in the UK at its record low level of 0.5% until at least late 2016.

- 1 Explain how the imposition of requirements on commercial banks to hold higher liquidity ratios reduces the risk of financial instability. (15 marks)
- 2 Discuss the view that the major way in which the Bank of England's activities affect the UK's real economy currently is through the setting of bank rate by the Monetary Policy Committee. (25 marks)



Topic 13

Fiscal policy & supply-side policies

Fiscal policy

Key concepts from Year 1

We introduced fiscal policy in Chapter 44 of the Year 1 companion textbook. In this chapter we explained the components of fiscal policy: government expenditure and taxation along with the budget balance. We discussed how fiscal policy can be used to attain macroeconomic policy goals through influencing aggregate demand and aggregate supply. We also explored the use of fiscal policy to achieve microeconomic policy objectives such as the avoidance or correction of market failure and the redistribution of income and wealth.

This chapter builds on the coverage of fiscal policy at AS level to consider the types of government expenditure and the reasons why governments engage in expenditure and levy taxes. We examine the principles of taxation and the role and merits of different UK taxes. In addition, we investigate cyclical and structural budget deficits and surpluses, and discuss the consequences of these for macroeconomic performance. Finally, we consider the significance of the size of the national debt as well as the roles of the Office for Budget Responsibility (OBR).

Fiscal policy

Fiscal policy comprises three elements: government expenditure, taxation and the budget balance. Unlike monetary policy, it is the responsibility of the government and decisions are taken by the Chancellor of the Exchequer, other ministers and officials at the Treasury. Fiscal policy has not been used by governments in the UK since 2010 to manage the level of demand in the economy. The priority on decisions relating to government expenditure and taxation has been to reduce the size of the government's budget deficit to levels that are considered more manageable.

Author tip

It may be worth rereading Chapter 44 of the Year 1 book to familiarise yourself once more with the material that we covered for the AS examinations.

Key terms

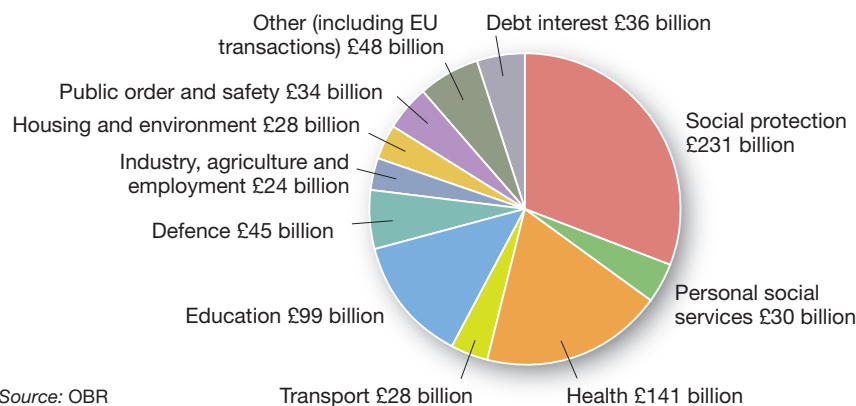
Fiscal policy refers to the government's manipulation of its expenditure, taxation and the budget balance to manage the economy.

Government (or public) expenditure describes spending by the government or its agencies on the provision of goods or services and spending on cash benefits.

Infrastructure refers to the basic facilities available to society that support economic activity such as transport and communication links.

Government expenditure

Government expenditure is sometimes referred to as public expenditure. It is spending by the central government (at Westminster) as well as local governments across the UK and the remaining public corporations such as the BBC. Local governments in the UK engage in spending to provide services such as social care and policing as well as transport, refuse collection and waste disposal services. Central government finances areas such as defence and healthcare. Figure 48.1 shows government spending in the UK for the fiscal year 2015–16 according to the main functions. Total expenditure is expected to be around £742 billion in 2015–16.



Source: OBR

Source: OBR

Government expenditure can be classified into three types: consumption expenditure, investment expenditure and transfer payments.

Government consumption expenditure

This type of government expenditure in the UK is also referred to as current expenditure. General government consumption expenditure includes all government current expenditure on purchasing goods and services, such as medicines for the NHS. It includes the payment of public sector employees, such as the wages of civil servants and lecturers in further education colleges. It also includes most forms of expenditure on the UK's defence.

Government consumption expenditure reached £354 billion in 2014. It rose steadily from the early 1990s until 2008 when it fell back as the UK government sought to reduce its spending and the size of the budget deficit. Since 2010 it has risen slowly in money terms, but fallen as a percentage of GDP.

Government investment expenditure

This is capital expenditure by the government – for example, on the construction of roads, bridges and hospitals. It involves investment in assets that will create economic growth in the future. Expenditure on the UK's infrastructure is one element of investment expenditure by the government. The government has argued that a higher level of investment is needed if the economy is to have the infrastructure that will enable the UK to be internationally competitive. However, it has proposed that 69% of investment in infrastructure in the UK should be funded by the private sector.

Figure 48.1 Government expenditure 2015–16 by main function

The Organisation for Economic Cooperation and Development (OECD) is an international economic organisation. It has researched the level of investment expenditure by governments in its member states over time. In 2013, on average, OECD governments invested 7.8% of total government spending, varying from 17.2% in Korea to less than 4% in Ireland. Investment expenditure by the UK government is towards the lower end of the range invested by member states: it was about 7% of GDP in 2014.

Transfer payments

One type of government expenditure that has received a lot of attention in the media recently is transfer payments. This form of expenditure simply represents transfers of money from one group to another. Transfer payments normally take the form of welfare payments, examples of which include pensions, universal credit and tax credits. Figure 48.1 shows that the UK government spent £231 billion on social protection in 2015–16. The majority of this would have taken the form of transfer payments.

This is an area of expenditure that governments in the UK have found difficult to control. Cuts in transfer payments can be politically sensitive. In 2015 the Conservative government faced considerable opposition to its plans to cut government spending on tax credits (a means of supplementing the incomes of low earners) and was eventually forced to abandon its proposals.

The reasons for government expenditure

Governments in the UK, as well as in other countries, undertake expenditure for a variety of reasons. To some extent, most of the reasons set out below could be argued to be correcting different forms of market failure.

1 To redistribute income

Most governments use their spending (primarily in the form of transfer payments), in conjunction with progressive taxes, to redistribute income from high income households to those in receipt of lower incomes. This is a means of reducing the level of poverty experienced by certain groups in society. Governments can use their spending to target particular social issues such as reducing poverty amongst households that contain one or more children.

Redistributing income using progressive taxes and transfer payments can have the effect of increasing the level of consumption and aggregate demand in an economy because lower-income households are likely to have higher marginal propensities to consume than higher-income households.

2 To provide essential services

A major reason for government expenditure is to provide essential services such as healthcare and education. These are key to providing the citizens of an economy with a decent standard of living as well as to increasing the productivity of the economy. They are merit goods that would otherwise be underproduced and underconsumed in a free market. The government also supplies public goods such as defence, which could not be provided otherwise because they are non-excludable and non-rival.

Key term

Progressive taxes take a higher proportion of taxpayers' incomes as their incomes increase.

3 To overcome other forms of market failure

Government spending on training and in supporting the establishment of start-up businesses can help to overcome other causes of market failure, such as geographical and occupational immobility. This can assist in reducing levels of unemployment and especially regional unemployment.

Taxation and why governments levy taxes

Almost every consumer and business in the UK pays taxes in one form or another. These are financial levies or payments imposed on a variety of economic activities. The UK government imposes direct taxes such as income tax on income and wealth as well as indirect taxes on spending such as value added tax (VAT). The UK and other governments levy such taxes for a variety of reasons.

1 To assist the government in achieving its macroeconomic objectives

The UK government levies taxes to assist it in controlling inflation, achieving a sustainable rate of economic growth and keeping unemployment at a minimal rate. Without levying taxes to fund the majority of its expenditure, its budget deficit would grow, possibly fuelling inflation as aggregate demand would rise too quickly. Levying taxes also enables the government to borrow money at a reasonable rate of interest and to maintain confidence in financial markets. This is essential in order to maintain a steady rate of economic growth.

2 To finance the government's expenditure

We saw in the previous section that the UK government is committed to spending in a number of different areas, such as defence, health and education. Such expenditure is essential to maintain the living standards of the economy's households as well as to maintain productivity through providing an efficient and competitive infrastructure and well-educated employees. Taxation provides the bulk of the finance for this expenditure: forecast figures suggest that it will provide £672 billion in 2015–16. Thus just over 90% of the UK's government's expenditure will come from taxation, with the remainder from borrowing. The expenditure plans show that the government intends to raise more from taxation than it spends by 2020–21. It intends to achieve this aim by reducing its expenditure through its programme of austerity.

3 To repay government debt

At the end of the 2015–16 fiscal year, the UK's national debt was in excess of £1500 billion. The current government plans both to cut its expenditure and to raise sufficient funds through taxation to achieve a budget surplus. This will enable it not only to halt the steady growth in national debt but also to reduce it.

Key terms

Taxation is a payment that has to be made to the government or other authority by households, firms or other organisations.

Direct taxes are those levied on income or wealth such as income tax.

Indirect taxes are paid on spending by firms, households and other organisations. The major indirect tax in the UK is value added tax (VAT).

Austerity describes government policies to reduce expenditure and increase revenues from taxation during periods of budget deficits.

4 To redistribute income and wealth

The UK's taxation system provides a stream of revenue, much of which is provided by households with higher levels of income and wealth. This finance is used, in part, to allow the government to support those on lower incomes through a range of welfare payments such as universal credit and tax credits. By transferring income from higher income and more wealthy households to less wealthy, lower income households, the government is able to achieve greater income equality in society.

The principles of taxation

Adam Smith set out the characteristics of a 'good' tax as early as 1776 in his influential book *The Wealth of Nations*. He termed these characteristics 'canons' of taxation and identified four canons on which the rules of a good system of taxation could be designed.

1 Equitable

Taxes should be fair and should be levied according to ability of taxpayers to pay. Taxpayers who earn the same incomes should pay the same amount of tax – a principle known as horizontal equity. Those receiving higher incomes should pay a higher proportion of their income as tax – known as vertical equity. Progressive taxes, which take a higher proportion of the incomes of those who receive higher pay, fulfil this particular canon or principle. Globalisation has made it more difficult for countries to impose equitable taxes on company profits as large multinational businesses can manage their affairs to generate the highest profits in countries with low rates of business taxation. Smaller and purely domestic businesses are unable to do this and may pay much higher levels of business taxation.

2 Economical

Taxes should be economical (i.e. cost effective) to collect. The cost of collection should be minimal relative to the revenue that is raised. The current Conservative government is seeking to reduce the cost of operating the organisation responsible for collecting taxes. It plans to make tax collection in the UK more cost effective by reorganising Her Majesty's Revenue and Customs (or HMRC) into fewer, larger administrative centres.

3 Convenient

In order to encourage the prompt payment of taxes, a tax system should be arranged so as to make it simple and convenient for taxpayers to make payments. Some indirect taxes, notably VAT, meet this canon by being included in the price of a product, and many taxpayers will not be aware that they are paying a tax.

Key terms

The **principles of taxation** set out the characteristics of a fair taxation system.

Hypothecation is a situation in which the revenue raised from a tax is ring-fenced for a specific purpose.

Globalisation is the trend for many markets to become worldwide in scope.

4 Certain

The operation of the tax should be clear to taxpayers and they should know how much they have to pay, the timing of the payment and the reason for the payment. For example, someone paying income tax should understand that the tax relates to their earnings, how the amount to be paid is calculated and the date by which it must be paid.

In addition to Adam Smith's four canons, other characteristics have been added, of which efficiency is an important example (as described below). Collectively these characteristics are now called the 'principles of taxation'.

5 Efficient

A tax should have minimal negative implications for other aspects of the economy. A poorly designed income tax, for example, may make it less worthwhile for those earning low incomes to continue working. In these circumstances, the level of voluntary unemployment may rise as the tax provides a disincentive to working for some employees.

The role and merits of the UK's major taxes

The UK's central and local governments impose a range of taxes on firms and households. These are illustrated in Figure 48.2, which also shows the forecast revenue from the individual taxes for the financial year 2015–16. Income tax, value added tax (VAT) and national insurance contribute the majority of the revenue from taxation in the UK. A summary of the role of each tax and their respective merits is provided in Table 48.1.

(We looked in detail at some of the UK's major taxes in Chapter 44 of the Year 1 companion textbook. You may wish to reread the relevant section of this chapter before studying the material below.)

Number crunching

Use the information in Figures 48.1 and 48.2 to calculate the government's budget balance for the 2015–16 financial year.

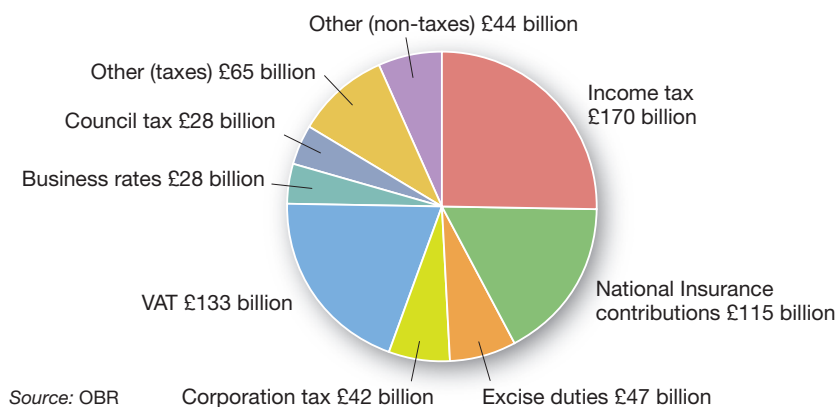


Figure 48.2 Revenue from taxation in the UK 2015–16

REALWORLD ECONOMICS 48.1

Netflix and UK corporation tax

Netflix, the online movie rental service, did not pay any UK corporation tax last year, despite generating an income of £200 million in the country. The world's biggest internet movie streaming service has five million subscribers in the UK alone and has achieved high sales by offering US shows like House of Cards and the BBC's Sherlock.

The company has been quick to defend its decision to avoid paying UK tax, stating that it is making 'overall losses' on its international ventures. It did, however, pay tax in Luxembourg

last year totalling £573,396.

Figures obtained by the *Sunday Times* showed that its Dutch arm – Netflix International BV – had a net turnover of £415 million and profits of £11.3 million. British customers, who typically pay £5.99 a month for the service, are given contracts from Netflix BV. This means that profits generated by Netflix International BV, which was Dutch based until 2014, included UK subscriptions.

Netflix had global revenues of \$5.5 billion and profits of \$267 million in 2014. The firm, founded in California in 1997, is expected to

have 100 million global subscribers by 2018 and is the most popular service of its kind in the UK. It said that it does expect to pay some UK corporation tax this year.

There have been calls for the UK government to invest more in HMRC, which collects taxes in the UK, to ensure businesses pay all the taxes for which they are liable.

Source: Adapted from the *Daily Mail*, 20.12.15

Exercise

- 1** To what extent has globalisation made it more difficult for the UK to levy 'good' taxes on businesses? (12 marks)

Income tax

Income tax is a progressive tax levied on individuals and some small businesses (primarily those that are not companies). It is paid on a range of incomes including those from employment and self-employment, the profits of small businesses as well as pensions, some welfare benefits and interest on savings.

Table 48.1 *The role and merits of the UK's major taxes*

Tax	Role	Key merits
Income tax	A direct tax on earnings and profits of individuals and small businesses.	<ul style="list-style-type: none"> ■ Progressive (to some extent). ■ Can improve distribution of income. ■ Steady inflow of revenue. ■ Simple for many to pay. ■ Revenue can be forecast.
Value added tax	The UK's major indirect tax and one required for all EU members.	<ul style="list-style-type: none"> ■ Can correct market failure. ■ Can influence spending patterns. ■ May promote consumption of environmentally friendly products. ■ Difficult to evade.
Corporation tax	A direct tax on company profits and a significant 'pull or push' factor for multinational businesses.	<ul style="list-style-type: none"> ■ Can be used to attract foreign direct investment. ■ Has the potential to help the government to meet its macroeconomic objectives.

Most people pay income tax each week or month from their earnings through a system known as Pay As You Earn (PAYE). This provides the government with a steady flow of income throughout the financial year, providing it with funds to spend on its regular and major commitments, such as social protection. Income tax can also be strongly progressive and form the basis of an effective method of redistributing

income to those who earn lower amounts. The UK's income tax system is not strongly progressive because the maximum rate that is levied is only 45%. However, there is a substantial tax-free allowance (£10,800 for most people in 2016–17), which gives very low or zero rates of tax for those on low incomes.

The existence of PAYE means that income tax is a relatively simple tax for many to pay, and HMRC plans to make more effective use of technology to simplify the process for most taxpayers. The government can forecast the likely effect on revenue raised from a change in the rate of income tax.

Value added tax (VAT)

VAT is an indirect tax levied on spending. Most goods and services sold in the UK have VAT at a rate of 20% added to their price. Some, such as car seats for children, have a 5% rate while other products, including most foods, are zero rated and VAT is not levied on them. This tax is imposed in all member states of the European Union, although rates vary.

VAT is a simple tax to pay for the final consumer because it is included in the purchase price of many goods and services. This also makes it a difficult tax to evade. It can be used to influence the pattern of demand by households for goods and services. Levying higher rates on environmentally unfriendly products and demerit goods may increase economic welfare, for example. As such, VAT can be used to overcome some forms of market failure.

VAT is also unlikely to act as a disincentive to working because it is levied on spending rather than on any form of earnings. Thus, a greater reliance on VAT rather than, say, income tax could be expected to encourage people into employment and help to generate supply-side improvements.

Corporation tax

Corporation tax is paid by companies in the UK on their profits; these profits can be from trading, investments or the sale of assets. The UK government is committed to creating a very competitive tax regime for companies in the expectation of attracting international businesses to the UK. In 2016, the main rate of corporation tax in the UK is 20% and it is set to reduce to 18% by April 2020.

Having a low rate of corporation tax encourages footloose businesses to relocate to the UK. In 2014 the UK received about 20% of all foreign direct investment entering Europe from the rest of the world, a much higher rate than any other European country. It is thus possible for a reduction in the rate of corporation tax to increase the amount of revenue received by the government.

The relationship between the rate of tax and the revenue received from that tax is shown in Figure 48.3. It is illustrated using a Laffer curve, which shows that there is an optimal rate of taxation ('t' in Figure 48.3) that delivers the highest tax revenue. The American economist, Arthur Betz Laffer, reputedly created the Laffer curve in a meeting in 1974. A central implication of the curve is that increasing tax rates beyond a certain level will reduce tax revenue. The underlying concept is often used in support of the implementation of supply-side policies.

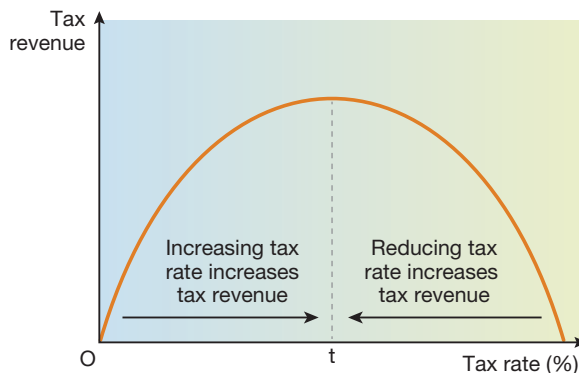
The process of globalisation has arguably made the revenue from corporation tax more sensitive to the rates levied. The UK's decision to reduce the rate of corporation

Discussion point

Do you think that the amount of revenue received is the main influence on the Chancellor of the Exchequer's decisions regarding the rate of corporation tax in the UK?

tax has resulted in increased investment, employment and production as well as the potential for higher receipts from this tax.

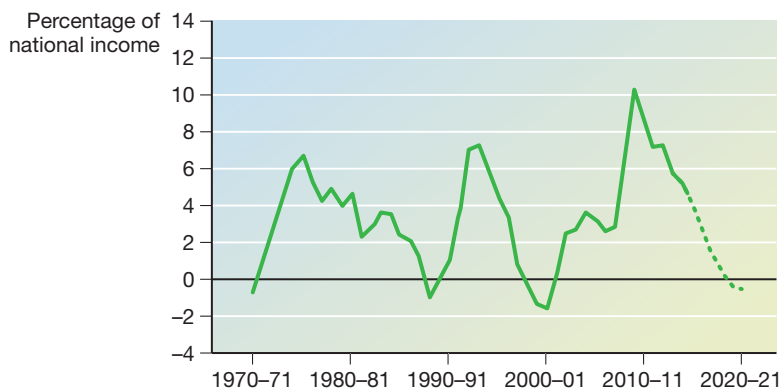
Figure 48.3 The Laffer curve



Budget deficits and surpluses

The budget balance is the difference between government spending and revenue over the financial year. Whether the government operates a budget deficit or a budget surplus (Figure 48.4) has implications for the performance of the economy.

Figure 48.4 The UK's budget deficits and surpluses as a percentage of national income, 1970–71 to 2020–21



Source: House of Commons Library, Briefing Paper 06167, 3.12.15

- Budget deficits:** These exist when government spending during the financial year exceeds the revenue received from taxation. This has been the situation in the UK for most years since 2000 and is forecast to continue until the 2020–21 financial year. A budget deficit can act as a stimulus to the economy as it is likely to stimulate aggregate demand. Furthermore, increases in government spending will initiate the multiplier effect.

Key terms

The **budget balance** is the difference between government spending and revenue over the financial year.

Budget deficits exist when government spending during the financial year exceeds the revenue received from taxation.

Budget surpluses occur when receipts from taxation exceed government expenditure over the financial year.

A **structural budget surplus or deficit** arises because of a fundamental imbalance between government receipts and expenditures, rather than short-term factors associated with the economic cycle.

- **Budget surpluses:** These result when receipts from taxation exceed government expenditure and they are less common in the UK. The UK last had a budget surplus in the 2001–02 financial year. Budget surpluses can have a deflationary effect on an economy because they reduce the level of aggregate demand.

Cyclical and structural deficits and surpluses

Budget surpluses and deficits can be associated with two influences: the state of the economic cycle as well fundamental structural factors such as a government's spending and taxation policies.

The size of a budget surplus or deficit is influenced by the stage of the economic cycle that a particular economy has reached. In a peak, for example, tax receipts will be relatively high while spending on welfare such as unemployment benefit will be at a comparatively low level. This reduces the level of borrowing by the government and is more likely to help to improve the budget balance. In a trough, borrowing tends to be high while receipts from tax revenues are reduced as profits and incomes fall. The government's budget balance will weaken at this stage of the economic cycle.

However, the stage of the economic cycle is not the only influence on the budget balance. An economy may experience a budget deficit or surplus irrespective of the stage of the economic cycle and this will be determined by underlying factors within an economy. A structural deficit or surplus is the imbalance that would persist in the long-term as the economy moves through all the stages of its economic cycle. The structural component of the budget balance is a good indication of the effectiveness of the government's financial management. This component indicates the underlying balance between long-term government revenues and expenditure, as it removes factors linked to the economic cycle.

Figure 48.4 reveals that the UK has experienced a structural deficit since around 2000–01 because the budget has been in deficit for approximately 15 years at the time of writing and is forecast to continue to be in deficit until the 2020–21 financial year.

The structural deficit is an elusive concept and cannot be directly measured. As a result, it has to be estimated. The UK's structural budget deficit is currently estimated to be approximately 90% of the overall level of borrowing in the 2015–16 financial year. Estimates of the size of a structural deficit are notoriously inaccurate, however.

Not all economists think that cyclical and structural deficits can be clearly separated because there are too many variables involved to allow a clear distinction to be made. Some economists argue that it is particularly difficult to make this distinction when analysing a government's current budget balance, as opposed to a historical one. It has been suggested by some economists that the distinction between cyclical and structural deficits and surpluses may be made mainly for political reasons.

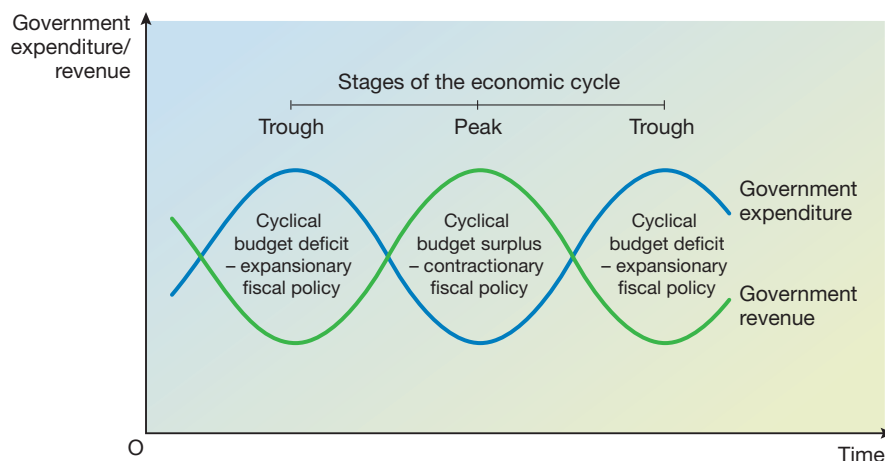
The consequences of budget deficits and surpluses for macroeconomic performance

The consequences of a budget deficit or surplus will depend upon whether it is cyclical or structural.

Cyclical deficits and surpluses

This type of budget surplus or deficit may not be regarded as a problem as, by its nature, it is likely to disappear over time. Many governments are content to operate a deficit in the recession and trough stages of the economic cycle, if they are followed by equivalent surpluses during the upswing and peak stages of the cycle. In fact, this element of fiscal policy can act as an automatic stabiliser, helping to reduce the effects of the economic cycle on macroeconomic performance. These broad effects are illustrated in Figure 48.5, which shows how changes in the cyclical budget balance alone can affect the nature of the government’s fiscal policy.

Figure 48.5 A simplified illustration of the effects of cyclical budget deficits and surpluses



During a recession, for example, a government budget deficit will, in effect, act as an expansionary fiscal policy and boost aggregate demand, helping to offset the deflationary effects of the recession. Conversely, during the peak stage of the economic cycle, government expenditure may be lower than its revenue and create a contractionary fiscal policy, thereby helping to reduce aggregate demand and lessen the likelihood of inflationary pressures emerging.

In October 2015 the UK government passed legislation to enforce a rule that future governments should operate a budget surplus during ‘normal’ times of economic performance. In effect this means that future UK governments are obliged to raise more from revenue than they spend, unless the economy is in a recession or trough. This will impose a legal barrier to prevent future governments from spending more than they receive in tax revenue when the economy is growing, unless, of course, they change the law.

Key terms

Automatic stabilisers are elements of fiscal policy that occur independently as an economy moves through its economic cycle.

The **national debt** is the total of all past government borrowing that has never been repaid.

Crowding out theories suggest that high levels of activity by the public sector can reduce the level of private

sector activity. For example, heavy borrowing by a government will use up much of the available credit and increase the interest rates charged (to the private sector) for what remains.

Structural deficits and surpluses

Structural deficits and surpluses can be more serious for governments because they cannot be expected to correct themselves over time. A significant structural budget deficit will result in economic problems for any government in the long term. There are a number of reasons for this.

A government will finance a structural budget deficit by borrowing either domestically or from overseas. This is currently the case in the UK and USA. As a consequence, the level of government debt will increase, raising the level of the economy's national debt. This will result in a rise in the country's national debt relative to its GDP. Those lending funds to finance a structural deficit tend to keep a close eye on the debt-to-GDP ratio because, if it rises too much, the government may be unable to repay the debt. Alternatively, the government may allow the rate of inflation to rise in the economy to reduce the real value of this debt and the value of repayments to creditors – an equally unattractive outcome for those lending money. A high debt-to-GDP ratio may make it more difficult for a country to repay these debts and the rate of interest on government loans may rise. In extreme circumstances, with very high debt-to-GDP ratios, the government may be unable to borrow funds to finance its deficit. We will consider the significance of the size of a country's national debt more fully in the next section.

A further possible negative consequence of budget deficits is 'crowding out'. Rising levels of government borrowing lead to an increasing demand for borrowing and a shortage of funds. This will increase interest rates. Higher interest rates are likely to lead to lower levels of investment and to reduce the international price competitiveness of the economy through rising exchange rates.

Structural surpluses are less common nowadays, but some countries such as Kuwait do operate them. Income from oil has allowed Kuwait to record a budget surplus every year since 1995. The macroeconomic effects of a structural surplus are the reverse of the effects that have been described above: they can be contractionary, but they may help to keep interest rates low and are unlikely to create inflationary pressures.

The significance of the size of a country's national debt

An economy's national debt is the total of all past government borrowing that has never been repaid. Most economies operate with substantial national debts, and an element of government spending is the payment of interest on these accumulated loans.

The UK's national debt increased by £71.9 billion during the 2015–2016 financial year to reach a total of over £1.5 trillion, according to data released by the Office for National Statistics (ONS). Other major economies operate with large national debts: the USA's amounted to more than \$18 trillion in 2015 and Japan's was \$11 trillion in the same year. However, a better way to assess the size of a country's national debt is to express it as a percentage of GDP. This is the debt-to-GDP ratio that we encountered in the previous section. This ratio provides a clearer indication of the extent of the burden of an economy's national debt by comparing it to the size of the economy in question. The ratios for a selection of countries are shown in Table 48.2.

Table 48.2 Debt-to-GDP ratios for a selection of major economies (%)

Country	Current ratio	Highest recorded	Lowest recorded
United Kingdom	89.4	89.4	31.3
United States of America	103.0	122.0	31.7
Euro Area	92.1	92.1	64.9
China	41.1	41.1	20.0
Japan	230.0	230.0	50.6
Germany	74.7	80.3	55.6
Italy	132.3	132.3	90.5
India	66.1	84.3	65.8
Russia	17.9	99.0	7.9
Spain	97.7	97.7	16.6

Source: Trading Economics

It is possible to identify a number of implications for an economy of having a high debt-to-GDP ratio.

1 Crowding out

A high debt-to-GDP ratio draws money away from private investment in creating new productive capital over the long term. This results in the economy having a lower productive capacity, reducing the growth rate and income levels within the economy compared with what may have been the case. This occurs because the funds used to buy government securities are unavailable for private investment projects. The result is a smaller stock of capital, and therefore lower output and income, than would otherwise have been the case, *ceteris paribus*.

2 Lower growth rates

Following on from the problems associated with crowding out, a study by Manmohan Kumar and Jaejoon Woo of the International Monetary Fund (IMF) illustrated that once countries reach higher debt-to-GDP ratios, their rates of economic growth slow. The work carried out by Kumar and Woo showed that advanced economies with high debt-to-GDP ratios grew 1.3 percentage points more slowly each year than countries with debt-to-GDP ratios below 30%. Kumar and Woo argued that the adverse consequences of high debt-to-GDP ratios increase steadily as the ratio increases from 30% to 90%.

3 Expenditure on interest payments

Government expenditure on interest payments will inevitably rise along with the size of the national debt. This may require the government to increase taxes or to reduce its spending, or possibly both. Without these actions it may not be possible for the government to manage its budget and national debt. If interest payments take up an increasing proportion of government expenditure, it will be less able to fund other forms of expenditure. A high debt-to-GDP ratio would also limit the ability of the authorities to use fiscal policy to respond to unexpected recessions or financial crises. If the government is committed to high interest payments, it may be less able to take decisions in the best interests of its citizens, such as increasing spending on health or defence during periods of crisis.

REALWORLD ECONOMICS 48.2

IMF offers advice on the size of a country's national debt

Public debt in rich countries exploded between 2007 and 2012, rising from an average of 53% of GDP to nearly 80%. Some people think this is a problem and say that governments need to do their best to cut it. However, that view has been challenged in a new paper from the International Monetary Fund (IMF), which suggests that 'paying down the debt' (or, in the words of George Osborne, Britain's Chancellor of the Exchequer, 'fixing the roof while the sun is shining') is not the most sensible approach.

The IMF's economists reckon that if a government could choose between having high or low debt today, then, *ceteris paribus*, they would (and should) choose the latter. After all, when debt is high, governments have to impose unpleasant taxes to fund spending on debt interest payments. These taxes are a drag on the economy.

However, when a government is faced with a high debt load, is it better to impose austerity and pay it down, or take advantage of low interest rates to invest? The answer depends on the amount of 'fiscal space' a government enjoys. This concept refers to the distance between a government's debt-to-GDP ratio and an 'upper limit', calculated by Moody's, a

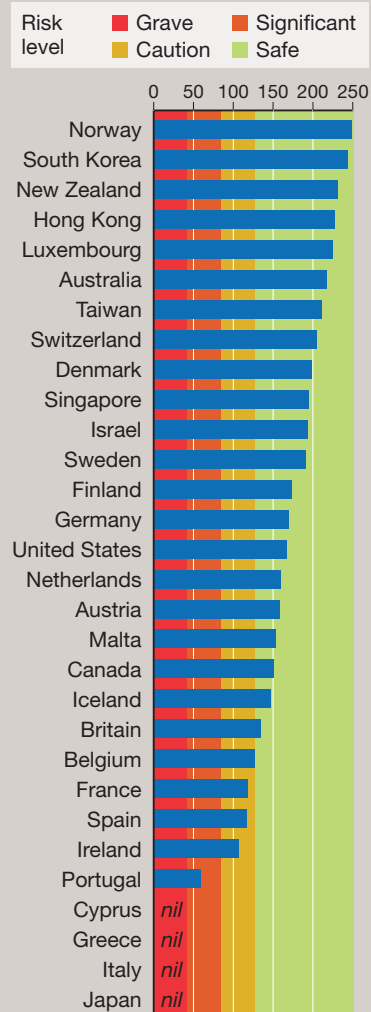
ratings agency, beyond which action would have to be taken to avoid default. Based on this measure, countries can be grouped into categories depending on how far their debt is from their upper threshold (Figure 48.6): safe (green), caution (yellow), significant risk (amber) and grave risk (red). It is a decent measure of how vulnerable a government's finances are to a shock.

For those countries with no headroom (in the red or amber zone on the chart), the IMF's paper is not much use: they need to take action to reduce their borrowing levels. For countries well into the green zone (of which America is a star performer and Britain is a somewhat marginal case), however, the IMF's analysis has a clear message: do not worry about your debt.

Source: *The Economist*, 3.6.15

Exercises Total: 14 marks

- Why might George Osborne (the UK's Chancellor of the Exchequer) have argued that the UK needed to reduce the size of its national debt? (6 marks)
- Why might a country with a large national debt, such as America, be relatively unaffected by it? (8 marks)



Source: Moody's Analytics

Figure 48.6 Fiscal space – distance to estimated limit in debt-to-GDP ratio, 2014 (percentage points)

4 Economic crises

High debt levels can precipitate a financial or economic crisis. An economy's national debt is normally financed through the issue of government bonds on its capital markets. If the institutions and individuals begin to doubt the ability of the government concerned to continue to finance the interest payments, then interest

rates on bonds will rise and reduce the government's ability to fund other areas of its expenditure. Ultimately the government may not be able to borrow at an affordable rate. In these circumstances, it has a choice: default on its debts (i.e. not repay them), seek help from organisations such as the IMF, or increase taxes and reduce spending dramatically and immediately. Whatever course of action is taken, the effects on the economy will be highly negative.

The roles of the Office for Budget Responsibility

The Office for Budget Responsibility (OBR) was established in 2010 to provide independent analysis of the UK's public finances. (Many other economies have equivalent organisations.) It engages in a range of research projects and produces publications to inform people about its work as well as providing briefings on public finance statistics to help people interpret the latest data in the light of its most recent forecasts. The OBR's website sets out its five main roles.

1 Economic forecasting

The OBR produces five-year forecasts for the economy and public finances twice a year. The forecasts accompany the Chancellor's Budget Statement (usually in March) and the Autumn Statement (usually in late November) and they incorporate the impact of any tax and spending measures announced by the Chancellor. The details of the forecasts are set out in the OBR's Economic and Fiscal Outlook publications.

2 Assessing the government's economic performance

The OBR uses its public finance forecasts to judge the government's performance against its fiscal targets. The government has set itself two medium-term fiscal targets: first, to balance the current budget; and, secondly, to have public sector net debt falling in 2015–16. In its Economic and Fiscal Outlook the OBR assesses whether the government has a greater than 50% probability of hitting these targets under current policy.

3 Scrutinising public finances

The OBR scrutinises the government's costing of tax and welfare spending measures. During the run-up to budgets and autumn statements, it subjects the government's draft costings of tax and spending measures to detailed challenge and scrutiny. It then announces whether it endorses the costings that the government finally publishes as estimates.

4 Sustainability

The OBR assesses the long-term sustainability of the public finances. Its annual Fiscal Sustainability Report (FSR) sets out long-term projections for different categories of spending, revenue and financial transactions, and assesses whether they imply a sustainable path for public sector debt.

5 Expenditure on welfare

The OBR judges the government's performance in terms of welfare spending against its welfare cap – a self-imposed limit on welfare spending. It publishes information on the trends in and drivers of welfare spending.

Review questions

Total: 34 marks

- 1 The payment of the state pension is an example of which of the following types of government expenditure?
 - A Consumption expenditure
 - B Investment expenditure
 - C Current expenditure
 - D Transfer payments

(1 mark)
- 2 Explain how the levying of progressive taxes can help a government to redistribute income between households in the economy.

(5 marks)
- 3 Define the term 'direct taxation'.

(3 marks)
- 4 Explain, with the aid of an example, how taxation can be used to correct market failure.

(5 marks)
- 5 Which of the following terms describes a situation in which the revenue raised from a tax is ring-fenced for a specific purpose?
 - A Austerity
 - B Hypothecation
 - C Resolution
 - D Hysteresis

(1 mark)
- 6 Explain why a cut in corporation tax might result in the government receiving more revenue from this tax.

(5 marks)
- 7 Explain the difference between a structural budget deficit and a cyclical budget deficit.

(5 marks)
- 8 Explain why governments may not regard a cyclical deficit or surplus as a problem.

(6 marks)
- 9 Define the term 'national debt'.

(2 marks)
- 10 Which of the following is the most likely consequence of an economy having a high debt-to-GDP ratio?
 - A Lower interest rates charged on bonds sold on capital markets
 - B Crowding out
 - C A more rapid rate of economic growth than would otherwise have occurred
 - D A structural surplus

(1 mark)

Supply-side policies

Key concepts from Year 1

We introduced supply-side policies in Chapter 45 of the Year 1 companion textbook. In that chapter we explained the difference between supply-side policies and supply-side improvements and the relationship that exists between the two. We examined how supply-side policies may increase the economy's potential output and its underlying trend rate of economic growth. We also considered how supply-side policies may affect macroeconomic objectives such as unemployment, inflation and the balance of payments. We concluded the chapter by assessing the strengths and weaknesses of supply-side policies.

This chapter builds on the material covered in Chapter 45 of the Year 1 companion textbook. It examines how and why governments use free market supply-side policies such as tax cuts and some labour market reforms as well as interventionist supply-side policies such as education and training. It also considers how supply-side policies can be used to reduce the natural rate of unemployment. Finally, we examine the ways in which supply-side policies can have microeconomic as well as macroeconomic effects.

Introduction

Supply-side policies focus on factors that influence aggregate supply and aim to increase the economy's productive potential. If successful, they can move an economy's production possibility boundary to the right or shift its long-run aggregate supply curve to the right. Governments in the UK gave greater emphasis to supply-side policies from the 1980s onwards in contrast to the dominance of Keynesian policies of demand management that were favoured in the period after 1945.

Supply-side policies are sometimes described as reducing the level of government intervention in an economy. We shall see that this can be the case, but they can also increase intervention to overcome limitations in the way markets work. Thus it is possible to divide supply-side policies into two categories: free market supply-side policies and interventionist supply-side policies (Figure 49.1). We shall consider each of these categories in turn.

Key terms

Supply-side policies are intended to increase an economy's productive potential by improving the efficiency with which markets operate.

Privatisation is the transfer of state-owned organisations to the private sector, where they are owned by individuals and private firms.

Deregulation is the reduction of the extent of state or government control over a business activity.

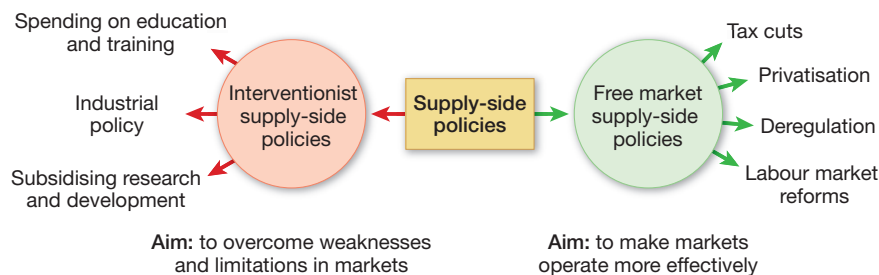


Figure 49.1 Different categories of supply-side policies

Free market supply-side policies

Free market supply-side policies are intended to make markets operate more efficiently. This approach to implementing supply-side policies can result in the use of a broad range of techniques, including those listed below.

1 Tax cuts

Economists who favour the use of supply-side policies believe that high rates of direct taxation (income tax, for example) can act as disincentives to households and firms and dissuade them from engaging in productive activity. Firms may not invest in increasing productive capacity if profits are subject to high rates of tax. Similarly, employees may not wish to be a part of the labour force if their earnings are subject to high taxation rates. Thus reducing rates of direct taxation can help to increase aggregate supply by making the labour market work more efficiently.

2 Privatisation

The supporters of privatisation argue that the removal of state control from an organisation and the introduction of the profit motive increase efficiency and competitiveness in a firm or in an entire industry. Thus privatisation can increase aggregate supply. This argument is more valid when privatisation creates genuine competition within a market, but less so when it simply turns a state monopoly into a privately owned one.

3 Deregulation

Deregulation entails a reduction in control over some business activity. Deregulation of markets can encourage new producers, increase competition and improve cost effectiveness, making the market operate more efficiently. Examples of markets that have been deregulated in the UK include bus and legal services, and plans are in place to deregulate the taxi and minicab industry.

The UK government passed the Deregulation Act in 2015, which was designed to reduce the legislative burden on businesses in order to assist in making product and labour markets operate more effectively.

4 Labour market reforms

Some reforms to the ways in which labour markets operate can improve the efficiency with which they work. A reduction in the power of trade unions, for example, can allow pay to vary between regions, rather than being settled nationally. This can help to reduce geographical labour immobility. Some areas such as London and the

Southeast have high housing costs and employers may need to offer higher wages to attract workers. If a wage rate is agreed nationally, firms may be unable to pay higher rates. As a result, some employees may not receive the market rate of pay and may be unable to move to better paid jobs in more expensive housing areas. The labour market works imperfectly because of the power of trade unions to negotiate national wage rates.

5 Achieving greater labour market flexibility

Some supply-side policies can make labour markets more flexible (i.e. allow firms to 'hire and fire' employees more freely in order to meet their production targets). This can be achieved by, for example, offering a lower degree of legal protection to those in employment. If firms can dismiss employees more easily at times when sales fall and they have excess productive capacity, then they will be more willing to hire when demand is rising. As a consequence the labour market responds more flexibly to changing conditions.

Interventionist supply-side policies

Interventionist supply-side policies are implemented with the aim of overcoming weaknesses and limitations in the way markets work. This category of supply-side policies can help to overcome market failure and can be used to encourage labour and product markets to operate in ways that assist the government in the achievement of its macroeconomic objectives.

1 Government spending on education and training

Expenditure by governments on education and training can create workforces that are more skilled and flexible in the face of changing demands from employers. This form of expenditure should improve the performance of the workforce in terms of labour productivity as well as in terms of creativity. Higher levels of productivity will increase the productive capacity of the economy directly and may attract multinational businesses, further increasing the potential output of the economy. If an economy has creative employees, they can develop new products and processes that help to enhance the competitiveness of firms and the economy as a whole.

Expenditure on education and training in the UK peaked at £96 billion (in 2013–14 prices) in 2010–11 and at 5.9% of GDP in 2009–10. Figure 49.2 shows that expenditure on training and education in the UK rose strongly in real terms until the financial crisis in 2007–08. However, the level of expenditure when expressed as a percentage of GDP has been relatively flat.

Key terms

Productivity measures the efficiency with which an economy uses its factors of production to produce goods and services.

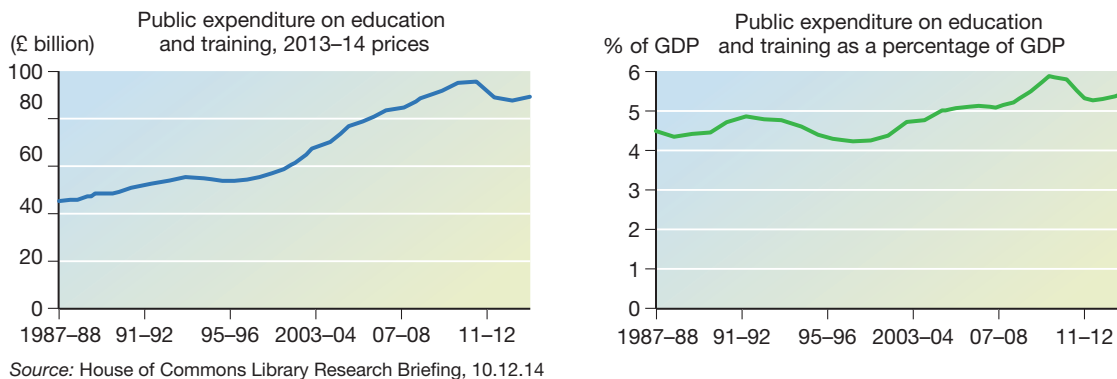
Industrial policy refers to government intervention that seeks to support or develop some industries to enhance economic growth.

Research and development (R&D) occurs when firms undertake projects that seek

to achieve an advance in science or technology.

Innovation is the process of turning an idea or invention into a saleable product or a more efficient method of production.

Regional policy is a series of government initiatives intended to raise employment and incomes in less prosperous parts of the UK.



2 Industrial policy

Industrial policy refers to government intervention that seeks to support or develop some industries with the aim of enhancing the rate of economic growth. In the 1970s industrial policy was based on state provision of financial support for industries identified as ‘future champions’. It was abandoned by Conservative governments in the 1980s and was not revived until after 2010 when the coalition government saw it as a way to support industrial sectors with significant potential for growth.

The current UK government has committed itself to supporting a range of UK industries including robotics, nanotechnology and life sciences. In order to achieve this aim, the UK government separates its industrial policy into two elements:

- **horizontal policies**, which address issues in many product markets. The government provides the resources and helps to create an economic environment that makes it easier for businesses and individuals to be productive. Examples of this type of policy include establishing tax regimes that favour business (e.g. lowering rates of corporation tax), taking actions to encourage innovation, and the development of more skilled workforces;
- **sectoral policies**, which include any policies directed towards one specific sector in the economy. Examples include subsidies to support the production of low carbon technology, providing funding for high technology industries or groups of firms, and support for research and development in particular industries.

3 Subsidising expenditure on research and development

Research and development (R&D) occurs when firms undertake projects that seek to achieve an advance in science or technology. This may be achieved by overcoming scientific or technological uncertainty or limitations – for example, the production of electric batteries that are light but hold substantial amounts of power. R&D projects include the improvement of existing products, processes or services, as well as devising new ones.

The R&D Expenditure Credit (the RDEC) was introduced by the UK government as a subsidy for suitable R&D expenditure in April 2013. In the UK there are currently two schemes in operation, one for small and medium-sized enterprises (SMEs) and one for larger businesses. Firms that invest in expenditure on qualifying R&D projects receive tax credits, reducing the amount of corporation tax or income tax that they have to pay.

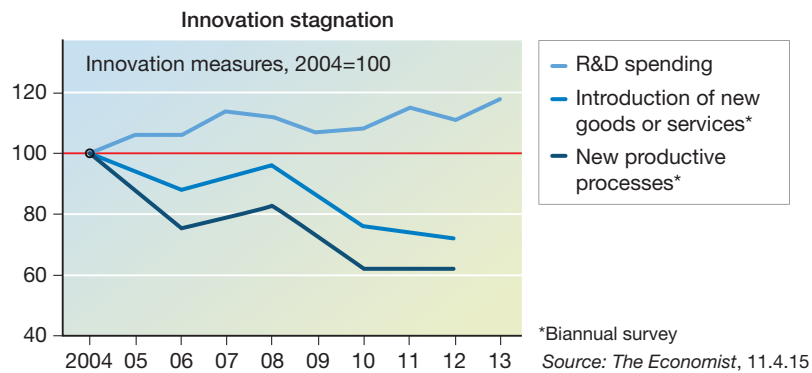
Figure 49.2 Public expenditure on education and training in the UK 1987–88 to 2013–14 in real terms and as a percentage of GDP

The UK government is committed to supporting innovation in the economy. Data from the Office for National Statistics show that R&D spending decreased to 1.72% of GDP (£27 billion) in 2013 from 1.77% the previous year. This is below the European average of 2.06% and far below the government's target of 2.5% by 2014. The R&D tax credit scheme is a vital part of the government's interventionist supply-side policy designed to stimulate investment R&D and to promote innovation.

The European Commission also supports expenditure on R&D in its member states. The Commission published a report in 2014 on R&D tax incentives within the Eurozone, showing that R&D tax incentives have a significant effect on business growth. It has a goal of increasing R&D spending to 3% of GDP by supporting member states with research.

The result of effective R&D expenditure is innovation. If expenditure on R&D is successful, firms are able to develop new and attractive products or more productive methods of producing existing goods and services. Innovation can therefore be a highly effective means of increasing an economy's competitiveness and productivity and stimulating economic growth. Figure 49.3 shows that spending on R&D in the UK has risen relatively slowly since 2007, though there has been a more significant increase since 2012. The decline in the introduction of new goods and services (and especially new productive processes) is apparent and a reason for the UK government increasing its financial support for R&D expenditure.

Figure 49.3 UK
 spending on R&D and
 the introduction of new
 products and processes,
 2004–13



Supply-side policies and the natural rate of unemployment

The natural rate of unemployment is the level of unemployment that remains when an economy's labour market is in equilibrium, as we saw in Chapter 41. It comprises two types of unemployment: structural unemployment and frictional unemployment.

Key terms

The **natural rate of unemployment** is the level of unemployment that exists when the labour market is in equilibrium.

Frictional unemployment exists when workers are in the process of moving to a new job.

Structural unemployment is the loss of jobs resulting from the long-term decline of specific industries.

Occupational immobility of labour exists when workers cannot transfer easily to employment in a different type of job.

Geographic immobility of labour occurs when workers cannot move freely to take employment in a new location.

REALWORLD ECONOMICS 49.1

A lack of skilled workers and managers drags the country down

The Dyson facilities in Wiltshire are dominated by a futuristic-looking 'campus' that is set to grow. Although production shifted to Singapore and Malaysia over a decade ago, all products are still designed and tested in Malmesbury, where the company employs around 1000 engineers. Now a vast new building is taking shape on the site, providing space for thousands more workers.

It should be a success story. However, as Sir James Dyson (pictured) explains, much will depend on whether he can find the 3000 engineers needed. The



James Dyson, vacuum cleaner entrepreneur

UK produces around 25,000 engineering graduates a year and so he could face a huge shortfall in recruitment, slowing his expansion plans. And Dyson is not alone. In the most recent survey of British firms that employ engineers and IT staff by the Institution of Engineering and Technology (IET), over half reported that they could not find the employees they were looking for and 59% said that the shortage would be 'a threat to their business in the UK'. Engineering UK, a lobby group, has warned that the UK currently has a shortfall every year of around 55,000 people with engineering skills.

According to the IET survey, the 'skills gap' has worsened for the ninth year in a row. It is a cause of the UK's terrible labour productivity. The UK has always lagged behind other countries in this respect, though before the financial crisis of 2008 it was catching up. Since

then, however, the productivity gap has widened again. In the long run it is increased productivity that will raise standards of living.

Economists have called Britain's productivity woes a puzzle, with poor infrastructure and stubbornly low investment cited as possible explanations. Most businessmen, however, are not puzzled at all: without skilled workers, it is hard for firms to grow, especially in the manufacturing and technology sectors.

Source: The Economist, 11.4.15

Exercises Total: 15 marks

- 1 Explain how the UK government might use supply-side policies to overcome the problems associated with the 'skills gap'. (7 marks)
- 2 Are free market or interventionist supply-side policies more likely to be effective in addressing the UK's skills gap? (8 marks)

We also saw in Chapter 41 that a number of factors determine the level of an economy's natural rate of unemployment. These include:

- geographical and occupational immobility of labour;
- inadequate labour market information for firms seeking workers as well as the unemployed;
- unemployed workers who do not possess suitable employment skills;
- insufficient expenditure on training by the government and firms;
- activities by trade unions that restrict the supply of labour into certain occupations.

Both free market and interventionist supply-side policies can be used to influence these factors and hence to reduce the natural rate of unemployment. Government expenditure on education and especially training can be used to improve the skills of a workforce. For example, it can reduce the degree of occupational immobility of labour by providing unemployed workers with relevant skills and knowledge. It can also help to overcome a lack of spending on training in the private sector. As a

consequence, the levels of structural and frictional unemployment can be expected to fall, reducing the natural rate of unemployment.

Supply-side policies can be used to address other aspects of frictional unemployment, most notably the lack of labour market information available to firms and unemployed workers. By providing information through job centres or websites, the government can match up unemployed workers to suitable vacancies as quickly as possible. This will reduce the duration of unemployment of those who are frictionally unemployed.

The microeconomic effects of supply-side policies

Supply-side policies have microeconomic effects and it is possible to argue that they are really microeconomic policies because they are implemented with the aim of improving the operation of product and labour markets. Policies such as privatisation and deregulation are intended to make markets work more efficiently with less government intervention. This can increase output and reduce average costs, enhancing efficiency. In this way, firms and industries in the UK can be expected to become more competitive and thus increase sales in domestic and global markets.

Supply-side policies can reduce the extent of market failure. For example, policies designed to reduce trade union power and influence over labour markets can assist labour markets in working more perfectly. Trade unions can restrict the supply of labour and can force up wages. As a consequence, the level of employment is lower than would otherwise have been the case, and wages higher. They can also create unemployment in labour markets by generating an excess supply of labour. Measures to reduce the power of trade unions can overcome at least some of these problems.

Review questions

Total: 26 marks

- 1 Define the term 'free market supply-side policies'. (3 marks)
- 2 Government policies that result in a reduction in the extent of state or government control over a business activity are best described as which of the following?

A Privatisation	C Industrial policy	
B Deregulation	D Regional policy	(1 mark)
- 3 Explain how privatisation can increase aggregate supply. (5 marks)
- 4 Explain how making labour markets in the UK more flexible might result in a rise in real GDP. (5 marks)
- 5 Distinguish between research and development (R&D) and innovation. (5 marks)
- 6 Which of the following types of unemployment is most likely to be associated with the natural rate of unemployment?

A Frictional unemployment	
B Cyclical unemployment	
C Involuntary unemployment	
D Seasonal unemployment	(1 mark)
- 7 Explain one way in which supply-side policies may be used to reduce the natural rate of unemployment. (6 marks)

Topic 13 Exam-style questions

A-LEVEL PAPER 2

SECTION A Context – deficit, debt and supply-side policies

Extract A Deficit and debt

The financial crisis and the recession led to a sharp increase in borrowing by the UK government. The annual budget deficit peaked at more than £150 billion, or 10% of national income or GDP in the 2009 financial year. It is forecast to be £75 billion for the 2015–16 financial year.

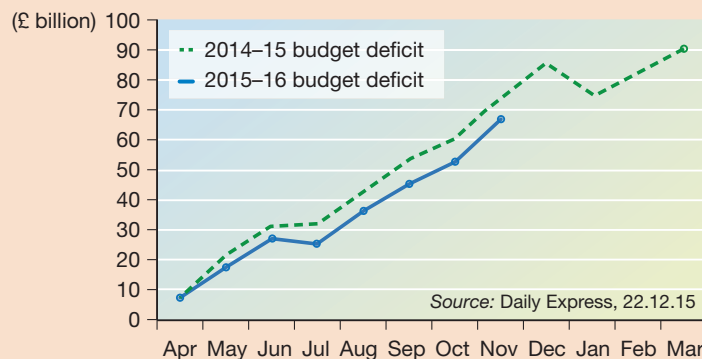


Figure A UK budget deficits 2014–15 and 2015–16, monthly cumulative (£ billion)

However, the total of accumulated debt (national debt) continues to rise, reaching £1.5 trillion by the end of the 2014–15 financial year. This amounts to nearly 90% of GDP.

Source: Adapted from BBC News, 27.10.15

Extract B Structural and cyclical deficits

A government can impose an austerity programme and still spend far more than it receives in the form of taxes. Indeed, the British coalition government had a deficit of 9.3% of GDP in the first year of austerity, a very high figure by peacetime standards. However, since this was less than the 11% of GDP in the year before, it counts as austerity.

What economists generally mean by austerity is a reduction in the structural deficit of the government (i.e. ignoring the effects of the economic cycle). Austerity usually involves a government trying to counteract the effect of the automatic stabilisers. Critics warn that this is counterproductive. The overall deficit could actually rise if a government tries to reduce the structural deficit at a time when the economy is weak; if the effect is to reduce GDP, tax revenues might fall even further while spending on benefits goes up.

It is not easy to determine how much of the deficit is structural and how much is cyclical. That requires economists to take a view on the trend rate of growth of an economy and on the level of spare capacity, particularly in terms of employment.

Source: Adapted from *The Economist*, 20.5.15

Extract C Supply-side policies in the UK

The UK is currently not at full employment. While the employment rate is very nearly at a historical high, it does appear that the labour market could sustain yet higher levels of employment, as there is little evidence of inflationary pressure.

Supply-side issues inform much of the UK government's proposed policy agenda on full employment. The coalition government implemented a range of supply-side measures over the years 2010–15. The list of policies can be summarised as cutting red tape, more apprenticeships providing training, investment in the UK's infrastructure, and low taxes. Chancellor George Osborne adds cutting benefits to this list.

Figure B UK
unemployment rate
2011–2015



Source: Adapted from an article on the Clarity website by Adam Tinson, 19.5.15

Infrastructure investment has a potentially much healthier return in terms of employment and growth, particularly if it is financed by borrowing. The IMF supports this approach. Tax cuts, on the other hand, may well be fine and good, but when accompanied by sizeable public spending cuts they are likely to find their impact blunted.

Questions

Total: 40 marks

- Using the information in Item A, calculate the percentage fall in the UK's budget deficit between 2014–15 and 2015–16, assuming the forecast for the 2015–16 financial year is correct. (2 marks)
- Use the information in Extract A to explain why reducing the UK's budget deficit does not reduce its national debt. (4 marks)
- Use the information in Extracts A and B to analyse why the UK government should be more concerned about its structural deficit than its cyclical deficit. (9 marks)
- Using the data in the extracts and your economic knowledge, evaluate the view that the UK government should implement further supply-side policies with the aim of achieving its objective of full employment. (25 marks)

SECTION B Essays

Total: 40 marks

The UK government has set out spending and taxation plans intended to eliminate its budget deficit by the 2020–21 financial year through a continuation of its policy of austerity. Despite this, some economists believe that certain tax rises are inevitable. The Institute for Fiscal Studies states that 'a one percentage point rise in all rates of income tax would raise £5.5 billion'.

- Explain the possible economic consequences of an increase in direct taxation in the UK. (15 marks)
- Evaluate the view that the UK government's budget deficit means that it must depend upon the use of supply-side policies to achieve its macroeconomic objectives. (25 marks)



Topic 14

The international economy

Globalisation

Key concepts from Year 1

There are no relevant concepts from Year 1 as this is an entirely new topic.

This chapter explains the major causes of globalisation and its main characteristics. It examines the consequences of globalisation for less developed and more developed economies. Finally, we consider the role of multinational corporations in globalisation.

What is globalisation?

The term ‘globalisation’ first emerged in the 1960s. Since that time it has been defined in a range of ways in different dictionaries. Collins says it is the ‘process enabling financial and investment markets to operate internationally, largely as a result of deregulation and improved communications’. Webster, an American dictionary, defines it more generally: ‘to make worldwide in scope or application’. Many economists would summarise it as the trend for many markets to become worldwide in scope.

From the 1980s onwards, financial markets became globalised as financial products were traded internationally because communications improved and made such trading possible. Capital and short-term financial products as well as insurance policies were sold in international rather than just national markets. Since that time, commodity markets (coffee and copper, for example) and final product markets (motor vehicles, mobile telephone services and consumer electronics, for example) have become increasingly global.

The globalisation of markets has created closer economic ties between countries as large firms (termed ‘multinational corporations’) have spread their production facilities across several countries.

Most economists would agree that economies throughout the world are becoming increasingly interconnected as a result of globalisation. It is possible to perceive that there is a widening and deepening of the connections that exist between people in different countries in all aspects of life. This includes cultural, sporting and fashion products, as well as concern for the environment and human rights. The world is

Key terms

Globalisation is the trend for many markets to become worldwide in scope.

Electronic retailing or e-tailing is the sale of goods and services through the internet.

The **World Trade Organisation** is an international body established with the aim of increasing world trade. It has 160 member countries.

being shaped by the forces of globalisation and driven by technology; more political and social experiences are being shared; and goods and services are becoming increasingly similar.

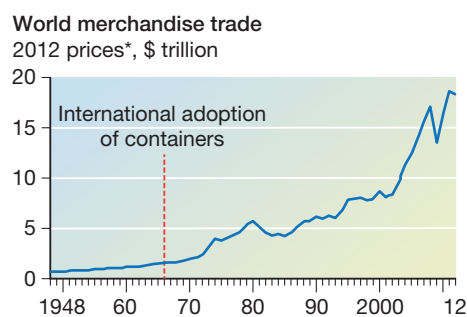
The causes of globalisation

Globalisation has its opponents, many of whom fear loss of traditional jobs and also of distinct cultures in particular countries and regions. However, there are a number of powerful forces behind globalisation.

1 The falling cost of international transport

In a globalised world, no economy is self-sufficient. Trade is essential to sell surplus products and to buy what is needed from other specialised producers. Improved transport links have made it easier and more cost effective to trade and thus to produce as well as to sell products in global markets.

Two trends in transport are particularly important. The fall in the cost of air travel in real terms has made it feasible to transport some products such as fruit and vegetables by air, enabling growers in Chile and Peru, for example, to sell products to wealthy consumers in North America and Europe. Simultaneously, the development of containerisation and much larger ships have brought down costs of sending products across the world by sea, dramatically increasing trade (Figure 50.1). When containers were first introduced they cost just \$0.16 per tonne to load, compared with \$5.83 per tonne for loose cargo. This reduction in costs has supported China's rise as the world's premier manufacturing economy.



Source: *The Economist*

Note: *Deflated by US consumer prices

Figure 50.1 *World trade and containerisation 1948–2012*

2 Developments in technology

Developments in communications technology have also made it quicker, simpler and less costly to send information around the world and to provide services. This has made it possible, for example, for many UK businesses to establish customer service facilities in countries such as India, where wage costs are significantly lower. The use of communications technology to send design and other information to other countries and to stage video conferences has encouraged firms to base production facilities in countries where costs are lower. Consequently production has become global rather than national.

The development of the internet has meant that consumers can evaluate products that are available from producers internationally before making a purchasing decision.

Even relatively small businesses are able to sell products to consumers worldwide through the use of electronic retailing or e-tailing.

3 The growth of global trading blocs and the reduction of barriers to trade

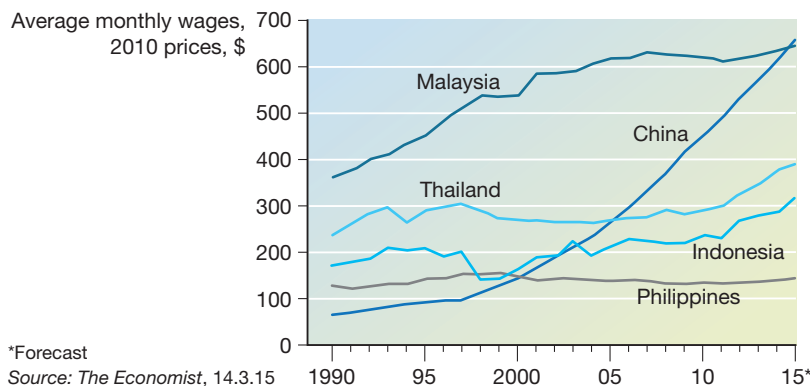
The European Union, one of the world's largest free trading blocs, has expanded to 28 countries with a total population of over 507 million in 2014. This allows many businesses to sell internationally without hindrance. Other international trading blocs have also been established and expanded including the North American Free Trade Association (NAFTA) and the Association of Southeast Asian Nations (ASEAN), with similar consequences for trade. At the same time, despite some setbacks, the World Trade Organisation has successfully reduced barriers to global trade.

Further major trade negotiations are taking place involving large numbers of developed and developing countries. For example, it may be that in 2016 the United States and the European Union agree a highly significant trade arrangement. The Transatlantic Trade and Investment Partnership (TTIP) is a free trade arrangement between the EU and the USA. If the negotiations are concluded successfully, TTIP is expected to promote higher rates of economic growth and economic interconnectedness. In October 2015, the USA concluded broadly similar trade negotiations with Japan, Canada, Australia and eight other Pacific countries – this is the Trans-Pacific Partnership or TPP. Among other things, the TPP agreement contains measures to lower barriers to trade, such as tariffs, with the aim of increasing rates of economic growth.

4 Increasing global incomes and growing demand for goods and services

Many millions of consumers across the world are enjoying rising incomes as a result of economic growth. For example, as illustrated in Figure 50.2, average incomes in China increased by approximately 650% between 1997 and 2015. Real wages in other developing Asian economies such as Thailand and Indonesia have risen sharply since 2010. This gives Chinese consumers, as well as those in other Asian economies, the ability to buy many products produced elsewhere in the world and thus fuels international trade and the process of globalisation.

Figure 50.2 Average monthly real wages for a selection of Asian economies, 1990–2015



The main characteristics of globalisation

It is possible to argue that the main characteristic of globalisation is that it promotes rapid and possibly increasing rates of change. It has extended social, political and

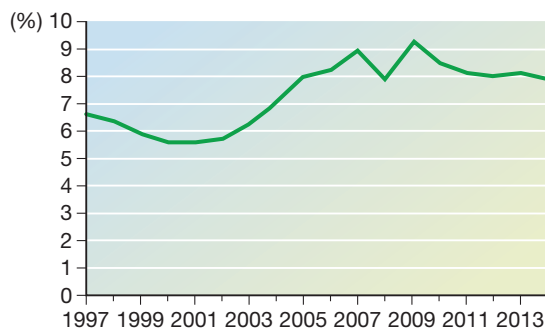
economic activities across national frontiers. Social forces, such as the desire to protect the environment, are global in scope, as they need to be. To some extent, the boundaries between national and global affairs are becoming blurred.

We can identify four main characteristics of globalisation. These overlap to varying extents with the consequences of globalisation for different types of economies that we consider in the next section.

1 Greater specialisation

Globalisation has resulted in firms and economies facing greater competition from overseas. This in turn has resulted in greater specialisation as economies have sought to focus on producing goods and services in which they have the greatest competitive advantage. For example, the finance sector (which includes banking and insurance) is very important to the UK economy, despite the difficulties associated with the financial crisis of 2007–08 and its aftermath.

Figure 50.3 shows that this sector contributes around 8% of the UK economy's gross value added (GVA). GVA measures the contribution of a sector to the economy, less the value of goods or services used in production. GVA therefore measures the value of a sector's goods or services to the economy.



Source: House of Commons Library Standard Note, 26.2.15

Figure 50.3 Percentage of the UK's gross value added (GVA) accounted for by financial and insurance services, 1997–2014

Other economies are noted for specialising in other products, industries or sectors. For example, the mining and quarrying sector is vital to the Australian economy and manufacturing is a core activity in China.

2 Increased international trade

Arguably the most prominent characteristic of globalisation is the increase in international trade that has accompanied it. Trade agreements and the reduction in barriers to trade have played important roles as most economies have increasingly engaged in trade. From the mid-1980s until about 2007, the value of international trade grew at an annual average rate of approximately 7%. Over the same period

Key terms

Gross value added (GVA) measures the contribution of a sector to the economy, less the value of goods or services used in production.

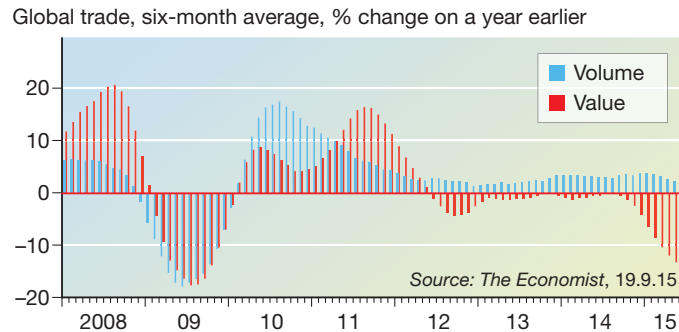
Specialisation occurs when an individual, firm, region or country concentrates on producing a limited range of products.

Multinational corporations (or companies) are businesses with productive facilities in more than a single country.

Economies of scale are the advantages that a firm gains due to an increase in its size or in the size of the industry in which it operates.

the ratio of the value of international trade in goods and services to GDP rose from under 20% to 31%. The latest figure for 2014 stood at 30%.

Figure 50.4 *The volume and value of global trade, 2008–15*



However, the financial crisis and the recession which followed saw a dramatic fall in the volume of trade, though it has grown at a broadly declining rate since 2010, as shown in Figure 50.4. In the first half of 2015 the volume of world trade grew by just 1.7%, below the long-term average of 5%.

The fall in the value of trade after 2011, which Figure 50.4 shows to have been most pronounced since the third quarter of 2014, has been due to some extent to falling prices of commodities that are traded globally. This has also suppressed the prices of final products that are traded which contain these commodities because this has held input costs and prices down.

3 Larger firms enjoying the benefits of economies of scale

Globalisation has increased the size of markets to which individual firms and especially multinational corporations can sell. As firms and economies have increasingly specialised, they have been able to reap the financial advantages of size (i.e. economies of scale). Some economists have argued that the minimum efficient scale (the level of output at which long-run average total costs are minimised) has increased over time. If this assertion is true, then targeting global markets (rather than just domestic ones) offers multinational corporations the opportunity to take full advantage of their scale.

Globalisation encourages the production of goods and services in different locations across the world, with each stage of production carried out in the location that offers the lowest costs. Thus the supply chains used by multinational corporations have become global and more complex as well as more efficient. The benefit of global production accrues to larger businesses in particular, allowing them to reduce their unit costs and selling prices or profit margins.

Dyson, a UK multinational corporation, manufactures vacuum cleaners, washing machines and other electrical products. It is a business that illustrates the nature of large-scale production in a globalised economy. The company was founded in 1993 and initially it conducted all of its research and production in Wiltshire in the UK. In 2002 it took the decision to manufacture its vacuum cleaners in Malaysia; later washing machine manufacture was moved to Malaysia as well. A factory was built in a joint venture between Dyson and the Singapore-based Meiban Group Ltd.

There were two key reasons for this move:

- manufacturing costs were reported to be 30% lower in Malaysia; and

- many of Dyson's suppliers were located in Asia and this reduced transport costs, especially when final products were sold to Asian consumers.

Research and development facilities remain primarily in Wiltshire, although the company has teams of engineers elsewhere, including Singapore and Chicago. Dyson illustrates the nature of globalised production: a complex global supply chain; various elements of production located in countries that offer the greatest benefits; a joint venture with another company; and sales in markets across the world.

4 Greater resource mobility

The level of interconnectedness brought about by globalisation has increased in other ways apart from trade. One fundamental characteristic of globalisation is that resources have become much more mobile. Capital and labour have both moved more freely around the world in pursuit of the highest returns as barriers to mobility have been reduced.

Migration has received much attention in the media over recent years. Much migration is due to 'push' factors such as war and disease. However, a further cause of migration is the lure of better living standards elsewhere. Globalisation has been accompanied by an unprecedented increase in human mobility. In 2013 there were an estimated 232 million international migrants.

The geography of migration flows is changing in line with changes in the global economy. Migrants are attracted to countries that are experiencing high rates of economic growth. These include developing countries such as India and Brazil and well as developed countries in North America and Europe. The process of globalisation has stimulated growth in many countries and created more diverse flows of migrants, as illustrated in Figure 50.5.



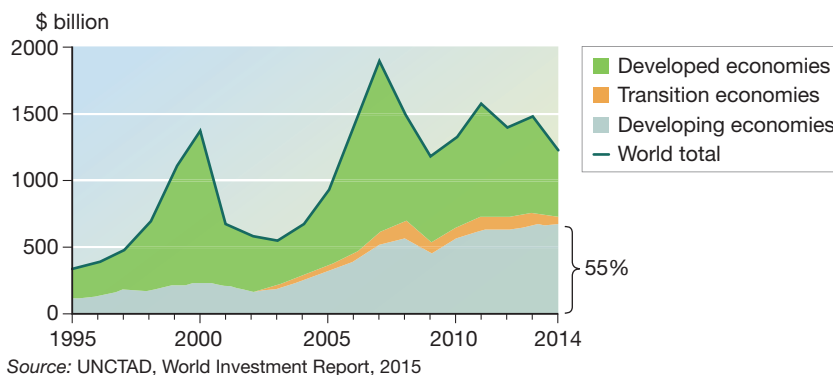
Figure 50.5 Global diversification of migration destinations

Source: International Organisation for Migration, World Migration Report, 2015

Increased capital flows between countries have been a feature of the process of globalisation, particularly foreign direct investment or FDI. FDI is defined by the Organisation for Economic Cooperation and Development (OECD) as 'cross-border investment by a resident entity in one economy with the objective of obtaining a lasting interest in an enterprise... in another economy'. Although FDI is very

sensitive to changes in the economic cycle, it has more than tripled since the mid-1990s, as illustrated in Figure 50.6. This graph also shows the growing importance of developing economies as a destination for FDI.

Figure 50.6 FDI inflows by groups of economies, 1995–2014 (\$ billion)



Sovereign wealth funds are state-owned investment funds that are invested in a range of assets including shares, bonds, hedge funds and property. These funds are normally financed from surpluses on an economy’s balance of payments, funds raised through privatisations and fiscal surpluses. Many of them were established after 2000. In December 2015, the Sovereign Wealth Fund Institute estimated that the total value of assets held in these funds was \$7204 billion. A government often uses sovereign wealth funds to generate profit to benefit the nation’s economy and its citizens, and globalisation has made it easier for the managers of these funds to take advantage of investment opportunities throughout the world.

The consequences of globalisation for economies

The consequences of globalisation have been profound and varied. It has social, cultural and political as well as economic effects. Some individuals and groups have benefited from globalisation, while others have been disadvantaged. The impact of globalisation has also varied, between less and more developed economies as well as within these categories of economies.

This section refers to less and more developed economies. The degree of a country’s economic development is not simply measured in terms of its GDP per capita (or income per head) of its inhabitants, and thus the rates of economic growth that it has experienced in the past. Economic development also measures the extent to which the inhabitants of an economy are able to benefit from improvements in their living standards. Thus it measures not just income levels but also educational standards, the extent and quality of healthcare and other factors that might affect the quality of people’s lives.

Key terms

Migration is the movement of people between different countries, though it can also refer to movement between regions within a country.

A **transition economy** is one that is moving from being centrally planned, where the state decides on production patterns, to a free market.

Foreign direct investment (FDI) records the transfer of ownership of UK or foreign businesses (i.e. productive assets) between residents in different countries.

Sovereign wealth funds are state-owned investment funds that invest in a range of assets including shares, bonds, hedge funds and property.

Less developed countries would be expected to have lower incomes per head and the economies would be unlikely to have sufficient resources to fund, for example, effective health and education programmes for all of their citizens.

The consequences for less developed economies

The term ‘less developed countries’ encompasses economies at quite different stages of development. It can refer to relatively undeveloped economies, many in Africa such as Ethiopia and South Sudan, or countries that are significantly more developed such as Brazil and India. It is important to appreciate that the impact of globalisation has not been the same for all less developed economies. Some economies, such as India, China, Mexico and South Africa, have benefited far more from the process. The term used for this group of countries is ‘developing economies’. Other, more undeveloped economies have frequently been affected in more negative ways by globalisation.

Author tip

We look in greater detail at the concept of economic development in Chapter 54. This chapter also examines how economic development can be measured, as well as the factors that promote development and the barriers that can exist.

1 Economic growth

The rise of developing economies is inextricably tied up with globalisation. Globalisation has made international trade easier by reducing political and legal barriers and by improving international communications and transport links. These changes have allowed developing economies such as China, India and Mexico to thrive. Freer trade and political systems have allowed businesses to succeed in developing economies and have encouraged established multinational corporations to locate in these countries. At the same time, domestic businesses in these countries have grown rapidly. Such changes, in total, have helped some developing economies to achieve very high rates of economic growth.

China is probably the most immediately recognised of the so-called developing economies. It has received much attention in the media for its spectacular rates of economic growth. Although this is beginning to slow, its economy still experienced an economic growth rate of 6.9% in 2015. In 1980 the Chinese economy was about 25% of the size of the UK’s; by 2015 it was more than twice as large. China has been extraordinarily successful at lifting its people out of poverty, removing more people from poverty than anywhere else in the world: its per capita income increased 25-fold between 1990 and 2010, from \$200 to \$5000, moving China into the ranks of middle-income countries.

Other developing economies include India, Russia, Brazil, Turkey, Mexico, Poland, Indonesia, Saudi Arabia, Taiwan, Vietnam, Iran, Argentina and Thailand.

However, not all less developed economies have enjoyed rapid increases in their rates of economic growth as a result of globalisation. Some of the world’s poorest nations have experienced few or no benefits from globalisation in terms of average incomes. They may have few natural resources to exploit and do not attract multinational corporations to locate there. At the same time, their most talented people often migrate to more prosperous countries. These economies may rely heavily on exporting

Key terms

Economic development

is a process whereby average incomes rise in an economy along with other measures of economic wellbeing such as standards of education and healthcare.

Developing (or emerging) economies

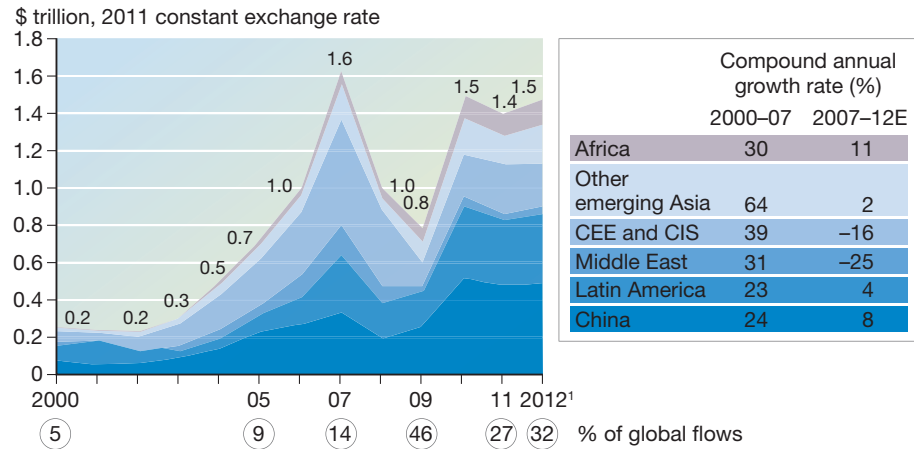
are those that have relatively low incomes per capita but generally enjoy high rates of economic growth.

agricultural products, the prices of which can fluctuate wildly, and their ability to negotiate improved prices in global markets may be very limited.

2 Capital inflows

Figure 50.7 illustrates that FDI in developing economies increased from approximately \$250 billion in 1995 to around \$1500 billion in 2014. These inflows underpinned rapid increases in the rates of economic growth enjoyed by these economies, financing investment in developing an efficient infrastructure, enormous increases in productive capacity and improvements in the skills of workforces. This investment has been essential in allowing these economies to improve the living standards of millions of their people.

Figure 50.7 Capital inflows to developing economies by region, 2000–12 (\$ trillion)



¹ Estimated based on data through Q2 2012. For countries without quarterly data, we use trends from the institute of International Finance.

Sources: IMF Balance of Payments; Institute of International Finance; McKinsey Global Institute analysis

The size of the flows of FDI has varied across less developed economies, as shown by Figure 50.7. It is apparent that China received the lion’s share of this investment, but that Latin America was also a major recipient with Brazil, Mexico and Argentina as notable beneficiaries.

3 Environmental impact

Globalisation has increased rates of economic growth in many less developed countries. This growth has meant the building of new factories and offices, new roads and airports, as well as dramatic increases in manufacturing and use of fossil fuels. These trends have brought about diverse environmental problems including air, soil and water pollution, water shortages, industrial waste, and depletion of natural resources such as timber and copper.

The exploitation of oil reserves and minerals remains a priority for many governments of less developed economies. The environmental consequences of such activities can have significant impacts on local communities that rely upon the natural environment to sustain a traditional lifestyle. Powerful multinational companies often exploit weak laws relating to the environment in these countries, as well as corrupt officials and an unwillingness of the authorities to enforce laws that do exist. For example, the multinational mining corporation Rio Tinto has been accused of polluting the air

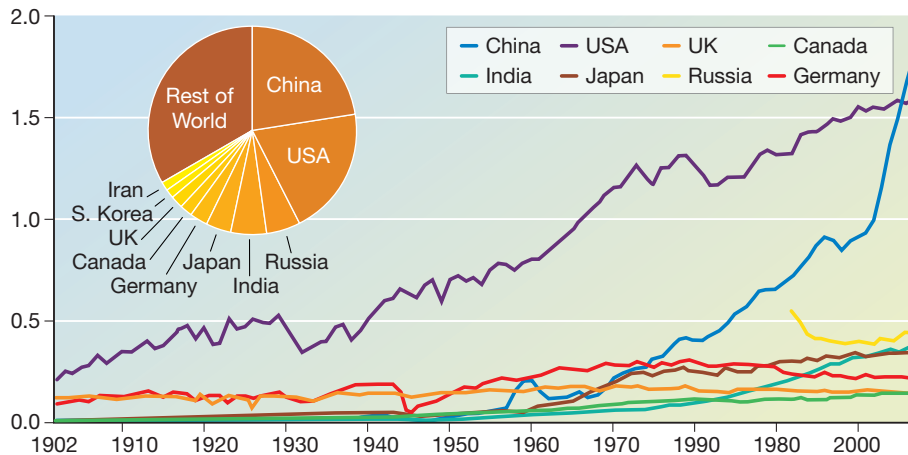


A woman covers her nose with a scarf on a Beijing street during severe air pollution, January 2013

and water surrounding its operations in developing economies such as Indonesia and South Africa. Norway has prohibited its sovereign wealth fund from investing in Rio Tinto because of concerns about the company's environmental performance.

Figures 50.8 and 50.9 highlight the increasing problem faced by some less developed economies with regard to pollution. China has become the world's largest source of carbon emissions, while emissions from India are growing quickly and well in excess of emissions from large developed economies such as Germany and the UK.

Billion metric tonnes of carbon



Source: Reuters

Figure 50.8 Global carbon emissions, 1902–2009

Figure 50.9 illustrates the extent of pollution in countries throughout the world based on a survey and information from organisations such as the World Health Organisation. The results are on a scale from zero (the best and green) to 100 (the worst and deep red). Developing economies such as Russia, China, India, Indonesia and Mexico are shown to be suffering badly from the effects of pollution.

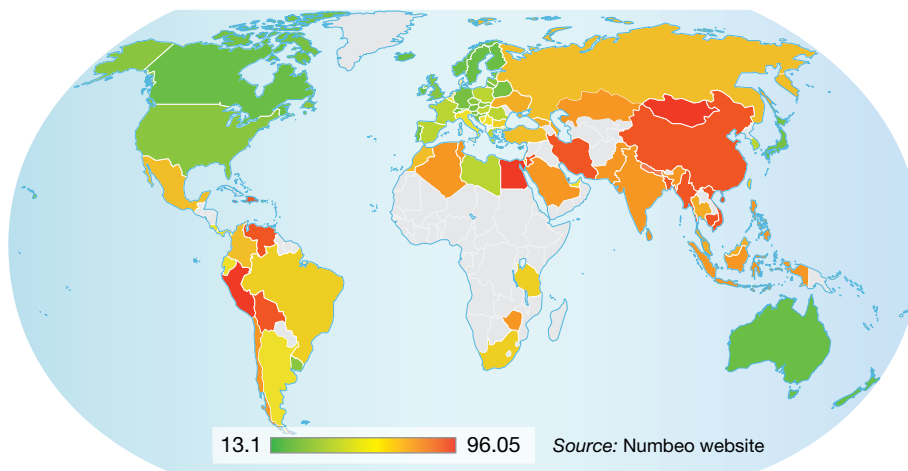


Figure 50.9 The countries of the world shown on a pollution index, 2016

However, the level of environmental damage taking place in less developed economies could be reduced in the near future. The 2014 Low Carbon Economy Index compiled by PricewaterhouseCoopers (PwC) has provided some optimism for the future. It showed that the E7 group of developing economies (China, India,

Brazil, Mexico, Russia, Indonesia and Turkey) appears to have recognised the benefits of environmentally friendly methods of economic growth. The E7 group is currently reducing carbon emissions substantially faster than the G7 (the group of seven developed economies including the UK and USA) for the first time.

The consequences for more developed economies

1 Deindustrialisation

Key term

Deindustrialisation

describes the process of removing manufacturing industries from an economy and the economic and social consequences, such as unemployment, which follow.

Globalisation has played a part in the deindustrialisation of the UK as well as other more developed economies. It has made industries in more developed economies vulnerable to competition from overseas and has undoubtedly contributed to structural unemployment. As a result, manufacturing in advanced economies such as the UK and USA has declined (Table 50.1). In the UK, approximately 2.6 million people work in manufacturing compared to over 5 million in the late 1970s. In 2015, 8% of the UK's workforce was in manufacturing – a substantial fall from a figure of 25% in 1978. Manufacturing only accounted for about 10% of the UK's GDP in 2015.

Globalisation has resulted in increased imports of manufactures and the loss of jobs as many manufacturers have moved operations to countries with lower wage costs, such as Poland and Vietnam. At the same time, developing nations have established their own manufacturing multinational corporations which have been increasingly successful in global markets, often benefiting from lower wage costs, making it more difficult to export manufactures. Deindustrialisation has impacted particularly on those carrying out relatively unskilled work and has resulted in increased unemployment, often in specific areas or regions.

Table 50.1 Share of manufacturing value added as a percentage of GDP for selected countries, 1995–2011

Country	1995	2000	2007	2011
Developing economies				
Brazil	20.7	19.5	19.6	17.5
China	35.1	33.8	34.8	33.4
India	20.7	18.3	18.9	16.8
Russia	18.6	21.8	19.4	18.1
Developed economies				
Germany	23.7	24.3	25.4	24.2
France	15.2	17.4	13.8	11.7
UK	21.7	17.7	12.7	12.4
Italy	23.1	22.0	20.1	17.7
Japan	22.7	21.3	21.1	18.8
USA	15.7	14.5	12.4	12.5
World	20.1	19.0	17.9	17.9

Source: Government Office for Science, Future of Manufacturing Project: Evidence Paper 34

Economies such as the UK have experienced deindustrialisation since the 1980s. Globalisation has been a key cause, and some advanced economies have suffered balance of payments difficulties as a consequence. The UK has had a large deficit in goods and this has resulted in the economy suffering a large and long-lasting deficit on the current account of its balance of payments. It still has some manufacturing

industry, but it has become a net importer of manufactured goods such as televisions and tractors.

However, not all developed economies have deindustrialised to the extent that has taken place in the UK and USA. Germany remains an important manufacturer: the country specialises in manufacturing vehicles, machinery, electrical equipment and chemicals. In 2014, Germany achieved annual sales of vehicles totalling €385 billion or 14% of its GDP. Table 50.1 shows the extent to which Germany has maintained its manufacturing sector, while other developed economies have lost out to less developed economies.

2 Increased competition for most industries

Developing economies like Mexico and China can produce goods and services cheaply. Developments in information and communications technology (ICT) have lowered the costs of logistics, so that globalisation makes it harder for developed nations to compete with lower cost producers. However, businesses based in less developed economies also have access to cheaper sources of labour and resources, enabling them to produce more price competitively as well. Thus there has been a downward pressure on final selling prices in many industries, which has contributed to historically low rates of inflation in developed economies.

A recent trend resulting from the process of globalisation is the increasing market influence of firms based in less developed economies. They have benefited from adopting modern production techniques and operating in economies in which infrastructure is improving, often rapidly. When combined with supplies of labour that are cheap in global terms, these businesses are becoming powerful rivals to those based in more developed economies.

3 Migration

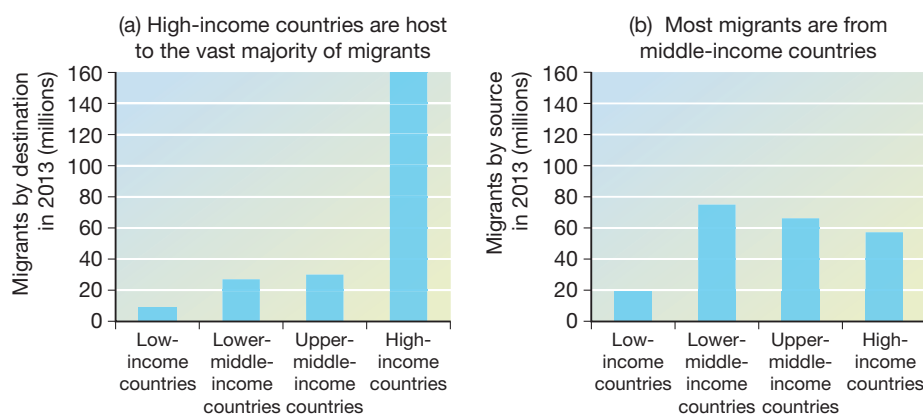


Figure 50.10 Hosts and sources of migrants in 2013, categorised according to income levels of economies

Source: Global Monitoring Report 2015/2016: Development Goals in an Era of Demographic Change

We saw earlier in this chapter that increased migration has been a major feature of globalisation. One significant element of this is flows of labour from less developed to more developed economies attracted by higher earnings and the prospect of improved living standards. A recent report from the World Bank and the IMF predicts that large-scale migration from poor countries to richer regions of the world will be a permanent feature of the global economy for many years ahead. The publication of

‘Global Monitoring Report 2015–16: Development Goals in an Era of Demographic Change’ in October 2015 showed the extent to which developed economies have been hosts to the majority of migrants. In 2013, high-income developed countries received approximately 160 million migrants. Around 100 million of these came from less developed economies, as illustrated in Figure 50.10.

REALWORLD ECONOMICS 50.1

Winners and losers from globalisation

It is generally thought that two groups are the big winners from the past two decades of globalisation: the very rich (those at the top of national and global income distributions); and the middle classes of developing economies, especially in China, India, Indonesia and Brazil.

The top 1% of the global income distribution have seen their real income (adjusted for inflation) rise by more than 60% over those two decades. What is far less known

is that an even greater increase in incomes was realised by those parts of the global income distribution that lie around the median. They achieved an 80% real increase in incomes.

Broadly speaking, the most significant consequences of globalisation for incomes have been as follows:

- the bottom third, with the exception of the very poorest, have become significantly better off, and many people there have

escaped absolute poverty;

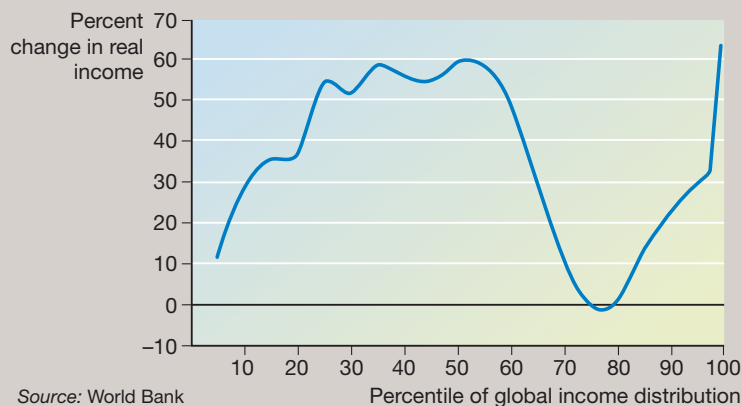
- the middle third (or more) have become much richer, seeing incomes rise, in real terms, by approximately 3% per capita annually;
- while the top 1%, and somewhat less so the top 5%, have gained significantly, the next 20% in the global income distribution have either gained very little or faced stagnant real incomes.

This has created polarisation among the richest quartile of the world’s population, allowing the top 1% to pull ahead of the other rich people. It has also reaffirmed — both in fact and, even more so, in public perception — the top 1%’s clear role as winners of globalisation.

Source: The Globalist website

Exercise

- 1 On the basis of the information provided here and your broader knowledge of globalisation, have less or more developed economies benefited most from globalisation? (9 marks)



Source: World Bank
Figure 50.11 Percentage change in real income at various percentiles of global income distribution, 1988–2008

The role of multinational corporations in globalisation

Multinational corporations (MNCs) have played a central role in globalisation (Figure 50.12). We have already seen in this chapter that many have been quick to take advantage of the greater economic and political freedoms that have driven the process. They have uncovered and utilised new sources of productive resources including raw materials and labour. This has resulted in a number of outcomes that have supported globalisation.

- The activities of MNCs have encouraged a greater volume and diversity of international trade. MNCs have established production facilities in countries where production can be carried out more cheaply and have helped national governments to exploit the resources that exist within their boundaries.

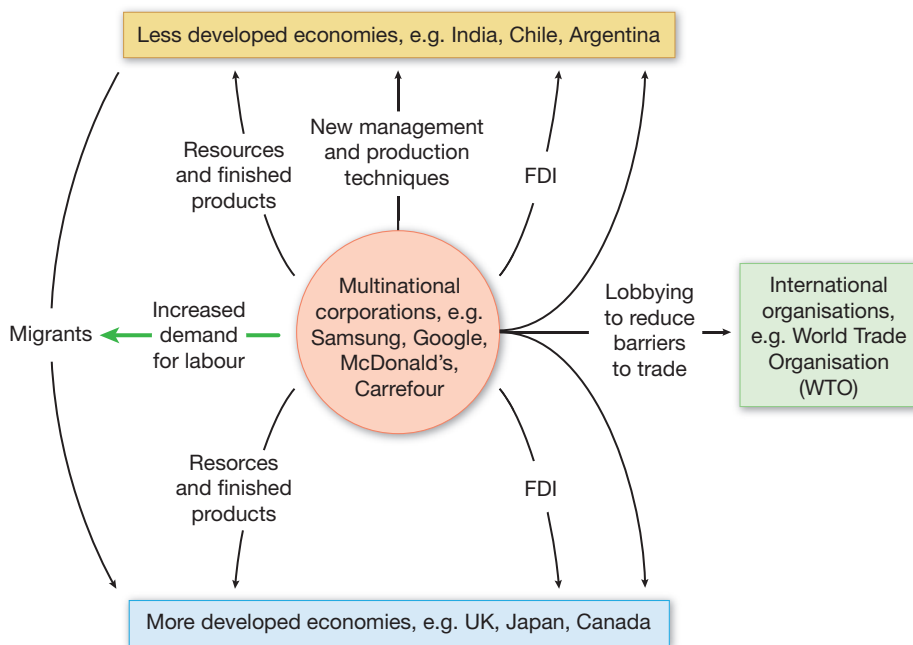


Figure 50.12 *The central role of multinational corporations in globalisation*

- MNCs have spread knowledge and understanding of modern management and production methods as they have established facilities in less developed countries. This has helped domestic firms in those economies to become more efficient and to compete effectively in global markets. In part this explains the rise of new businesses such as the Chinese computer technology company Lenovo. MNCs have also provided training that has enhanced the skills and productive capabilities of workforces in less developed economies. This has helped, in turn, to make these economies more attractive to other MNCs.
- By lobbying governments in less developed economies, MNCs have presented the case for globalisation and set out the benefits to these economies, such as rising employment rates and increased rates of economic growth. This intervention has helped to persuade governments to implement policies that have fostered globalisation. MNCs have also lobbied international organisations such as the World Trade Organisation (WTO) to reduce barriers to trade. Free trade has been a cornerstone of globalisation.
- Many MNCs have established supply chains that are truly global. This has created opportunities in a wide range of economies, both more and less developed, creating more interconnected economies.
- The activities of MNCs have been crucial in encouraging migration, both within countries and between countries. Labour has been essential in order to achieve the high rates of economic growth that have accompanied globalisation. MNCs have sought to persuade governments to allow suitably skilled employees to move to locations where they have needed workers.

Review questions

Total: 44 marks

- 1 Define the term 'globalisation'. (2 marks)
- 2 Explain why transport is an essential element of globalisation. (5 marks)
- 3 Countries that concentrate on producing a limited range of goods and services are engaging in which of the following?
A Deindustrialisation
B Specialisation
C Diversification
D Hysteresis (1 mark)
- 4 Explain why the rise in incomes in countries such as India has been a factor contributing to globalisation. (6 marks)
- 5 Explain one way in which the managers of sovereign wealth funds have benefited from the process of globalisation. (4 marks)
- 6 Explain how the process of globalisation has affected flows of migrants between countries. (6 marks)
- 7 Explain why some less developed economies may have had relatively few benefits from the process of globalisation. (7 marks)
- 8 Why has globalisation led to deindustrialisation in some developed economies? (6 marks)
- 9 Explain why some multinational corporations based in developing nations such as India and China have been increasingly successful in global markets. (6 marks)
- 10 Which of the following is *not* a reason for the growth in the size of multinational corporations?
A An increase in the number of joint ventures
B An increase in minimum efficient scale in many industries
C An increase in the size of markets in which businesses trade
D Downward pressure on prices, making it essential for firms to reduce production costs (1 mark)

Trade

Key concepts from Year 1

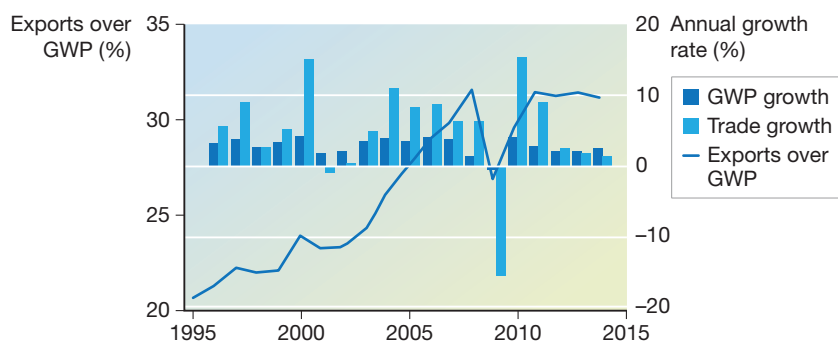
This chapter applies the concept of specialisation that we discussed in Chapter 13 of the Year 1 companion textbook. We also introduced the principle of comparative advantage in Chapter 41 and we build upon that introduction here.

In this chapter we examine the benefits to economies of engaging in international trade, including the model of comparative advantage. We balance this by considering the costs of international trade. We describe the pattern of the UK's trade with the rest of world and the main changes that have taken place in recent times and explain the reasons for these changes. We also outline the major protectionist policies that can be used by governments and analyse the reasons why countries adopt such policies as well as the consequences of their use. This chapter goes on to look at the main features of a customs union and the key characteristics of the European Union's Single European Market. It analyses the consequences for the UK of its membership of the EU before considering the role of the World Trade Organisation.

The benefits and costs of international trade

Most economists believe that international trade is beneficial to individual economies, even though it does have costs (as we shall see later). The volume of international trade (i.e. the quantity of goods and services exchanged, rather than their value) has increased by an annual average of around 5% over the last 30 years, despite fluctuations brought about by the economic cycle. The growth in international trade since 1995 is compared to the growth in the world's GDP (gross world product or GWP) in Figure 51.1.

It is apparent from Figure 51.1 that international trade is affected by the economic cycle – the recession of 2008–09 provoked a 12% fall in the volume of goods that



Source: United Nations Conference on Trade and Development (UNCTAD), Key Statistics and Trends in International Trade, 2015

Key terms

International trade is the exchange of goods and services across national frontiers.

International merchandise trade is the exchange of goods (for example, cars and televisions) between two or more countries.

Figure 51.1 Growth in international trade, gross world product (GWP) and exports as a percentage of GWP, 1995–2015

were traded. It is also apparent that a positive relationship exists between trade and the world's GDP – shown as GWP in Figure 51.1. We will consider the relationship between trade and GDP in more detail later.

Another interesting feature of Figure 51.1 is the steady increase in exports as a percentage of GWP up until the recession of 2008–09. This indicates the increasing importance of trade to economic growth as well as the impact of globalisation on macroeconomic performance.

The benefits of international trade

Some economists do not believe that all economies benefit from trade and they frequently cite the disadvantages of international trade to underdeveloped countries. Despite this, there is a widespread view that economies benefit overall from engaging in international trade.

The theory of comparative advantage

Some countries have an absolute advantage in the production of goods and services. Such countries are able to produce a unit of output of a good or service using fewer resources than others. The principle of absolute advantage explains some trade between developed and less developed economies. A developed economy may be more efficient at producing complex manufactured products while a less developed economy may be more productive in producing minerals in which it is abundant. Specialisation and trade between these two countries would benefit both economies.

In the nineteenth century the British economist David Ricardo developed what is possibly the most fundamental argument in support of international trade. His theory of comparative advantage states that countries benefit from international trade even if they are relatively inefficient at producing goods and services in comparison to their competitors (i.e. if they do not have an absolute advantage). Ricardo's theory showed that output would be greater if an economy concentrated on producing products at which it is most efficient, even if other economies are more productive in producing those goods and services. He argued that an economy should produce those goods and services in which it has the least disadvantage. Ricardo described this as a comparative advantage.

The principle of comparative advantage states that countries achieve the highest possible levels of GDP by using the resources available to them to specialise in producing goods and services in which they have a comparative advantage and then by engaging in international trade.

Key terms

Absolute advantage exists when one country requires a smaller quantity of inputs to produce a good than another country.

Comparative advantage occurs when one country can produce a good or service at a lower opportunity cost than another.

Specialisation occurs when an individual, firm, region or country concentrates on producing a limited range of products.

Opportunity cost is the next best alternative that was given up when making a choice.

A **production possibility diagram** shows the maximum combination of products that can be provided by an economy during a given period of time with the resources available.

We can illustrate the workings of Ricardo's theory by developing a numerical example based on a simple scenario. Assume that the world comprises just two countries, Eldorado and Utopia, and that they are able to produce just bread or beer. These two countries have the same amount of resources and, using their resources fully and divided equally between producing the two products, the production levels of the two countries would be as shown in Table 51.1.

	Utopia	Eldorado	World
Bread	300	100	400
Beer	300	200	500

Table 51.1 Production assuming self-sufficiency

In these circumstances, Utopia has an absolute advantage in the production of both products. However, Ricardo showed that it was still possible for world production (and living standards) to be greater if Eldorado specialised in producing the product at which it had a comparative advantage. An economy has a comparative advantage if the opportunity cost of producing the good is lower there than in other economies. If an economy only produces two products, as in our example, then the opportunity cost of producing one product is the amount of production of the other product that is foregone following this decision. The opportunity cost ratios for Utopia and Eldorado are set out in Table 51.2.

Country	Opportunity cost ratio
Utopia	1 unit of bread = 1 unit of beer, or 1 unit of beer = 1 unit of bread
Eldorado	1 unit of bread = 2 units of beer, or 1 unit of beer = 1/2 a unit of bread

Table 51.2 Opportunity cost ratios for Utopia and Eldorado

Table 51.2 shows that Eldorado has a comparative advantage in the production of beer. It only has to give up half a unit of bread for each unit of beer that it produces. Utopia has to give up a full unit of bread to produce an additional unit of beer. Thus, the opportunity cost of producing beer is lower for Eldorado than Utopia. Ricardo argued that it was such differences in opportunity cost ratios that allow countries to benefit from specialisation and trade.

	Utopia	Eldorado	World
Bread	600	0	600
Beer	0	400	400

Table 51.3 Total specialisation based on the principle of comparative advantage

Table 51.3 illustrates the effects if Eldorado specialises in the production of beer, in which it has a comparative advantage, and Utopia just produces bread. The outcome is a significant increase in the world production of bread but a fall in beer production. This occurs because the less efficient economy has focused on producing this product. However, if Utopia only specialises to a limited degree, as shown in Table 51.4, then world production of both products has risen above the self-sufficiency level shown in Table 51.1.

	Utopia	Eldorado	World
Bread	480	0	480
Beer	120	400	520

Table 51.4 Limited specialisation

The final stage of the process is that international trade should take place. This is shown in Table 51.5. We assume Eldorado and Utopia agreed to trade 120 units of bread for 190 units of beer. As a result, both countries have greater volumes of both products, offering the potential for a higher standard of living for all their citizens following specialisation and trade.

Table 51.5 The result following international trade

	Utopia	Eldorado	World
Bread	360	120	480
Beer	310	210	520

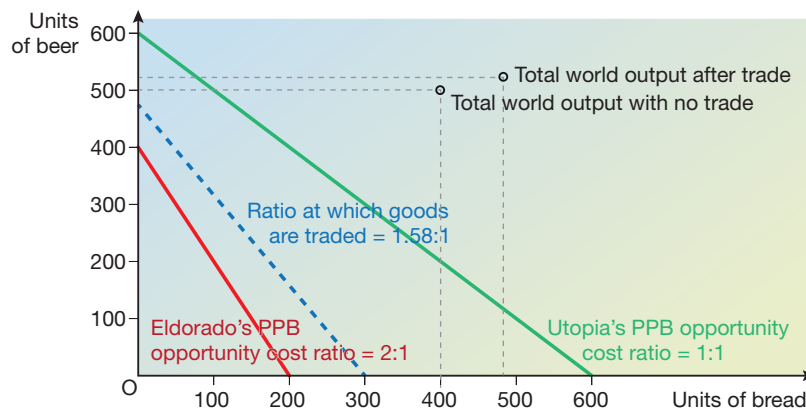
The rate at which the two products were exchanged in international trade plays an important role in bringing about benefits for both countries. The rate of exchange under which the international trade was conducted was 120 units of bread for 190 units of beer. That equates to 1 unit of bread for 1.58 units of beer or 0.63 units of bread for 1 unit of beer. The reason that this trade provides benefits to both countries is that it lies between the two countries' opportunity cost ratios for these products under self-sufficiency. We saw earlier that these were:

- Eldorado: 1 unit of bread to 2 units of beer;
- Utopia: 1 unit of bread to 1 unit of beer.

Thus Eldorado benefits because it receives 1 unit of bread for each 1.58 units of beer it trades. If it were to produce the bread itself, it would have to forego 2 units of beer for a unit of bread. Similarly, Utopia receives 1 unit of beer in exchange for 0.63 units of bread. If it were self-sufficient, it would forego 1 unit of bread to produce an additional unit of beer. Hence, even though Eldorado only has a comparative advantage in the production of beer, specialisation and trade are mutually beneficial. If, however, the two countries had the same opportunity cost ratios (which is extremely unlikely), meaning their production possibility boundaries were parallel, no comparative advantage would exist and there would be no benefits from trade.

The two countries' opportunity cost ratios can be shown using a production possibility diagram (Figure 51.2).

Figure 51.2 Production possibility boundaries for Eldorado and Utopia



This model is, of course, a simplification of the real world in a number of ways.

- It assumes only two countries and two products, and that no resources are used in transporting products across national frontiers. In reality these could be substantial and would reduce, or possibly eliminate, the benefits of trade.

Furthermore, international transport is likely to generate negative externalities that could reduce the benefits of trade.

- The model ignores possible currency fluctuations that could take place and may reduce the benefits accruing within the model.
- It assumes that costs (as shown by opportunity cost ratios) are constant – hence the straight-line production possibility boundaries in Figure 51.2. It would be reasonable to assume that specialising in either beer or bread would allow Utopia and Eldorado to receive the benefit of economies of scale as they devote more resources to the production of a single type of product. Thus unit costs of production should fall as output increases.
- The model assumes that perfect competition exists within the economies and that factors of production are perfectly mobile.

Achieving macroeconomic objectives

The theory of comparative advantage tells us that specialisation and trade can allow a country to increase the volume of goods and services that it produces. Trade can therefore help governments to fulfil their macroeconomic objectives. Increases in production increase the level of an economy's GDP and can create higher levels of employment. Higher employment in exporting industries is likely to create a positive multiplier effect, further increasing incomes and employment.

Recent figures for the UK emphasise the contribution that trade can make to rates of economic growth. Figure 51.3 shows that the UK's poor trade performance (net trade) in the third quarter of 2015 contributed to a slowing in the economy's rate of growth of GDP.

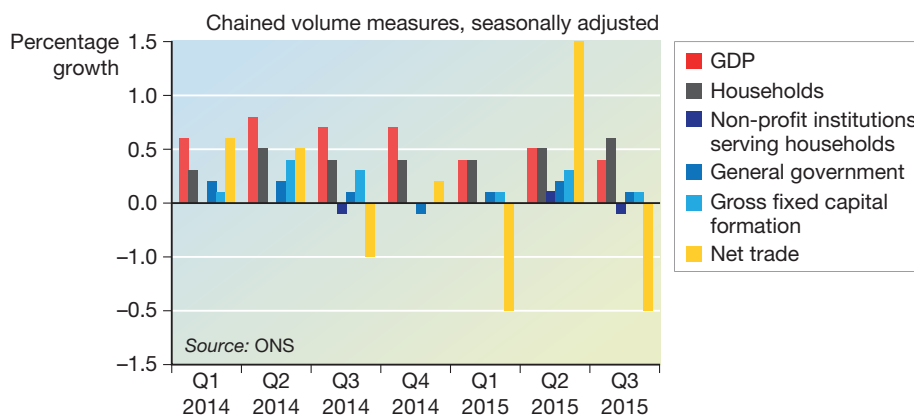


Figure 51.3 UK expenditure components' percentage contribution to GDP growth, Quarter 1 (Jan–Mar) 2014 to Quarter 3 (July–Sep) 2015

International trade can also assist in controlling inflation. It is perhaps no coincidence that the rise in the volume of international trade that has accompanied globalisation has provided a period of very low rates of inflation in many countries. By increasing competition and encouraging efficient producers to dominate markets, prices tend to be lowered.

Another dimension of trade is that it tends to lead to the ending of inefficient domestic monopolies as consumers have the option of purchasing goods and services from efficient firms based in other countries. This adds to the downward pressure on prices in many markets as monopoly profits are competed away.

Benefits to consumers

International trade offers a range of benefits to consumers. It enables them to purchase products that may not be supplied by domestic firms or that are only available at exceptionally high prices. International trade allows consumers in the UK to purchase products such as tropical fruit that might not otherwise be available or affordable.

Consumers benefit from the lower prices and the higher quality that tend to accompany international trade. Prices are lowered as less efficient producers lose market share and possibly cease trading. Lower prices increase the consumer surplus enjoyed by consumers as well as increasing their purchasing power, providing a hike in their standards of living.

The costs of international trade

While international trade brings many benefits, significant drawbacks can also occur.

Structural and regional unemployment

Businesses in some domestic industries may struggle when facing competition from efficient firms in other countries. Overseas firms may have advantages in terms of lower production costs, especially labour, which can provide a major competitive advantage. We saw in Chapter 50 that many manufacturing businesses in developed economies such as the UK have failed due to competition from businesses located in low-cost developing economies. This process is called deindustrialisation.

The outcome of the failure of these businesses can be unemployment, which is centred on particular industries or locations. Structural unemployment occurs as a result of job losses from the long-term decline of specific industries. It is not unusual for these industries to be located in specific regions of an economy, thereby imposing heavy burdens on particular areas. The coal industry in the UK failed because of competition from lower-cost mines in other countries and it was based in a few locations in South Wales, the central belt of Scotland, the Midlands and the North of England. The effects on the local economies have been devastating – the loss of dozens of coal mines resulted in a strong reverse multiplier effect.

Too much specialisation

There is a risk that economies become too specialised. If the industries on which an economy depends experience an external shock, such as a dramatic increase in competition from overseas, the impact on employment and GDP in the economy can be strongly negative. The UK suffered heavily as a result of the financial crisis in 2007–08 because its economy is so dependent upon financial services.

Difficulties developing new industries

Some industries, in particular in developing economies, may struggle to establish themselves in the face of fierce international competition. International trade can destroy infant industries before they have a chance to grow and to benefit from recognised brand names and economies of scale. This can make it very tough for developing economies to diversify their economies and reduce their reliance on supplying primary products such as cotton and copper.

The UK's pattern of trade with the rest of the world

It is possible to analyse changes in the UK's patterns of trade in terms of the countries with whom the UK trades as well as in the goods and services that are exported and imported.

Changes in the geographical pattern of trade

There have been significant changes in the relative importance of the UK's trading partners over recent years.

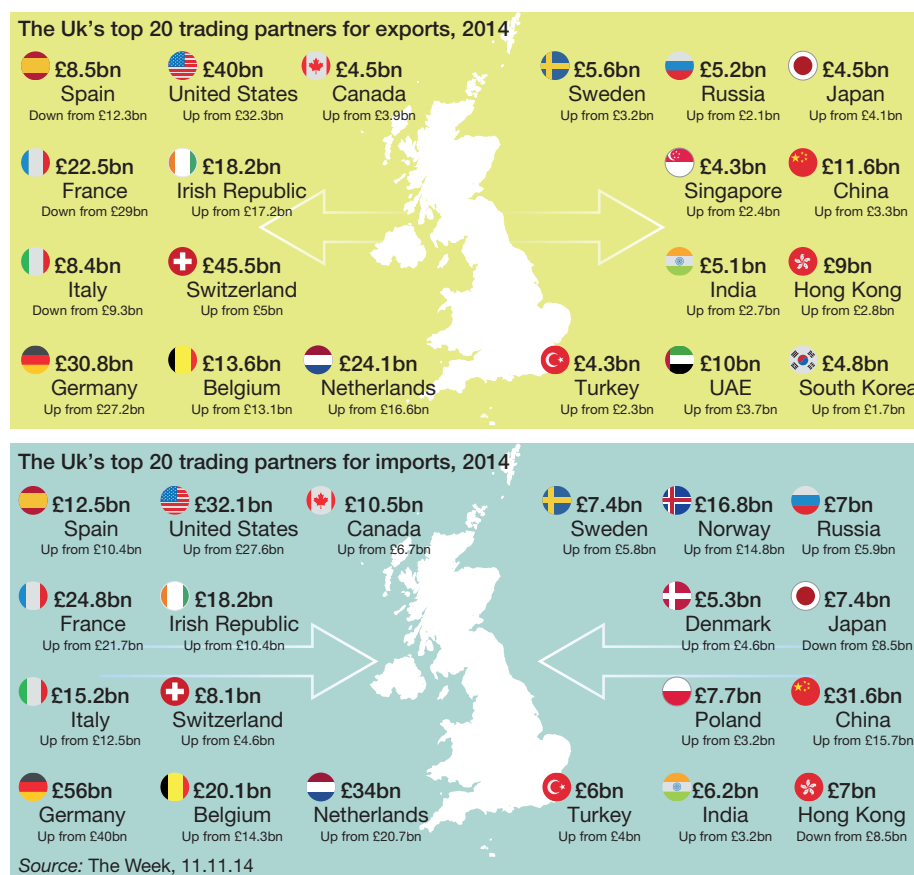


Figure 51.4 Changes in the UK's geographical pattern of trade, 2006 and 2014

Figure 51.4 shows that the UK's trade with many countries has increased markedly and Table 51.6 highlights the major changes in the UK's trading patterns that have taken place since 2008.

The UK has a number of trading partners that have been important for many years, and continue to be so. Germany, France and the United States fall into this category. The 27 other member states of the European Union (EU) remain a vital destination for UK exports, with 44% sold there. However, within the EU there are significant changes taking place in terms of countries that are buying increasing values of UK exports. The former communist countries of Eastern Europe, such as Poland and Bulgaria, are becoming important markets for UK businesses as consumers' incomes in these countries rise rapidly. The EU is the most important source of imports for

Key term

The **eurozone economy** is the economic region comprising the 19 EU countries that use the euro as their national currency.

the UK. In 2015, just over 53% of the UK's imports arrived from other EU countries. Table 51.6 makes it clear that the UK runs a substantial trade deficit with the rest of the EU.

Table 51.6 Value of exports and imports for selected countries, 2008 and 2015

Country/ region	Exports			Imports		
	2008 (£ million)	2015 (£ million)	Average annual percentage change	2008 (£ million)	2015 (£ million)	Average annual percentage change
Germany	28,535	30,352	0.8	44,504	60,679	4.0
Poland	2932	3600	2.6	4241	8094	8.4
Bulgaria	201	346	10.2	248	346	4.3
Lithuania	277	270	-0.3	343	769	10.6
Greece	1615	877	-7.3	616	714	1.9
Total EU	141,068	133,365	-0.7	167,821	218,503	3.4
China	4878	18,071	17.8	21,969	36,103	6.4
India	4187	3911	3.5	4267	6063	4.5

Sources: HMRC, Overseas Trade Statistics

Other significant changes in geographical patterns of trade relate primarily to developing nations such as China, India and Russia. All three countries are important trading partners for the UK, although trade with Russia has declined sharply since sanctions were imposed in 2013 following its invasion of Ukraine. India and China are vital markets for many UK products as consumers in these countries enjoy rising incomes. At the same time, the UK imports vast quantities of manufactures from China and services from India.

Number crunching

Use the information in Table 51.6 to answer the following question. Was the rate of increase in the UK's deficit greater for its trade with the EU or non-EU countries between 2007 and 2014?

Changes in the commodity pattern of trade

The pattern of the UK's trade in goods has changed relatively little in recent years. Table 51.7 provides some evidence that the current Conservative government's aim to expand the role of manufacturing in the economy appears to have made a little progress in terms of exports, but the inexorable rise in imports of manufactures has continued. It is the continuing loss of competitiveness of the UK's manufacturing industry that is the most notable feature of the data in Table 51.7. Firms in the UK can compete in certain complex manufacturing processes, but they are unable to compete in the production of a wide range of goods with low-cost producers overseas. In addition, some UK firms have moved manufacturing facilities overseas to take advantage of cheap labour. Thus, although the exports of manufactures rose by 5% in volume terms between 2012 and 2014, import volumes have risen by 13% over the same period. The UK car manufacturing industry has, however, shown a

REALWORLD ECONOMICS 51.1

UK seeks to target new export markets

China and India are expected to be increasingly important trading partners for the UK in the future as their economies continue to develop and their huge populations exercise increasing spending power. Some economists also forecast that Africa will become more significant as a source of UK imports and a destination for its exports due to a combination of high rates of population growth and rising incomes.

The importance of China as a trading partner to the UK economy has increased steadily since 2004, with both imports and exports rising. Following the growth of imports from £11.4 billion to £36.1 billion in 2015, China has become the UK's second largest import partner behind the USA, accounting for over 7.0% of UK imports in 2015 compared with 3.3% in 2004. Compared with the rise in imports, exports to

China have increased at a more subdued rate over the same period, from £4.0 billion to £18.1 billion in 2015. This now accounts for 3.2% of UK exports. The UK's goods trade deficit of £22.1 billion in 2014 is the second highest, behind Germany. This compares with a surplus in services of £2.7 billion. These trends are shown in Figure 51.5.

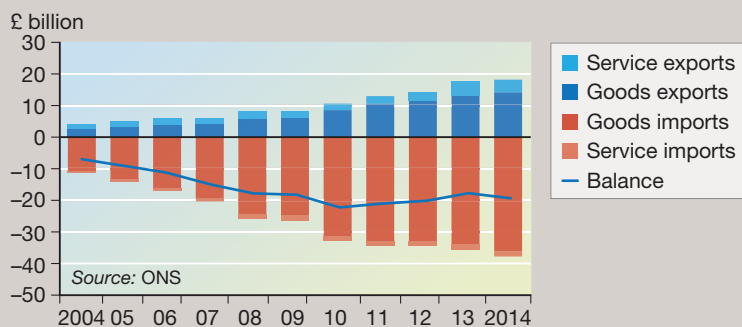


Figure 51.5 Exports, imports and trade balance with China, 2004–14

Exercises

Total: 15 marks

- 1 Explain how the UK benefits from its trade with China. (7 marks)
- 2 Is it a problem that the UK has a large trade imbalance with China? (8 marks)

Discussion point

Is the availability of cheap labour to manufacturers in some foreign countries the major reason why the UK has a deficit in trade on goods?

strong rise in export sales. Major multinational corporations such as Nissan and Honda have established production facilities in the UK and export vehicles to the EU and elsewhere.

Exports	Volume indices (2012 = 100)			Value 2014 (£ million)
	2004	2009	2014	
Oil	113	114	107	32,688
Chemicals	77	91	95	47,975
Cars	62	55	95	25,918
Total manufacturing	83	81	105	147,461
Imports				
Food, beverages & tobacco	87	97	103	38,490
Chemicals	75	83	100	52,681
Consumer goods	82	85	119	82,417
Total manufactures	85	79	113	210,782

Source: ONS, data from the Pink Book, 2015

Table 51.7 UK imports and exports of selected goods by volume indices, 2004–14, and by value, 2014

It is noticeable from Table 51.7 that imports of consumer goods, such as televisions and furniture, have increased strongly as the economy has recovered from the deep recession of 2008–09 and consumers' incomes have shown signs of increasing in real terms.

The UK performs much more strongly in international markets for services than those for goods. It has enjoyed a steadily rising surplus on trade in services over the last ten years. Its surplus on trade in services was £27 billion in 2004. In 2014 it reached £89 billion, having increased from £81 billion in 2013. The UK has a number of service sectors in which it has a competitive advantage, such as financial services. This advantage is based on a highly skilled workforce and a reputation for expertise in these areas. However, although it is very important, the financial services sector has made a relatively stable contribution to the UK's trade in services. It has generated a surplus between £9 billion and £11 billion each year since 2008.

UK businesses have also been consistently successful in selling a range of other business services including management consultancy, advertising, market research, engineering and research and development. Earnings from these services (termed 'other business services') have risen steadily since 2004. By 2014, earnings from 'other business service' exports exceeded expenditure on imports of similar services by more than £20 billion. The UK has enjoyed success in selling these services to developing economies as well as developed ones.

Protection

Protection refers to any measures used by governments with the intention of restricting trade in goods and services. It may seem that, in an era during which the world economy has become more interconnected as a consequence of globalisation, barriers to trade have become less relevant. However, this is not the case. Although tariffs and other forms of protection have been steadily reduced since 1947, barriers to trade still remain.

The Organisation for Economic Cooperation and Development (OECD) estimates that scrapping all tariffs on merchandise trade and reducing trade costs by 1% of the value of trade worldwide would boost global welfare by more than £120 billion a year. This would add the equivalent of up to 2% to GDP for some economies.

Some governments still favour turning to protection during difficult economic times, despite the likelihood that any imposition of barriers to trade would almost certainly lead to retaliation by other countries as well as a diminishing of world trade and global prosperity. When many economies were moving into recession in 2008, governments implemented protectionist measures. Governments of some of the world's largest economies imposed nearly 50 policies designed to restrict trade during 2008 and 2009.

Key terms

Protection refers to a range of measures used by governments to restrict the free entry of imports into an economy.

A **tariff** is a tax on an imported product.

A **quota** is a limit on the volume or quantity of a product that can be imported into an economy.

Export subsidies are payments to exporters of goods and services to

enhance the price competitiveness of the firms concerned.

Non-tariff barriers are a range of techniques designed to restrict imports, but not in the form of a tariff.

Dumping is the sale of a product in an export market for less than the cost of production or at a lower price than it is sold in the exporter's domestic market.

The arguments for protection

1 To protect vulnerable industries

Some governments may opt to impose restrictions on the import of certain products to protect fledgling industries. This is known as ‘the infant industry argument’. The protection is intended to be relatively short term to give the industry time to develop some form of comparative advantage and allow it to survive the rigours of global competition. However, it can be difficult for a government to remove such protection once it is given.

Similarly, some governments may opt to protect industries that are reaching the end of their lives and to protect their economies, for a time at least, against the negative consequences of decline, such as increases in structural and possibly regional unemployment. In the 1980s there were calls to protect the UK coal mining industry from cheap foreign imports for this reason.

2 To prevent unfair competition, such as dumping

Methods of protection may be used to prevent domestic firms facing unfair competition from overseas. One example of such competition is dumping, whereby firms sell products in export markets at prices lower than the cost of production or the domestic price. In 2015 there were allegations that Chinese steel producers were dumping surplus stocks of steel onto global markets, depressing prices and threatening the survival of producers in other countries. Several steel producers in the UK closed factories, claiming that they were unable to compete with Chinese imports, and some politicians proposed imposing tariffs or quotas on imports of steel from China.

3 To reduce large and persistent trade deficits

Governments can use protection to reduce persistent current account deficits on their balance of payments. Methods of protection such as tariffs and quotas can be deployed to reduce imports of goods and services. However, this may provoke retaliation by other governments and is unlikely to address the underlying reasons for the economy’s lack of competitiveness in global markets. We look in more detail at the policies that governments can use to correct imbalances on their balance of payments in Chapter 52.

Methods of protection

1 Tariffs

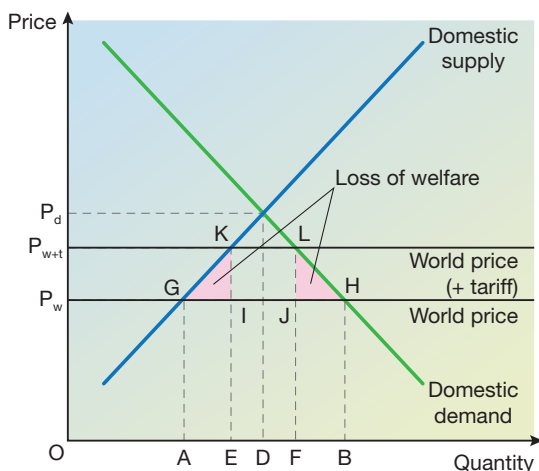
Tariffs are taxes imposed on imported goods and services. The aim is to increase their price relative to those of domestically produced products and to make them less attractive to consumers and firms. There are two types of tariff:

- a specific tariff which is levied at a certain amount per unit of a product;
- an ad valorem (literally ‘by value’) tariff which is levied as a percentage of the price of the import.

There have been many reductions in the use of tariffs in recent years. Nevertheless, they continue to influence trade patterns. Tariffs on agricultural products are on average much higher than those on industrial products, although there is considerable variation between countries.

Tariffs have the potential to bring disadvantages as well as advantages to countries that impose them. Figure 51.6 can be used to help to analyse the effect of the imposition of a tariff on imports of a specific product. In this situation, in the absence of any international trade, price and quantity would be determined by the interaction of domestic supply and demand. Thus the quantity bought would be OD and the market price would be P_d .

Figure 51.6 The effects of imposing a tariff



If we assume that initially free trade takes place without restrictions, then the world price for the product plays a key role. World supply is assumed to be perfectly elastic because sales in this particular country are deemed too small to have an impact on global prices. Thus global firms can supply any quantity to the market at price P_w . With free trade, the most efficient domestic producers can undercut the world price and so OA is supplied domestically and AB (or GH) is imported.

If we now suppose that the government imposes a tariff on imports, the effect is to shift the world supply curve upwards, by the amount of the tariff, to P_{w+t} . This allows more (and less efficient) domestic producers to compete, but reduces the quantity of the product that is bought and sold on the domestic market. The new quantity falls to OF. However, domestic supply to the home market rises to OE and imports fall to EF.

What are the effects of the imposition of this tariff and, in particular, on the welfare or wellbeing of the people in the economy?

- Domestic producers have gained an increase in market share: they now sell OE units rather than OA. The producer surplus has risen by $P_w P_{w+t} KG$ while the consumer surplus has fallen.
- From the government's perspective, employment and national income should rise as a consequence of the tariff. The government also receives tax revenue amounting to KLJI, which it can use to help it attain its macroeconomic objectives.
- The overall effect of the tariff is that welfare is reduced. The increase in price means that the consumer surplus is reduced by $P_w P_{w+t} LH$. This can be offset to some extent by the increase in producer surplus and the government's tax revenue as noted above. However, there is a net loss of welfare, as shown by the two triangles GKI and JLH. This could, of course, be offset to some extent by the government's spending decisions with regard to its additional revenue. This is sometimes termed a deadweight loss.

Discussion point

Do you think that transferring consumer surplus to firms and the government by imposing a tariff, as shown in Figure 51.6, is a good move?

There are potential problems that can arise from the use of tariffs. They entail transfer of welfare from consumers to producers and to the government in the form of revenue. This may not be desirable. Their use invites retaliation by other nations and they may be against the rules of regional trading blocs, such as the EU or the North American Free Trade Agreement (NAFTA).

2 Quotas

Quotas place a limit on the volume or quantity of a specific product than can be imported into an economy. These are normally imposed through the use of a system of licensing. Firms can only import the product if they have a licence to do so, and the number of licences is strictly controlled. For example, in 2015 Japan imposed an import quota of 770,000 tons a year of rice, or around 7% of the total consumed by the country – any imports above this limit are subject to tariffs.

The use of quotas can involve a government in administrative expenditure and can result in officials engaging in corrupt practices or lead to an increase in smuggling. Quotas increase prices and allow less efficient producers to supply the market. Consequently firms benefit from quotas at the expense of consumers.

A significant difference between a tariff and a quota is that the welfare loss associated with a quota may be greater because there is no tax revenue earned by a government. Import quotas generate surpluses for firms that get the licence to import. This is one reason why quotas are less frequently used than tariffs.

3 Export subsidies

Export subsidies are payments to exporters of goods and services designed to increase export volumes. Most commonly applied to agricultural products, they can increase demand for products overseas by enhancing the price competitiveness of the exporters.

Export subsidies benefit consumers in the importing country at the expense of those in the exporting economy. Research presented at the Royal Economic Society's annual conference in 2013 claimed that if China removed its export subsidies, consumers around the world would see their real incomes fall by 1%, while China's national income would rise by 3%. China's export subsidies usually take the form of tax rebates that are dependent upon a firm exporting all or most of its output.

Customs unions

A customs union is an extension of a free trade area. The latter removes all barriers to trade between countries within a group, while a customs union also imposes a common external tariff. This means that imports to any member of the customs

Key terms

A **customs union** exists when a group of countries establish a free trade area with a common external tariff.

A **common external tariff** is the uniform tariff rate imposed on imports by all members of a customs union.

Trade creation occurs when consumers within a customs union purchase products from an efficient producer elsewhere in the union rather than from relatively high-cost domestic producers.

Trade diversion refers to trade diverted away from efficient global producers as a result of the creation of a customs union.

union are subject to the same rate of tariff. A customs union is a deeper form of integration than a free trade area because it generally requires more coordination and a greater loss of autonomy as member states surrender their separate international commercial policies and give up the right to sign trade agreements. Countries can achieve different stages of economic and political integration, as illustrated in Table 51.8.

Table 51.8 Stages of economic and political integration

	Removal of internal barriers to trade	Common external tariff	Free movement of products and factors of production	Single currency and harmonised economic policies	The creation of common political institutions
Free trade area	✓				
Customs union	✓	✓			
Common or single market	✓	✓	✓		
Monetary and economic union	✓	✓	✓	✓	
Political union	✓	✓	✓	✓	✓

The member states within a customs union are able to specialise to a greater extent and to develop their comparative advantage in the production of specific products. Businesses in states within the customs union are also able to enjoy the benefit of economies of scale because they are able to sell to the larger market throughout the union. Jacob Viner, a Canadian economist, developed the concepts of trade creation and trade diversion in relation to the establishment of customs unions.

Trade creation

The creation of a customs union can be expected to result in higher volumes of trade between member states. When it is established, consumers in member countries are likely to take advantage of the opportunity to buy tariff-free goods and services produced by efficient firms based within the union.

Figure 51.7 The effects of trade creation

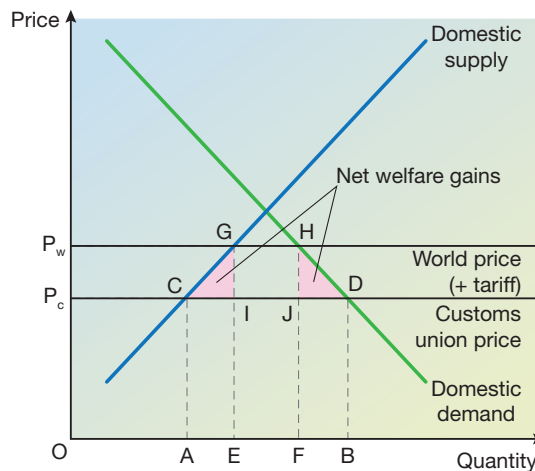


Figure 51.7 illustrates the effects of trade creation, which occurs when a country joins a customs union. Prior to joining the customs union, the country was able to buy supplies on the world market subject to paying a tariff. The world supply

was perfectly elastic at price P_w . Domestic supply was OE and imports were EF to fulfil domestic demand, which was OF. The government received revenue from tariffs equivalent to IGHJ.

Upon joining the customs union, the tariff is removed and the market price becomes P_c because the country's consumers can purchase supplies from efficient producers within the customs union. As a result, the country's consumers receive an additional consumers' surplus equivalent to $P_c P_w HD$. However, the country does not benefit to this extent because the area IGHJ was previously tariff revenue for the government, and $P_c P_w GC$ represents a fall in the producer surplus that occurred as a result of the fall in market price. Thus the trade creation has led to some redistribution of income from producers and governments to consumers. However, the areas CGI and JHD are the net welfare gains to the country resulting from the trade creation.

Trade diversion

The decision to become a member of a customs union may result in other changes to the trading pattern experienced by a country. It may be that as a result of joining a customs union, a country is not able to import products from highly efficient producers in countries outside the customs union. The goods and services supplied by these producers are now subject to the customs union's common external tariff and thus the new member imports supplies from less efficient producers within the union. This situation does not generate new trade: it merely diverts trade from a more efficient source in the rest of the world to a less efficient one within the union. The results can be more complex in terms of the effect on welfare within the new member state. There is no certainty that welfare will increase.

Dynamic effects

The trade creation and diversion effects that accompany membership of a customs union are termed static effects. These are mainly felt in the short term. In the longer term, dynamic effects can be expected to emerge from this integration. These include the following:

- Firms within the customs union benefit from selling in a larger 'domestic' market. They can be expected to gain the advantages of internal and external economies of scale, enhancing their price competitiveness.
- Research and development may increase as greater competition drives businesses to innovate in order to gain advantages over rivals. Reduced costs per unit may allow firms to fund more expensive programmes of research and development.
- Relatively inefficient national monopolies may be exposed to competition from other more efficient producers based elsewhere in the customs union. This should benefit consumers in terms of lower prices and greater choice as well as permitting greater allocative efficiency.

The European Union

The European Union is a major example of economic and monetary integration. When Croatia joined the EU in July 2013, it became its 28th member state. Figure 51.8 shows the countries that are members of the EU in 2016. It also shows those countries that are members of a free trade area – the European Free Trade Area or EFTA – as well as countries that would like to become EU members.

The EU has grown steadily since it was established in 1957 as a customs union (then called the European Economic Community or EEC) when Belgium, France, Italy, Luxembourg, the Netherlands and West Germany signed the Treaty of Rome.

Figure 51.8 The 28 member states of the EU, 2016



The Single European Market

From the outset, the founder member states of the EEC had an ambition to integrate their economies further to establish a common or single market. This required the member states to agree the rules to permit the free movement of products (goods and services) as well as factors of production such as labour and capital.

The Single European Market (SEM) came into operation in January 1993. The SEM eliminated restrictions on the movement of products and resources as well as the removal of non-tariff barriers by agreeing common standards for products, employment laws, competition policies and government financial support to industry. In order to regulate the single market, the European Commission has authority over a range of areas of economic policy. Its regulations are most often passed down to national governments via directives that are then adopted into national law.

The SEM is based on the so-called ‘four freedoms’. These are the free movement throughout the 28 member states of:

- **goods** – firms can sell their products anywhere in the member states and consumers can buy wherever they wish;
- **services** – businesses providing services such as insurance, architecture and market research can sell their products in any member state;
- **capital** – capital can flow freely between the 28 member states and European citizens can use financial services from any member state;
- **people** – citizens of all the member states are free to live and work in any state and their professional qualifications should be recognised and accepted.

Key terms

Competition policy

refers to a range of laws and regulations used by governments to prevent the abuse of monopoly power by firms.

The European Commission

is the executive body of the EU with responsibility for proposing new legislation, monitoring the operation of EU treaties and overseeing routine operations.

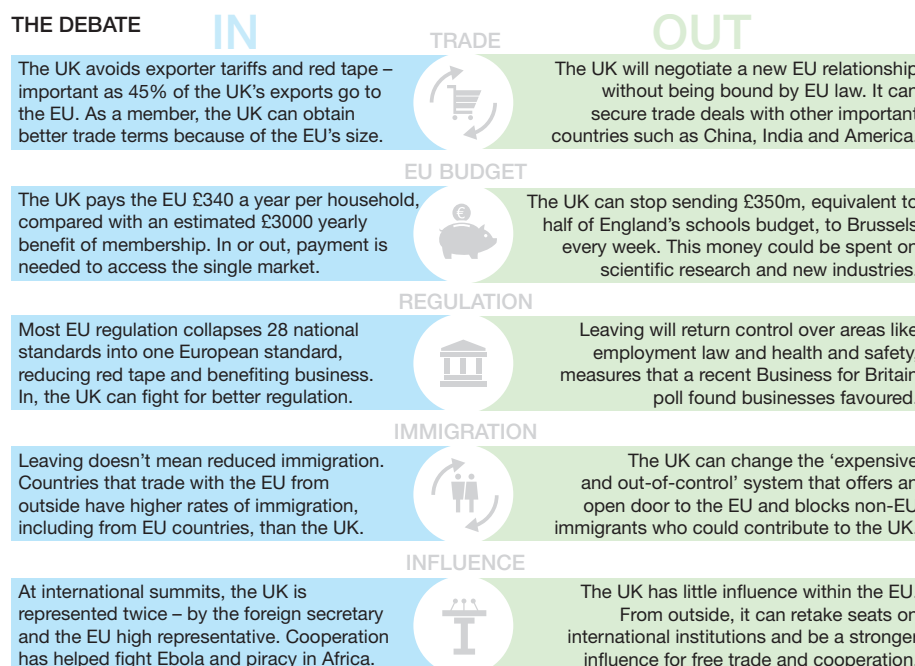
The SEM is based on three other key principles:

- the operation of relevant laws, regulations and administrations within all member states that are approximately the same;
- a common competition policy that applies across the EU, which is administered by the European Commission;
- the Common External Tariff.

The SEM is designed to create economies of scale, allow the establishment of Europe-wide commerce and enable faster rates of economic growth by applying consistent rules across the EU. However, the EU cannot be described as a truly single market because it does not have a unified taxation or welfare system, the single currency is not used by all members of the EU, and some countries have opt-outs from certain EU rules.

The consequences of EU membership for the UK

We saw earlier that there are advantages and disadvantages to economies arising from joining customs unions. Single markets require a greater degree of integration and also bring drawbacks as well as benefits. The UK has been a member of the EU since 1973 and its membership has been a source of controversy for much of the intervening period. The UK held a referendum on its continued membership of the EU in 1975 (which resulted in a strong majority in favour) and there will be another EU referendum held in June 2016.



Source: Adapted from *The Economist*, 19.10.15

Figure 51.9 summarises some of the main arguments relating to the UK's membership of the EU. We will supplement this by analysing the consequences of membership of the EU in relation to a number of key macroeconomic issues.

Figure 51.9 The main arguments for and against the UK's membership of the EU

Author tip

This chapter only provides a brief outline of the arguments for and against the UK's membership of the EU. However, there is an enormous amount of material available to you online and elsewhere on this topic.

1 The impact on the UK's GDP

Unsurprisingly, there are diverse views on the likely effects on the UK's GDP if the UK opts out of EU membership. One side of the argument is that the UK receives many economic benefits from its EU membership. A number of foreign multinationals have chosen to locate in the UK because it is within the EU's common external tariff (CET) and they can sell products throughout the EU from manufacturing bases in the UK. Decisions by multinationals such as Honda, Nissan and GlaxoSmithKline to locate in the UK have created thousands of jobs directly and indirectly, contributing significantly to UK production and to taxation revenue. If the UK were to withdraw from EU membership, it might not be able to operate within the CET and thus it would be a less attractive location for such multinational corporations. Their absence from the UK would have implications for a large supply chain and could reduce GDP significantly.

On the other hand, the UK makes a substantial net contribution to the EU and is subject to its regulations.

The UK made a net contribution to the EU budget of around £10.4 billion in 2015. Much of this money pays for the Common Agricultural Policy, while a sizeable amount goes towards the Structural Funds, which transfer money to poorer areas of the EU. Some economists argue that if the UK could retain and invest the funds that are sent to the EU, it would benefit and this could finance much needed improvements in the UK infrastructure and employee skill levels, provoking a strong multiplier effect.

Open Europe is an organisation that promotes ideas for economic and political reform of the European Union. Its most favourable estimate arising from the UK leaving the EU is that the UK's GDP could be 1.6% higher a year by 2030 than it would be if it remained a member. In contrast, the Centre for Economic Performance at the London School of Economics offers a reduction in GDP of between 6.3% and 9.5% as its worst-case scenario.

It is a major challenge to predict the effect of leaving the EU on the UK's GDP, since there are so many variables such as the extent to which the UK is able to negotiate some favourable trade deals, particularly a free trade deal with the EU. The impact on GDP also depends on the timescale involved.

2 The impact on employment in the UK

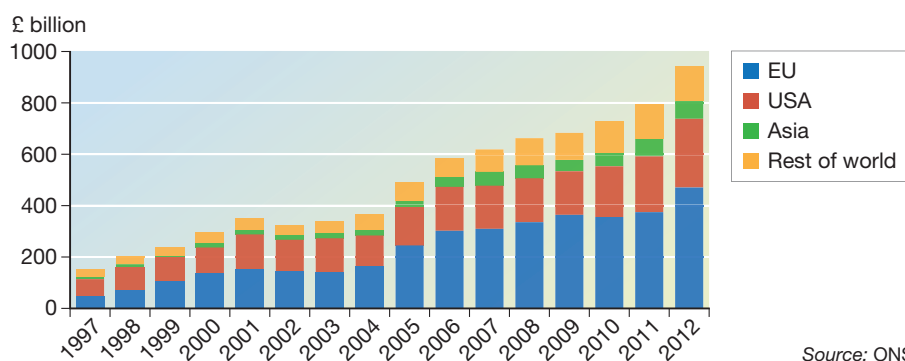
A number of studies conducted from 2000 onwards have suggested that anything from 2 million to 4 million jobs in the UK are dependent upon exports to the EU. There are worries that multinational manufacturers, especially from the car industry, might move operations to other parts of the EU. The financial services sector, which employs about 2.1 million people in the UK, also has concerns about the UK not being a member of the EU. The financial services industry depends heavily upon legislation relating to the Single European Market and some analysts fear the risks of an uncertain future on employment levels in this industry.

The Institute of Economic Affairs, which advocates the operation of free markets, has a very different view: ‘The UK labour market is incredibly dynamic, and would adapt quickly to changed relationships with the EU. Prior to the financial crisis [of 2007–08], the UK saw on average of 4 million jobs created and 3.7 million jobs lost each year... Indeed, the annual creation and destruction of jobs is almost on exactly the same scale as the estimated 3–4 million jobs that are associated with exports to EU nations’.

To some extent the level of employment would depend on the type of trade deals that could be negotiated (once again) and the labour market policies that the UK could pursue outside the EU. If the UK were to operate highly flexible labour markets alongside a low-tax environment for businesses, the impact on employment of leaving the EU would be more likely to be favourable, at least in the long term.

3 Other factors

The UK is an increasingly attractive destination for foreign direct investment (FDI), much of which comes from the EU (Figure 51.10). Foreign capital is a mobile factor of production and there might be an outflow (as well as fewer inflows) if the UK left the EU. Leaving the EU would almost certainly make the UK a less attractive destination for FDI even if the UK negotiated a favourable free trade area deal with the EU.



Source: ONS

Figure 51.10 Sources of UK inflows of foreign direct investment, 1997–2012

Countries such as Norway and Switzerland have performed very strongly in economic terms while not being members of the EU. Both economies have access to the SEM but are not bound by EU laws on agriculture, fisheries, justice and home affairs. This allows them greater flexibility and freedom from regulation. Opponents of the UK’s membership of the EU argue that the UK could be equally successful.

The role of the World Trade Organisation

The World Trade Organisation (WTO) was established in 1995 as a forum for resolving disputes between nations over trade and to promote and lead international negotiations with the aim of promoting greater freedom of trade. In January 2016, 162 countries were members of the WTO. The WTO has approximately 600 staff such as economists, statisticians and lawyers. They support the WTO’s members to help to achieve the resolution of trade conflicts and to ensure that the rules of international trade are applied.

The WTO exists with the aim of promoting international trade. It believes that its

Key term

The **World Trade Organisation** is an international organisation that promotes greater freedom of trade and resolves trade disputes between nations.

main function is, in its own words, 'to ensure that trade flows as smoothly, predictably and freely as possible'. It achieves this aim through two main activities.

Trade agreements

The WTO has negotiated a series of trade agreements. These negotiations can be lengthy and take place in 'rounds', each based in a particular location. The Doha Round is the latest round of trade negotiations among the WTO members. Its aim is to achieve major reform of the international trading system through the introduction of lower trade barriers and revised trade rules. It covers about 20 areas of trade and seeks to strengthen the position of developing nations in international trade and particularly to support the least developed economies. This round was officially launched in Doha, Qatar, in November 2001.

The WTO's trade agreements are conducted with the aim of reducing barriers to trade such as tariffs and quotas in order to help businesses engaged in international trade to carry out their activities more freely. Despite very slow progress, the Doha Round reached an agreement on 'trade facilitation', or measures to reduce trade costs by cutting red tape in customs procedures, in December 2013. Trade facilitation could cut global trade costs by an estimated 10%, raising annual global output by over £560 billion, with benefits flowing disproportionately to developing economies.

The WTO seeks to develop international trading rules that are fair through its negotiations. One of its key principles is that of most-favoured nation status. Under the WTO's rules, countries are prohibited from offering one country more favourable trade terms: all countries must be treated on an equal basis. Another is the so-called national treatment principle. This requires WTO members to avoid any form of discrimination against products imported to domestic markets. Once imports arrive in a domestic market, they must be treated in the same way as those supplied by local businesses. Since the WTO seeks to support developing countries, it frequently allows less developed economies additional time and flexibility to implement its agreements.

Trade agreements brokered by the WTO have to be agreed by all member states (which is why they can take so long) and to be ratified by the governments of the countries concerned. The agreements can relate to intellectual property as well as to goods and services.

Dispute settlement

The WTO's involvement in resolving trade quarrels is essential for enforcing the rules and therefore for ensuring that international trade flows are subject to minimal interruptions. Dispute settlement is vital for the success of a multilateral trading system and the stability of the global economy. Without a means of settling disputes, the rules-based system would be less effective because the rules could not be enforced.

Countries bring disputes to the WTO if they think their rights under the WTO's agreements are being infringed. Judgements by specially appointed independent experts are based on interpretations of the international trade agreements and individual countries' commitments. Trade disputes frequently relate to broken agreements. In December 2015, the USA made a complaint against China over unfair taxation of aircraft manufactured in that country.

Key term

Intellectual property

refers to the outcomes of human creativity including literary works, inventions as well as designs, symbols and names used by businesses.

Review questions

Total: 24 marks

- 1 Explain the difference between comparative advantage and absolute advantage.

(4 marks)

Use the following information to answer Questions 2 and 3. The world has two economies with the same quantity of resources for the production of two goods. On the assumption that they use half their resources to produce each good, the outcome is shown in the table below.

	A	B	World
Guns	2000	2000	4000
Butter	3000	4000	7000

- 2 Which of the following statements is true about this situation?

- A Trade would not be worthwhile as the opportunity cost ratios of producing the two products are the same in both countries
- B Country A has a comparative advantage in producing guns
- C Country B has a comparative advantage in producing guns
- D Country B has an absolute advantage in the production of both products

(1 mark)

- 3 If the two countries specialised completely according to their comparative advantage, which of the following statements would be true?

- A World output of guns and butter would both rise
- B World output of guns would rise and butter would fall
- C World output of guns would stay the same and output of butter would rise
- D World output of guns would rise and output of butter would stay the same

(1 mark)

- 4 Explain one benefit to domestic consumers that may arise from a country engaging in international trade.

(4 marks)

- 5 Which of the following statements about the UK's pattern of international trade is true?

- A In total, the UK exports more to the EU than to non-EU states
- B The UK has enjoyed a persistent trade surplus with non-EU states
- C The UK has a surplus on its trade in manufactures
- D Developing nations are becoming more important trading partners for the UK

(1 mark)

- 6 Economists use which of the following terms to refer to payments to exporters of goods and services to enhance the price competitiveness of the firms concerned?

- A Quotas
- B Export subsidies
- C Dumping
- D Tariffs

(1 mark)

- 7 Explain the difference between a customs union and a free trade area.

(3 marks)

- 8** What is the difference between trade creation and trade diversion? **(4 marks)**
- 9** Explain any two of the 'four freedoms' that underpin the Single European Market (SEM). **(4 marks)**
- 10** Which of the following statements about the UK's membership of the European Union is true?
- A** The UK makes a positive net contribution to the EU budget, which is generally rising over time
 - B** Flows of foreign direct investment (FDI) into the UK have declined steadily in recent years
 - C** The UK's trade deficit with the EU is becoming smaller over time
 - D** The UK's membership of the eurozone brings major disadvantages **(1 mark)**

The balance of payments

Key concepts from Year 1

Chapter 41 in the Year 1 companion textbook provided an introduction to the balance of payments. It outlined the overall structure of the UK's balance of payments and looked in some detail at the items that make up the current account. It considered the distinction between a deficit and a surplus on the current account of the balance of payments and the factors that determine whether a country records a current account deficit or surplus.

This chapter builds upon the material covered in the Year 1 book and examines the difference between the current, capital and financial accounts of the balance of payments. We analyse the consequences of investment flows between countries. The chapter also considers the policies that might be used to correct a balance of payments surplus or deficit, including expenditure-switching and expenditure-reducing policies as well as the effects that these policies might have on other macroeconomic objectives. We investigate the significance of balance of payments deficits or surpluses for an individual economy and the global implications of a major economy taking action to correct an imbalance on the current account of its balance of payments.

The structure of the balance of payments

The balance of payments records the financial and trading transactions of the UK with the rest of the world. These transactions are recorded in three main accounts in the balance of payments: the current account, the capital account and the financial account.

1 The current account

The current account comprises the trade in goods and services as well as flows of primary income and secondary income. The difference in the monetary value of inflows and outflows on these accounts is known as the current account balance.

There are four components of the current account of the balance of payments that include imports and exports of goods and services as well as certain financial flows.

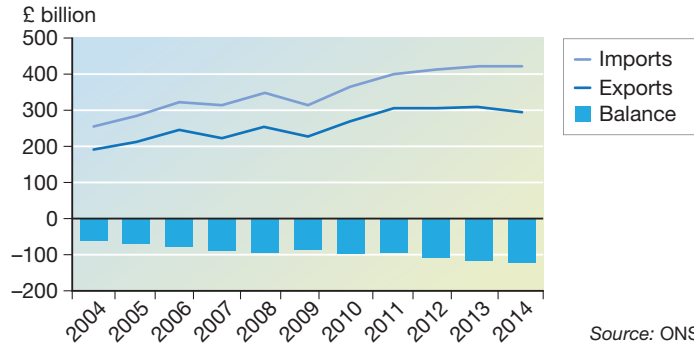
(a) Trade in goods

This component records the financial effects of UK residents buying imports of goods such as televisions. It also records the sale of goods overseas by UK firms, such as whisky from Scotland. This component of the current account is called 'visible' trade. The UK normally operates a deficit on this element of the balance of payments and, as shown in Figure 52.1, the deficit has increased steadily since 2004 to reach a record figure of £123.7 billion. The widening was due to a decrease in exports, with imports remaining relatively stable.

Key term

The balance of payments is a record of a country's trading and financial transactions with the rest of the world.

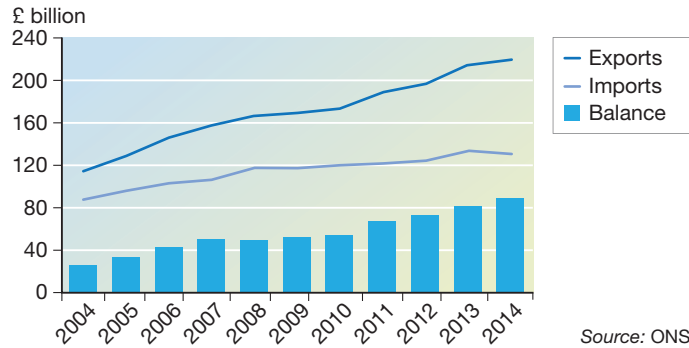
Figure 52.1 UK trade in goods, 2004–14



(b) Trade in services

The purchase of imports of services and the sale of services overseas are included within this element of the current account. Trade in services is termed ‘invisible’ trade.

Figure 52.2 UK trade in services, 2004–14

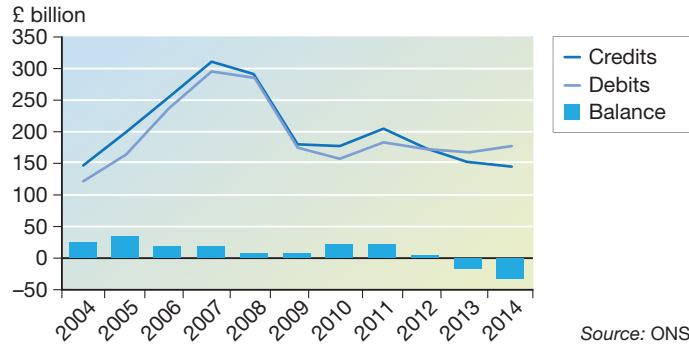


In contrast to trade in goods – or the visible balance – the UK has enjoyed a surplus on this element of the current account and one that has risen in recent years (Figure 52.2). The surplus in the trade in services balance for the UK increased by £8.1 billion, from £81.0 billion in 2013 to £89.1 billion in 2014.

(c) Investment income

Flows of investment are recorded on the financial account, but any resulting income, such as interest payments, profits or dividends, is investment income and recorded on this section of the current account. Thus the current account records the balance from investment income received and paid, which is known as primary income.

Figure 52.3 Net investment income for the UK, 2004–14



This deterioration in the UK's balance of payments on current account in recent years (Figure 52.3) can be partly attributed to the recent weakness in the primary income balance. This reached a record deficit in 2014: £33.1 billion compared with £16.8 billion in 2013. This deficit is due principally to a fall in UK residents' earnings from investment abroad, alongside broadly stable foreign resident earnings on their investments in the UK.

(d) Total current transfers

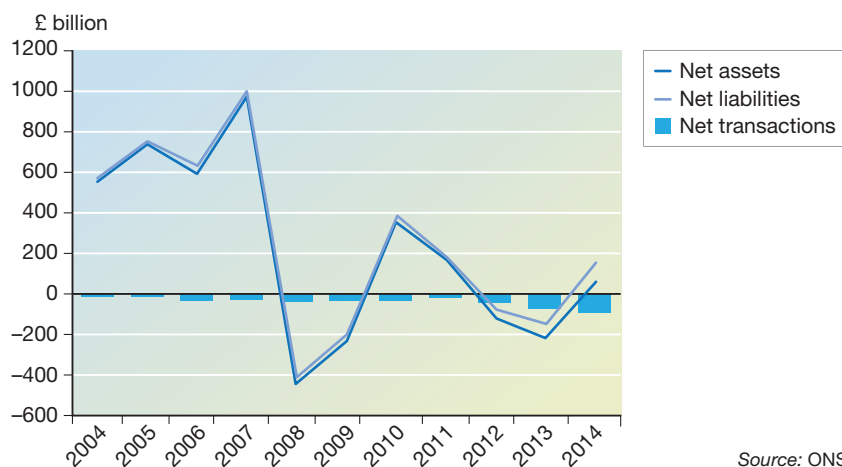
These are flows of money between countries such as international aid (disaster relief, for example) and gifts sent overseas or received from abroad. This is sometimes referred to as secondary income. The UK has suffered a deficit on this account since 2004, although the size of this deficit fell by £1.6 billion in 2014 to £25.2 billion.

2 The capital account

The capital account comprises two components: capital transfers between countries and the sale and purchase of non-productive and non-financial assets such as franchises, leases and copyrights. The capital account is relatively unimportant in terms of size and its impact on the UK balance of payments. In 2014, credits on this account (that is, capital flowing into the UK) were just over £2 billion and debits (or outflows) were significantly higher, giving a balance of -£0.625 billion. This is small compared to the size of the flows on the current and financial accounts.

3 The financial account

The financial account records transactions that result in a change of ownership of financial assets and liabilities between UK residents and non-residents. In 2014 the transactions recorded on the UK's financial account resulted in an inflow of currency. This is not immediately obvious from Figure 52.4. However, the graph does show that in 2014 foreigners increased their net liabilities (or net investment) into the UK by more than UK firms and households increased their overseas assets. Foreigners investing in the UK results in an inflow of funds, while UK investment overseas causes an outflow. Hence, in 2014 there was a net inflow of approximately £90 billion.



Source: ONS

Figure 52.4 The net transactions of the UK's financial account, 2004–14

The financial account of the balance of payments is made up of the three components outlined below.

(a) Net foreign direct investment

Foreign direct investment (FDI) records the transfer of ownership of UK or foreign businesses (i.e. productive assets) between residents in different countries. It also records money invested abroad for a new business venture. Net foreign direct investment is the balance between the purchase by UK firms of productive assets overseas less the purchase by foreign firms of these assets in the UK. This form of investment can involve a large investment in an enterprise, giving the investors some control of the firm and possibly a role in its day-to-day management. A UK firm buying a large stake in a manufacturing company in Poland is an example of outward direct investment, while an American bank investing in opening offices in London is an example of inward direct investment.

The UK's direct investment recorded a net inflow of £80 billion in 2014, as non-residents continued to invest in the UK, while UK residents disposed of more foreign assets than they bought. This net sale of assets will contribute to the UK's deteriorating primary income balance in the long term.

(b) Net portfolio investment

Portfolio investment refers to financial flows between economies arising from the purchase and sale of financial assets such as shares and bonds. Unlike direct investment discussed above, portfolio investment does not give the investors real control over the business. It is sometimes referred to as 'passive investment', allowing others to manage the enterprise. If someone in the UK buys shares in a Brazilian company, this represents an outflow of money from the UK in the portfolio investment section of the financial account. Similarly, if a Brazilian buys bonds issued by a UK company, the result will be an inflow of money into the UK, recorded in the same section of the financial account of the balance of payments.

The flows of income, for example dividends and interest, arising from portfolio earnings are recorded as primary income in the investment income section of the current account. Earnings on portfolio investments in other economies represent an inflow into the UK.

(c) Other investment

This mainly comprises short-term international financial flows such as movements of hot money by speculators. This section of the financial account can be quite large and the flows can be volatile: in 2007 it recorded a net inflow of over £600 billion; in the following year there was a net outflow of a similar size. In 2013 and 2014 this component recorded a surplus.

Key terms

Foreign direct investment (FDI) records the transfer of ownership of UK or foreign businesses (i.e. productive assets) between residents in different countries.

Portfolio investment refers to financial flows between economies arising from the purchase and sale of financial assets such as shares and bonds.

Hot money refers to funds that flow between financial markets as investors seek to earn the highest possible returns.

(d) Flows of reserve assets

Reserve assets are foreign financial assets that are managed by the UK authorities. The UK's holdings of reserve assets comprise gold and foreign currency assets as well as Special Drawing Rights (an accounting currency created by the International Monetary Fund). The UK's official reserves are held in the Exchange Equalisation Account, which is administered by the Treasury but is managed on a day-to-day basis by the Bank of England. In March 2015 the total assets held in the EEA were valued at more than £85.7 billion.

A country's reserve assets are used to finance deficits on the balance of payments. If a country suffers a deficit on its current account, it may be that flows on its capital and financial accounts will offset this. If this is not the case, then the sale by the Bank of England (or any other central bank) of some of its reserve assets will produce an inflow of pounds that will be recorded on the balance of payments.

The balance of payments always balances

A country's balance of payments will always balance. If a country runs a deficit on its current account, this will result in net outflows of the currency exceeding inflows. The funds to finance this deficit could be gained as a result of foreign firms and households investing in the UK, or by UK firms and households selling overseas assets, and money flowing into the UK as a consequence. If an imbalance remains, the authorities can use reserve assets to offset this, as outlined in the previous section.

Since the UK's capital account is relatively small in comparison to the other accounts on the balance of payments, it is broadly true to say that the current and financial accounts should offset one another. For example, if the current account runs a deficit, the financial account should record a surplus to offset it. We can see from Table 52.1 that the UK ran a deficit on its current account in both 2013 and 2014. However, this was covered to a large extent in both years by a surplus (i.e. a net inflow of money) on the financial account once adjustments had been made for the capital account and the balancing item.

The balancing item

Economic theory tells us that the balance of payments must balance. Any overall imbalance must be covered by an adjustment in the level of reserve assets held by the central bank. However, the collection of data when compiling the balance of payments is a huge and complex task and errors and omissions are certain to occur.

Errors and omissions may take place because of the inadvertent submission of incorrect data by firms or because some illegal activities are not recorded. Hence a balancing item is included to ensure that the accounts balance and in recognition that they include both errors and omissions. Over time, the balance of payments data for a particular period are likely to be updated as more information becomes available.

Number crunching

Use the data in Table 52.1 to complete the following calculation. Assume that the secondary income balance in 2014 was subsequently revised to be -£26,000 million. What would be the new figure for the balancing item in order to make the balance of payments balance?

Component	2013 (£ million)	2014 (£ million)	Comments
Trade in goods	-112,599	-119,605	The UK's deficit on visible trade is widening, despite the government's aim to increase output of manufactures.
Trade in services	78,866	85,863	The invisibles surplus is rising, but it is insufficient to offset the deficit on trade in goods.
Primary income balance	-15,760	-38,754	This widening deficit is the result of changes in patterns of investment and is a concern for many economists.
Secondary income balance	-27,162	-25,424	A deficit is recorded here as well, but it is stable over these two years.
Current account balance	-76,655	-97,920	This current account deficit is 5.5% of the UK's GDP and the largest annual deficit since modern records began in 1948.
Capital account balance	737	625	Surpluses are recorded here, but these are tiny in relation to the balances on the current and financial accounts.
Direct investment	40,078	80,143	There is an increasing inflow here as overseas net investment in the UK exceeds the UK's net investment abroad, creating an inflow of funds.
Portfolio investment	31,408	102,363	The UK is becoming an increasingly popular destination for overseas investors buying financial assets, such as shares and Treasury bills.
Other investment	-1098	-81,938	This component of the financial account is highly volatile and includes hot money flows.
Reserve assets	-4961	-7113	There are adjustments to reserve assets to fund the current account deficits.
Financial account balance	65,427	93,455	This offsets the current account deficit, once small adjustments are made for the capital account balance and the balancing item.
Balancing item	10,491	3840	This item allows for errors and omissions in the balance of payments accounts.
Balance of payments	0	0	The balance of payments always balances.

Source: ONS – The Pink Book, 2015

Table 52.1 Selected components from the UK's balance of payments for 2013 and 2014

The consequences of investment flows between countries

We saw in the previous section that flows of investment between different countries can be classified as direct investment and portfolio investment. Portfolio investment entails investment in financial assets, which are relatively easily traded and may be less permanent, and does not represent a controlling stake in an enterprise. Other investment flows include hot money and can be highly volatile.

In contrast, foreign direct investment (FDI) creates long-term flows and it has increased over recent years as markets have become more accessible due to globalisation. For example, OECD data show that the global total of FDI rose from \$57 billion in 1982 to an estimated \$1.9 trillion in 2007. It was \$1.2 trillion in 2014.

Benefits

1 Rising levels of income and employment

Positive net inflows of FDI can promote rising income and employment levels in an economy. For example, FDI is helping to generate economic growth in Africa: it reached an estimated \$55 billion in 2015, which was 20% higher than in 2010. FDI into Africa has a strong focus on technology, retail and business services and the potential to create substantial numbers of jobs as well as the basis for increased volumes of trade.

The UK attracts more FDI than any other European country. The government has estimated that FDI received during 2014–15 resulted in over 100,000 new and safeguarded jobs in the UK.

2 Potential to bring about supply-side improvements

FDI into many countries, and especially less developed ones, can result in supply-side improvements. The investment may increase the productivity of labour forces by improving skills, or the use of new production techniques based on technology can increase the economy's productive capacity. Taxes raised by governments from multinational corporations can also be used to develop and improve the economy's infrastructure, which can raise the productive potential of the economy further.

In 2015 the Prime Minister of India, Narendra Modi, announced the easing of restrictions on FDI into India with the aim of encouraging greater investment in technology, especially in the manufacturing sector, to increase the productive potential of the economy. The move is intended to replace imports with domestic production.

Drawbacks

1 Multinational corporations may exploit some countries

Multinational corporations invest directly overseas to take advantage of lower costs of production. This may be because factors of production such as labour or land are cheaper in some developing countries. For example, a high proportion of global manufacturing takes place in Asia and in Eastern Europe where wage rates are significantly lower. However, this concentration of manufacturing can result in the exploitation of non-renewable resources and heavy negative externalities, notably in the form of damage to the environment. For example, the multinational oil company Shell has received much criticism for the environmental damage it has caused as part of its operations in Nigeria as well as for allegedly covering up the extent of the damage inflicted. In addition, multinational corporations have been accused of paying unacceptably low wage rates, sometimes to young children.

2 Dangers of an interconnected global economy

Portfolio investment plays an important role in connecting the world's economies as it links investors in one country with economic activity in others. Investment by multinationals in productive capacity overseas brings close links too. The downside of this interconnectedness is that crises in one part of the global economy spread quickly to others. At the time of writing, the Chinese economy is experiencing a slowdown. Its deteriorating performance is having a negative impact on other economies, partly because they have invested heavily in China. Business confidence is low and global stock markets have suffered falls.

3 Implications for the current account of the balance of payments

An increase in inward investment will strengthen the financial account of the balance of payments because it creates inflows of capital. However, there is also a longer-term outflow associated with the payment of profits, interest and dividends on these investments. This can result in a country acquiring a negative balance on its primary income. This is the case with the UK at the time of writing and one of the causes of a long-term deterioration in its current account balance.

REALWORLD ECONOMICS 52.1

FDI flows fall in 2014 but rise strongly over the long term

Net inflows of foreign direct investment (FDI) fell by 16% globally in 2014, to \$1.2 trillion (Figure 52.5), as a result of economic weaknesses and uncertainty about global economic policies. FDI flows to developed

countries dropped by 28%, to \$499 billion, their lowest level in a decade. Inflows to the United States fell by 60%, to \$92 billion, while flows to Europe fell by 11%, to \$289 billion, at a time when many of the continent's economies

are performing poorly.

Due largely to Asia, developing economies' share of FDI inflows is rising steadily: it accounts for 55% of the global total. Services continue to attract the bulk of investment, accounting for 63% of the global FDI – more than twice the share of manufacturing.

Source: Adapted from *The Economist*, 27.6.15

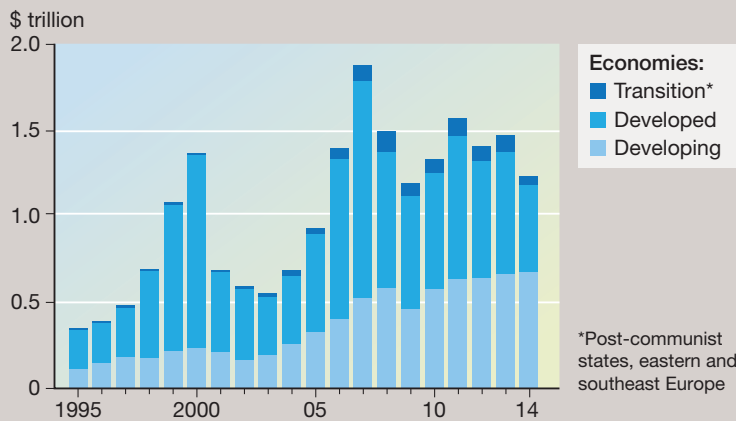


Figure 52.5 Global FDI (inflows, \$ trillion), 1995–2014

Exercises Total: 13 marks

- 1 Why might developing economies in Asia be pleased to be receiving a rising share of global FDI? (7 marks)
- 2 What might be the consequences for European economies of reductions in inflows of FDI into their economies? (6 marks)

Policies to correct a balance of payments deficit or surplus

As a result of international trade, it is always likely that some countries will suffer deficits and others surpluses on the current accounts of their balance of payments. Globalisation has increased the volume of trade and capital flows and thus made it possible for some countries to have larger imbalances on their balance of payments, before official financing by central banks. An economy's balance of payments cannot be looked at in isolation: one economy's surplus will represent deficits for one or more other economies.

Most attention is normally focused on the current account of the balance of payments.

Current account deficits and surpluses

The UK has not experienced a surplus on its current account since the 1980s. However, other countries such as Germany regularly record surpluses – it last experienced a deficit in 2003. Surpluses can create difficulties for economies because they tend to increase the value of the nation's currency and can lead to global financial instability as other countries implement policies to eradicate their matching current account deficits. They also tend to stoke inflationary pressures.

However, governments are more likely to take action when facing a current account deficit, especially if some of the criteria below are met:

- it represents a relatively high proportion of the nation's GDP. A figure over 4% is likely to be a cause for concern. Research has shown that governments tend to implement policies to reduce current account deficits when they reach about 5% of GDP;
- if the deficit is persistent. Many economies experience a series of deficits and surpluses over time. If these roughly cancel one another out, a problem does not exist;
- if the country is not able to finance this deficit by importing capital in the form of net inflows of investment recorded on the financial account of the balance of payments;
- if the central bank of the country concerned has relatively low levels of gold and foreign exchange with which to purchase its own currency and create compensating inflows.

A current account deficit is also regarded as more of a problem under a fixed exchange rate system where the value of the currency cannot adjust by falling in value to eliminate the deficit. With a floating rate system, where market forces determine the value of the currency, a deficit will lead to more of a currency being sold on foreign exchange markets than is purchased. The currency will fall in value as a consequence, making exports more price competitive and imports less so, helping to correct the deficit. However, the impact of a trade imbalance on the value of a currency is relatively small. Many currency transactions are conducted for reasons other than trade. About 90% of transactions in sterling on foreign exchange markets are for the purposes of investment or speculation.

It is simple to say that a current account deficit is bad for an economy, but this does depend upon the circumstances. It may be bad if it is due to a surplus of imports over exports and this is the result of a lack of competitiveness of domestic firms and perhaps a consumption binge by households at the expense of saving. However, this may not be the case if an economy is growing quickly and attracting investment that may lead to higher imports of producer goods, which will allow exports to increase in the future. In this case the deficit may not represent a long-term problem.

A country operating a persistent current account surplus is likely to come under pressure from other trading nations to implement policies to reduce or eliminate it. If one country has a surplus, it means others must have equivalent deficits. Furthermore, current account surpluses can create inflationary pressures because high export sales are likely to boost aggregate demand and initiate a multiplier effect.

Expenditure-reducing policies to correct imbalances

A government will enact expenditure-reducing policies to deflate the economy and to reduce consumer spending. By reducing consumer spending, the government would expect to reduce the level of expenditure on imports of goods and services and thereby to improve the balance of payments on current account. Export sales should not be affected by these policies. Consumers in the UK are generally thought to have a high marginal propensity to import. This means that expenditure-reducing policies are likely to be effective as a reduction in income and expenditure will result in a substantial fall in imports.

Key terms

A **current account surplus** exists when inflows of currency recorded on the current account exceed outflows over some period of time.

A **current account deficit** exists when inflows of currency recorded on the current account are smaller than outflows over some period of time.

The government can use a range of deflationary policies to achieve the desired fall in consumer expenditure on imports. It could implement deflationary fiscal policies by reducing government expenditure or increasing taxes, or a combination of the two. Alternatively it could use monetary policy and raise interest rates to reduce the level of consumer spending by making loans more expensive and saving more attractive.

An assessment of expenditure-reducing policies

Expenditure-reducing policies have significant side effects, many of which can be undesirable for governments in terms of achieving their macroeconomic objectives. Reducing consumer expenditure through deflationary policies to correct a deficit will reduce the level of aggregate demand and hence the level of economic activity. As a result, there will be a fall in real GDP. The consequences are likely to be a lower rate of employment (or a higher rate of unemployment) as real GDP declines. The average standard of living is likely to decline as well. It may be that the costs of correcting a balance of payments deficit in this way are too great.

Expenditure-increasing policies to correct a current account surplus

A government could use reflationary policies, such as reducing taxes or the rate of interest, to help to overcome a persistent surplus on the balance of payments. This would increase consumer expenditure and draw in more imports, helping to correct the surplus.

The correction of a balance of payment surplus through the use of reflationary policies would increase GDP and the rate of employment but also probably generate inflationary pressures in the economy, especially if it is operating near to full capacity.

Expenditure-switching policies to correct imbalances

These policies operate with the aim of persuading consumers to switch demand away from imported goods and services to those produced within the domestic economy. If this switch can be achieved to a significant extent, it will reduce the level of expenditure on imports, helping to improve the current account balance. The government can encourage consumers to change their expenditure patterns if it can increase the competitiveness of domestic industries. Increasing their price competitiveness might be a particularly effective option if demand for the products in question is price elastic.

There are two main policies available to a government with the aim of achieving expenditure switching.

Key terms

Expenditure-switching policies operate with the aim of diverting consumers' expenditure away from imports and towards domestically produced goods and services.

Expenditure-reducing policies are those designed to reduce aggregate demand in the economy and, with it, consumer spending on imports.

A **deflationary policy** is an action by governments or other authorities implemented with the intention of reducing aggregate demand.

The **marginal propensity to import** is the proportion of a rise in income that is spent on imported products.

Import controls

In Chapter 51 we saw that governments can use a range of methods of protection to deter its consumers and firms from purchasing imports. These methods of protection include tariffs and quotas. Such forms of trade protection have the potential to be effective. The implementation of tariffs will increase the price of imports relative to domestically produced products, and consumers may be expected to change spending patterns, especially if demand is price elastic. Quotas will restrict the quantity of imported products, and their scarcity may increase their price on the domestic market. The effect will be to reduce the volume of imports and, *ceteris paribus*, improve the current account balance.

An assessment of import controls

Import controls may seem like a quick fix for a current account deficit, but most governments are extremely reluctant to use them. There is the fear of retaliation. If one country imposes, say, a 10% tariff on imports from one or more other countries, those countries are likely to impose their own tariffs, probably at a similar rate. They will do this to protect their domestic industries from what they consider to be unfair competition.

The UK is unable to impose tariffs or quotas on imports from the 27 other members of the European Union because it is part of a customs union. Furthermore, the World Trade Organisation operates to reduce barriers to global trade, and the UK (as well as most other countries) is subject to many of its agreements, making it harder to impose import protection.

Import controls are unlikely to have positive effects on a country's current account balance in the longer term and may impact negatively on its GDP and rate of employment, especially if other countries retaliate.

Reducing the value of the currency

A government may decide to fix its exchange rate (i.e. to maintain a set value between its currency and other currencies). An alternative approach is to allow market forces to determine the value of the currency, which will change constantly as a result. This latter approach is termed a 'floating exchange rate'. In either event, it is possible for the government to engineer a fall in the value of the currency. This is called devaluation when the exchange rate is fixed and depreciation when it is floating. (We consider exchange rate systems in detail in Chapter 53.)

Devaluation and depreciation have the same effects on the price of exports and imports. For simplicity, we will refer to a fall in the value of a currency as depreciation. Depreciation of a currency might be expected to improve an economy's current account balance for the following two reasons:

- **Exports fall in price:** Depreciation of a currency means that fewer units of another currency are required to buy it. As a result, the price of exports in overseas markets will fall. Depending on price elasticity, the sales volume will rise.
- **Imports rise in price:** Imports become more expensive as more units of the domestic currency are needed to buy a unit of a foreign currency. Higher-priced imports are likely to be less attractive to domestic households and firms, which may result in a switch to home-produced products. Once again the overall effect will depend upon the price elasticity of demand for the various categories of imports.

Key terms

The **Marshall–Lerner condition** states that currency devaluation will only lead to an improvement in the balance of trade element of the current account if the sum of the elasticities of demand for imports and exports is greater than one.

The **balance of trade** is a part of the current account balance and measures the earnings from exports minus expenditure on imports.

An assessment of depreciation

There are two factors that will impinge upon the effectiveness of depreciation as a means of overcoming a current account deficit: the price elasticity of demand for exports and imports, and the timescale over which the effect is judged.

Depreciation reduces the price of exports and increases the price of imports. This change alone would further weaken the current account balance as the country is selling its products for less and paying more for imports. Therefore, in order to improve a country's balance of payments, depreciation must result in positive adjustments to the volumes of imports and exports. Depreciation will have a greater impact on the current account balance if demand for the imports and exports are sensitive to price changes – in other words, if demand is price elastic.

The English economist Alfred Marshall and a Romanian-born colleague Abba Lerner devised a rule to determine whether a depreciation of a currency would be effective in terms of improving an economy's current account balance. This is known as the Marshall–Lerner condition. This condition states that currency devaluation will only lead to an improvement in the balance of trade element of the current account if the sum of the elasticities of demand for imports and exports is greater than one. This condition holds true because if demand is sufficiently price elastic, then volumes will adjust by a sufficient amount to more than offset the adverse changes in price. If the Marshall–Lerner condition is not met (i.e. if the sum of the two elasticities is less than one), depreciation will worsen the balance of trade, and hence the current account, *ceteris paribus*.

In the short term, devaluation may weaken the balance of trade and the current account. This gives rise to what is known as 'the J-curve effect'. In the short term, demand for both exports and imports might be fairly inelastic because contracts may have been signed on trade deals for a period of time. Thus, during this time the volume of trade may not change by very much, but the price changes will work against the country that has depreciated its currency. It will buy a similar volume of imports but will pay a higher price for these as its currency has depreciated. At the same time, it will find that its volume of export sales changes little and it receives the same sterling price.

Taken together, these effects mean that the country's balance of trade may weaken in the short term, before it starts to improve over time as volumes of trade adjust to the new prices. This validates the Marshall–Lerner condition because in the short term, when the price elasticities are less than one, the balance of trade worsens. In the longer term, when they exceed one, the balance improves. This is illustrated in Figure 52.6.

Depreciation (or devaluation) may affect an economy's ability to achieve its macroeconomic objectives. Devaluation makes imports more expensive and, for an open trading economy such as the UK, a substantial depreciation in the currency could result in increasing inflationary pressures. Depreciation would cause cost-push inflationary pressures to emerge as imported fuels, materials, components and finished goods all rise in price.

The depreciation may also only have a positive impact on employment and GDP in the medium term. In the short term, as shown in Figure 52.6, sales of exports may decline and imports rise. This is likely to reduce aggregate demand. In the longer term

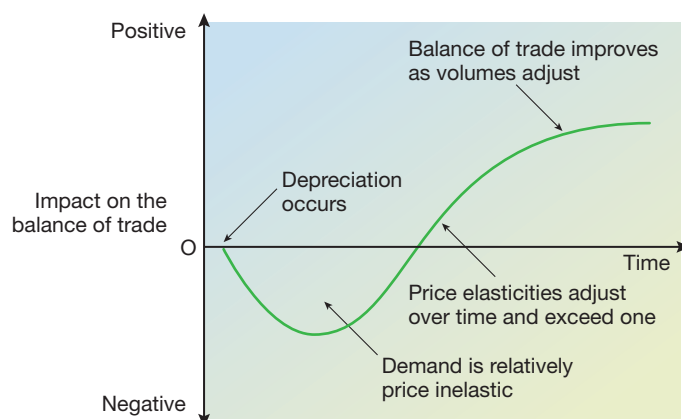


Figure 52.6 The J-curve effect following depreciation

the competitive advantage gained by domestic firms from the depreciation is likely to be eroded (at least to some extent) as import prices rise, increasing production costs. Thus the government may not be able to gain the improvements in GDP and employment that it hoped would accompany the improvement in the current account balance.

Expenditure-switching policies and current account surpluses

The analysis of expenditure-switching policies in the previous section was based on dealing with a current account deficit. An economy that is experiencing a surplus on its current account may opt to use expenditure-switching policies. However, it would need to reverse these policies. This could entail reducing or removing any existing forms of trade protection such as tariffs or quotas. This is unlikely to provoke any opposition and would most likely have the support of the World Trade Organisation.

An increase in the value of a currency would be required to eliminate a surplus on the current account of the balance of payments. For a fixed exchange rate system this is termed a revaluation, or an appreciation under a floating system of exchange rates. In either case the effect would be to increase the price of exports and to reduce the price of imports. This would provoke a reverse J-curve effect because the balance of trade element of the current account would improve in the short term due to price inelastic demand for imports and exports. In the longer term the volumes would adjust – higher imports and lower exports, reducing the surplus.

The implications of major economies with global imbalances

Table 52.2 reveals that some countries run relatively large current account deficits and surpluses, but that these vary enormously in size. The table expresses the deficits and surpluses as a percentage of GDP to make comparisons simpler. Kuwait has an eye-catching surplus as a result of being a major exporter of oil, and Germany's surplus is based on its prowess as a manufacturer. The UK's deficit is significant in global terms and rising quickly. In dollar terms the USA's deficit is large but, because of the huge size of the American economy, it represents a comparatively small proportion of the country's GDP. The deficit recorded by the Democratic Republic of the Congo, a low-income developing economy in Africa, is large. Although the country is richly endowed in mineral wealth, its trade deficit had occurred because of a sharp decline in world commodity prices and poor management of its economy.

Table 52.2 *Current account balances as a percentage of GDP for selected countries, 2013 and 2014*

Country	2013	2014
Brazil	-3.4	-4.4
China	1.9	2.1
Congo, Democratic Republic of	-10.4	-7.4
Denmark	7.2	6.3
France	-1.4	-1.0
Germany	6.5	7.6
India	-2.6	-1.5
Kuwait	40.9	32.5
Netherlands	11.0	10.8
Russia	1.7	3.1
Sweden	9.4	7.8
UK	-4.4	-5.9
USA	-2.2	-2.2

Source: World Bank

The global implications of eliminating a deficit

There could be significant consequences for the global economy if a major economy such as the USA decided to take corrective action to reduce its trade deficit. This would depend upon the policies it opted to adopt to eliminate the deficit and the timescale over which it sought to achieve a balance on its current account.

The imposition of import controls could have considerable negative consequences for the global economy. The USA is the world's largest economy in terms of nominal GDP and a decision to restrict imports could have huge consequences for exporters in other nations. For example, the value of China's exports to the USA was over \$440 billion in 2014 and a significant decline in these could have a strong deflationary effect on the Chinese economy. Globalisation has made the world's economies much more connected. Thus, a slowdown in the large Chinese economy would be felt in those economies that meet its huge demands for minerals and raw materials such as Australia and Brazil. The risk would be a slowdown in global growth. American import controls may result in retaliation by other countries, provoking a further reduction in trade and a more pronounced slowing of global growth rates.

The effects of the US government implementing deflationary policies to remove the current account deficit would be similar. Studies have suggested that this would be likely to slow the growth rates of the economy sharply and have a strong deflationary effect on the wider global economy. It could push some weaker economies into recession. The risks of such negative consequences would be much increased if the US administration were to introduce strongly deflationary policies with the aim of eliminating the current account deficit quickly. Research in 2001 by the United Nations indicated that a decision by the American government to reverse its current account deficit quickly could result in world GDP declining by 1.7%. It should, however, be noted that the USA's current account deficit was nearly 6% of GDP at the time, so the effects of a reversal would be greater than would be the case in 2016.

The global implications of eliminating a surplus

Germany is the world's fourth largest economy. In 2014 its GDP was \$3853 billion

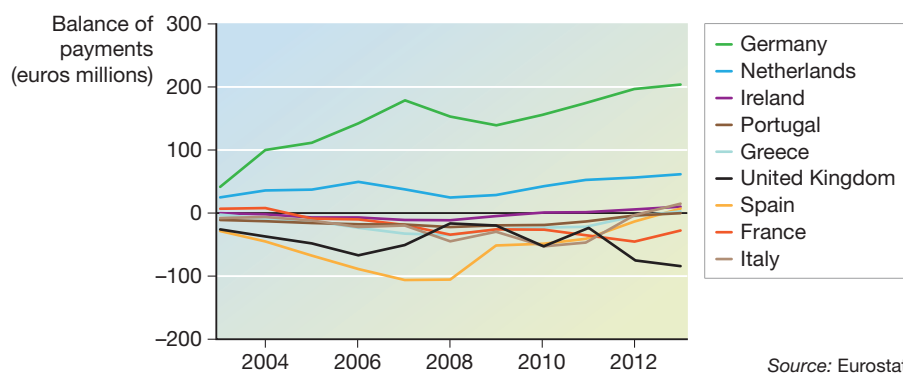


Figure 52.7 The current account balance of a selection of European countries, 2004–13

and its current account surplus was 7.6% of its GDP, or approximately \$293 billion. This creates similar-sized deficits in its trading partners.

A decision by the German government to eliminate its surplus would have significant consequences for the global economy because of the size of the surplus. Germany could implement a range of policies designed to draw in imports and to reduce or eliminate its trade surplus. Such policies could include reflating its economy through tax cuts or by investing in its infrastructure. These would boost consumption and aggregate demand and draw in a greater volume of imports. The policies would tend to increase the volume of world trade and, through the export multiplier in other countries, boost the GDPs of those countries as well as global GDP.

The elimination of the large German trade surplus would bring other benefits too. Germany is a member of the eurozone and is one of 19 countries that use the euro as their national currency. It has had a large and growing trade surplus since 2004 (Figure 52.7), creating problems for other members. Chronic surpluses are a way of taking aggregate demand from other countries. German surpluses export unemployment to other countries in the eurozone as well as slowing their rates of economic growth. As a consequence, any German moves to reduce or eliminate its trade surplus would be likely to improve the macroeconomic performance of other member countries of the eurozone.

Review questions

Total: 38 marks

- Which of the following transactions would be recorded on the current account of the UK's balance of payments?
 - A UK citizen purchasing shares in a French company
 - An Italian company establishing a factory in Newcastle
 - A Swiss company paying dividends to its UK shareholders
 - An American hedge fund purchasing a controlling interest in a UK firm

(1 mark)
- Which of the following transactions would be recorded on the financial account of the UK's balance of payments?
 - A UK citizen buying corporate bonds issued in America by an American company
 - The UK government transferring capital to Argentina
 - A payment from the UK to Africa to relieve famine
 - The purchase of UK insurance by an Algerian company

(1 mark)

- 3** Explain, with the aid of examples, the difference between foreign direct investment and foreign portfolio investment. **(6 marks)**
- 4** Explain why it is necessary to include a balancing item in the balance of payments. **(6 marks)**
- 5** Which of the following statements relating to the UK's balance of payments is true?
(i) The volume of inflows and outflows on the UK's current account is normally greater than those on its capital account.
(ii) Official financing transactions, using reserve assets, are recorded on the capital account.
(iii) In recent years the UK has tended to record a surplus on its financial account.
A (i) only
B (i) and (ii)
C (i), (ii) and (iii)
D (i) and (iii) **(1 mark)**
- 6** Explain how FDI inflows may help an economy to achieve supply-side improvements. **(5 marks)**
- 7** Explain, with the aid of examples, the difference between expenditure-reducing and expenditure-switching policies. **(6 marks)**
- 8** Which of the following is an expenditure-reducing policy?
A The imposition of a quota
B Increases in direct taxation
C A decrease in the rate of interest
D A 'Buy-British' campaign **(1 mark)**
- 9** What is the Marshall–Lerner condition? **(3 marks)**
- 10** Explain what is meant by the 'J-curve effect'. **(8 marks)**

Exchange rate systems

Key concepts from Year 1

We introduced the concept of exchange rates in Chapter 35 of the Year 1 companion textbook when discussing net exports as a component of aggregate demand. At that point we simply explained that an exchange rate is the price of a currency expressed in terms of another, and we analysed how changes in the exchange rate could influence the prices of exports and imports. In subsequent chapters we explained the reason why a direct relationship exists between the bank rate of interest in the UK and the exchange rate of the pound and how changes in the exchange rate can influence a government's ability to attain its macroeconomic objectives.

In this chapter we distinguish between fixed and floating exchange rate systems before explaining how exchange rates are determined within freely floating exchange rate systems. We investigate the ways in which governments can intervene in markets to influence the exchange rate. We then discuss the advantages and disadvantages of fixed and floating exchange rate systems and, in the final section of the chapter, we consider the advantages and disadvantages for a country of joining a currency union.

Exchange rates

An exchange rate is the price of one currency, expressed in terms of another. A currency can thus have many different exchange rates against the other currencies of its trading partners. On 21 January 2016, the UK's currency, the pound sterling, traded at the following values:

- Japanese yen: £1 = ¥165.09
- Chinese yuan: £1 = ¥9.29
- American dollar: £1 = \$1.41
- The EU's euro: £1 = €1.30
- Indonesian Rupiah: £1 = Rp19,648

Key terms

The **exchange rate** is the price of one currency expressed in terms of another – for example, £1 = \$1.45 or €1.28.

A **fixed exchange rate** system exists when a currency is maintained at a stable value in relation to other currencies.

A **floating exchange rate** operates when a currency's value is determined in the free market through the interaction of the forces of supply and demand.

A **trade-weighted index** shows how a single currency changes in value over time against a 'basket' of other currencies.

Number crunching

Assume that the value of the Japanese yen has risen in value against the pound from ¥165.00 to £1 = ¥150.00. What would be the impact of this rise on:

- the selling price of a car exported from the UK to Japan priced at £20,000;
- a television imported into the UK from Japan priced at ¥190,000?

Each of these exchange rates can change because of factors relating to the pound or those linked to the other currencies. They are examples of bilateral exchange rates (i.e. between just two currencies). The pound may fall against some currencies and rise against others on any day's trading on foreign exchange markets. A multilateral rate is the value of a currency against more than one other currency. In order to measure the average movement in the value of the pound, or any other currency, economists use a trade-weighted index.

Author tip

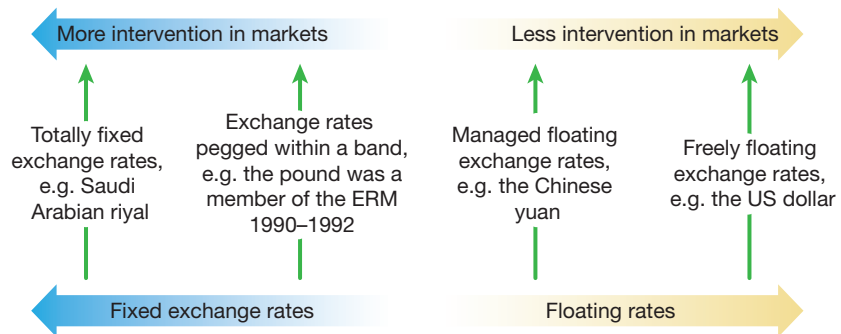
If you are unsure about weighted indices, you should reread Chapter 33 in the Year 1 textbook.

A **trade-weighted index** shows how a single currency, for example the pound, changes in value over time against a 'basket' of other currencies. A basket of currencies is a selection of currencies that are combined in the form of an index number to produce one value against which another currency is measured. The choice of currencies to be included in the basket normally reflects their importance in international trade for the country concerned.

Exchange rate systems

The exchange rate systems by which a country's monetary authorities allow the value of its currency to be determined can be divided into two broad categories: fixed and floating exchange rates. However, within each of these two categories there are slightly different versions in operation in economies throughout the world, as summarised in Figure 53.1.

Figure 53.1 The spectrum of exchange rate systems



Fixed exchange rate systems

Fixed exchange rate systems can involve a range of scenarios. In some circumstances, a country's currency may be totally fixed (or pegged) against other currencies. In

these circumstances the government would have to intervene in the foreign exchange markets to buy and sell its currency to maintain the agreed par value. For example, the Saudi Arabian riyal is fixed against the American dollar. It is fixed at a rate of \$1 = 3.75 riyals and has been since October 2003.

An economy's currency can be 'pegged within a band'. A country will fix its currency against a major global currency such as the dollar or euro (or possibly against a basket of currencies) with a central exchange rate. It may choose the currency of a major trading partner. However, the currency is allowed to move within a narrow band (possibly +1/-1%) or a wide band (up to +20%/-20%). The central peg and range may move up or down over time, normally depending on differences in inflation rates between countries and other changes in macroeconomic fundamentals.

Between 1990 and 1992 the UK was a member of the European Exchange Rate Mechanism (ERM). The value of the pound was fixed against the German Deutschmark with a central value of £1 = DM2.95 and a 6% fluctuation either side. This exchange rate was too high and the Bank of England was unable to buy sufficient sterling on foreign exchange markets to maintain the pound within the currency band. The UK government was forced to withdraw the pound from the ERM in September 1992.

Floating exchange rate systems

Floating exchange rates are used by the majority of world economies and by many of the world's major currencies including the US dollar, the euro, the pound sterling and the Japanese yen.

How floating exchange rates are determined

Floating exchange rates are determined by the sale and purchase of currencies on foreign exchange (or 'forex') markets. The exchange rate of a currency is simply its 'price' expressed in another currency, or possibly a basket of currencies.

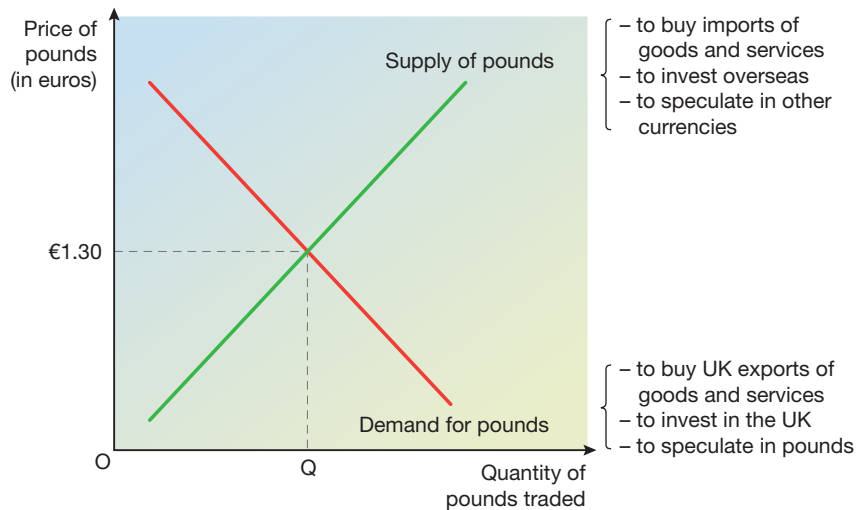


Figure 53.2 *The determination of a country's floating exchange rate*

The demand for a country's currency arises from three main sources, as shown in Figure 53.2. Thus, for example, there are three broad reasons why a demand for

pounds sterling might exist from consumers and organisations in the 19 countries that use the euro as their currency – the eurozone:

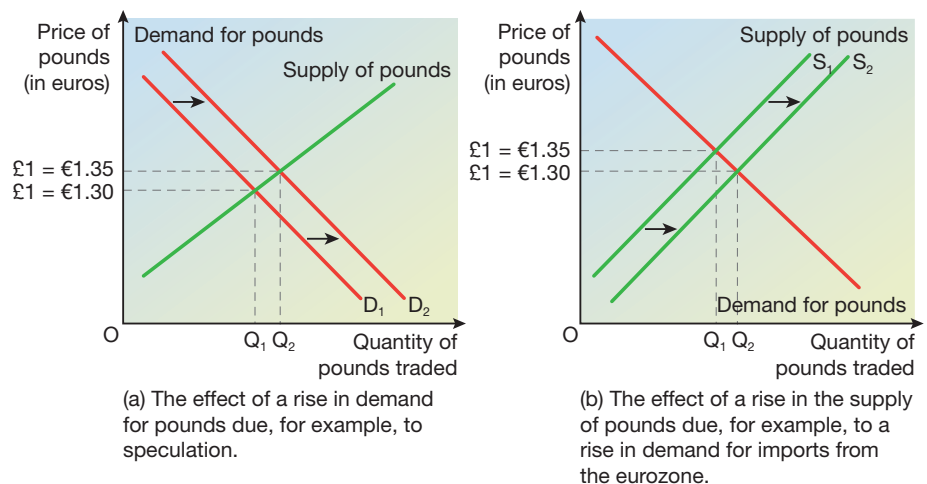
- **To buy UK exports:** Eurozone consumers, firms or governments wish to buy UK exports of goods and services. Exporters in the UK will normally expect to be paid in the domestic currency and thus eurozone buyers will need to sell their currencies on foreign exchange markets to buy pounds.
- **To invest in the UK:** Eurozone households and organisations may wish to invest in UK firms (either in the form of direct or portfolio investment) or place funds in UK banks if interest rates are attractive. In order to do so, they will sell euros and buy pounds sterling, thus creating a demand for the currency on foreign exchange markets.
- **For speculative purposes:** A demand for pounds may arise if speculators in the eurozone area believe that the exchange rate of sterling is likely to rise and they buy the currency in expectation of making a profit in this way. The effects of such an expectation are illustrated in Figure 53.3(a), which shows that the exchange rate would rise in such circumstances, *ceteris paribus*. Thus the expectation becomes a self-fulfilling prophesy.

Pounds are supplied onto foreign exchange markets when holders of the currency decide to exchange it for a different currency. Once again there are three broad reasons why holders of pounds might sell them to buy another currency (in our example, euros):

- **To buy imports from the eurozone countries:** Sellers in the eurozone area will expect to be paid in their own currency.
- **To invest in the eurozone area:** Holders of pounds might, for example, buy shares in Italian companies.
- **For speculative purposes:** This would entail the sale of sterling and the purchase of euros in the belief that the value of the euro would rise against sterling.

As shown in Figure 53.3(b), an increase in the supply of pounds reduces the exchange rate as they become less scarce on foreign exchange markets. It is worth noting that exchange rate transactions have reciprocal effects on the partner currencies. For example, the sale of pounds in our example results in an increased demand for euros that, *ceteris paribus*, would lead to a rise in its exchange rate.

Figure 53.3 Changes in the demand and supply of pounds sterling



Freely floating exchange rates

This is a relatively simple exchange rate model where market forces are allowed to determine the value of a currency. Thus decisions by economic agents across the globe to buy and sell the currency determine its value on world markets. In a few countries such as Sweden, Canada and the USA, central banks almost never intervene to manage the exchange rates.

Government intervention to influence exchange rates

Managed floating exchange rates

When a government intervenes in foreign exchange markets to influence the exchange rate of its currency, this is referred to as a managed floating exchange rate system. This exchange rate system allows a currency's value to be determined by free market forces to some extent. Thus, a government may allow its currency to float within an agreed range of values on foreign exchange markets and its central bank will intervene to buy or sell the currency or adjust interest rates to ensure that it is kept within a desired range. For example, if a government wishes to increase its exchange rate, it could opt to buy its currency on foreign exchange markets or raise its interest rate to attract inflows of currency, which will increase demand for the currency and raise its exchange rate.

Figure 53.4 shows situations in which a government or central bank might wish to manage the value of its currency to keep it within an agreed range (as measured on its trade-weighted index). Its management could take the form of buying and selling its currency on foreign exchange markets, changing interest rates or using other policies such as taxing foreign investments.

REALWORLD ECONOMICS 53.1

A divergence in monetary policy causes currency volatility

Several world currencies are under pressure. The Swiss franc rose by 30% against the euro in January 2015 and the Russian rouble has plunged against the dollar in the face of a declining oil price and economic sanctions by the West. This week the Reserve Bank of Australia announced a surprise interest rate cut, sending the Australian dollar down to its lowest level against the US dollar since May 2009. Denmark has had to cut interest rates three times, moving into negative territory, in order to discourage capital inflows that were

threatening its fixed rate against the euro.

What is behind this sudden burst of currency volatility, which follows a quiet period in foreign exchange markets? In large part, it is caused by a divergence in monetary policy among the big three central banks: the Federal Reserve, the European Central Bank and the Bank of Japan. The Federal Reserve may push up interest rates this year (it did so in December 2015), but the Bank of Japan is still implementing a policy of quantitative easing (QE) and the European Central Bank is

just about to start one.

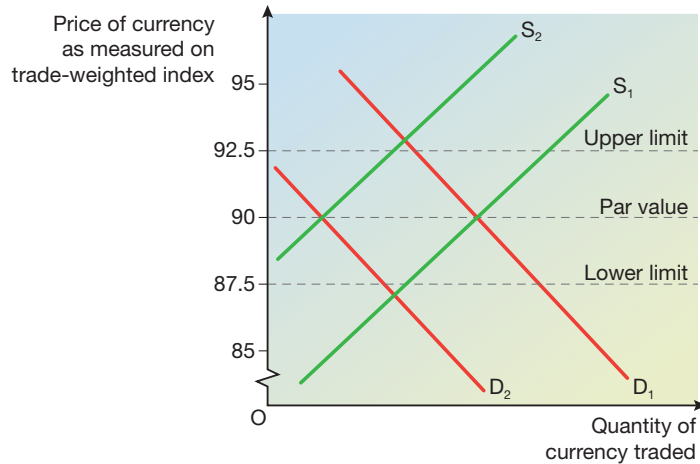
These diverging policies reflect economic fundamentals. The American economy is growing at a decent rate, but both Japan and the eurozone are struggling to generate a sustainable recovery. And, like Japan, the eurozone is teetering on the brink of deflation.

Source: *The Economist*, 7.2.15

Exercise

- 1 Explain why economies such as Japan may benefit from floating exchange rates in these circumstances. (7 marks)

Figure 53.4 *Managing a currency to keep it within predetermined limits*



If we assume that the currency was initially at a trade-weighted exchange rate of 90 (its par or central value) as a result of the interaction of D_1 and S_1 . Demand for its currency could fall to D_2 . If supply is unaltered, this would reduce its trade-weighted exchange rate to around 87, below its lower limit or floor. In such circumstances the government could seek to increase demand or reduce supply of its currency by the use of one or more of the following:

- It might sell its official reserves of gold and foreign currency to buy its currency on foreign exchange markets. This will increase demand and its exchange rate.
- It might raise its interest rate. This could provoke an inflow of investment as foreigners seek to gain the increased returns available within its economy. It may also reduce supply as investors decide to keep their funds within the economy.
- It might implement other policies such as reducing taxes on overseas investments in the country. This might increase demand for the currency as well as reducing supply.

A government or central bank may also intervene in the foreign exchange market if the value of its currency rises above the upper limit or ceiling. This might arise due to a fall in the supply of a currency, possibly because the country's domestic products are increasingly competitive. This is shown in Figure 53.4 by the move of S_1 to S_2 . This could be corrected by the central bank selling a currency to increase its supply, or by reducing interest rates to lower demand and increase supply for the currency.

The Japanese central bank has intervened regularly in foreign exchange markets to manage the value of its currency (the yen) with the aim of protecting its trade performance. In 2015 Haruhiko Kuroda, the Governor of the Bank of Japan, offered guidance to foreign exchange markets as to a range of values within which the yen would be allowed to move.

'Dirty' floating

A country may attempt to manipulate the value of its currency to gain an advantage in its international trade. For example, it may prevent the currency appreciating to maintain the price competitiveness of its exports on global markets. This is termed 'dirty' floating.

China is a notable example of a country whose currency (the yuan) is thought by

some economists to be managed in this way. The yuan is pegged to the US dollar within a fixed band (which is reviewed daily) to control the value of the yuan. The Chinese government prevents any possible appreciation of its currency by selling yuan and buying dollars as necessary. There have been widespread suggestions that the yuan is held at a substantially undervalued level to make Chinese exports more competitive in global markets.

The advantages and disadvantages of fixed and floating exchange rate systems

There are arguments in favour of, and against, both fixed and floating exchange rate systems.

Arguments in favour of a fixed exchange rate system

1 Supports trade by eliminating fluctuations in currencies

If a currency is completely fixed in value against other currencies, it removes the possibility of currency fluctuations provided the system succeeds. This offers benefits to both exporters and importers. Exporters know what their products will cost in terms of foreign currencies and can price accordingly. They are also reassured that a sudden appreciation in the currency, which would make their products uncompetitive in global markets, will not occur. Importers are able to predict the costs of raw materials and components as well as final goods and services with greater accuracy and will not face a sudden hike in costs following currency depreciation. This assists financial planning and decision making and helps to maintain the economy's long-term competitiveness and prosperity by encouraging trade.

2 Attractive to investors from overseas

Fixed exchange rates, especially if they have been fixed for a long period of time (for example, the Saudi Arabian riyal), can be attractive to overseas investors as they remove much uncertainty and risk related to investing in the country. Overseas investors have less reason to fear a fall in value of the currency of the country in which they have invested, which would reduce the value of their investment as well as the value of any flow of income from it. Thus countries with fixed exchange rates may be more successful in attracting inflows of FDI. FDI can bring supply-side improvements and can result in positive multiplier effects, boosting aggregate demand.

Arguments against a fixed exchange rate system

1 Conflict with other macroeconomic objectives

A government will inevitably have to take action from time to time to maintain its fixed exchange rate against other currencies. This might entail raising interest rates if the currency is in danger of slipping below its fixed rate or through the floor of its range of acceptable rates. Increasing interest rates will impact upon other aspects of the country's economic performance. It may, for example, have a substantial deflationary effect, slowing rates of economic growth and raising the rate of unemployment. Equally, lowering interest rates at a time when the exchange rate shows signs of unacceptable rises might provoke inflationary pressures.

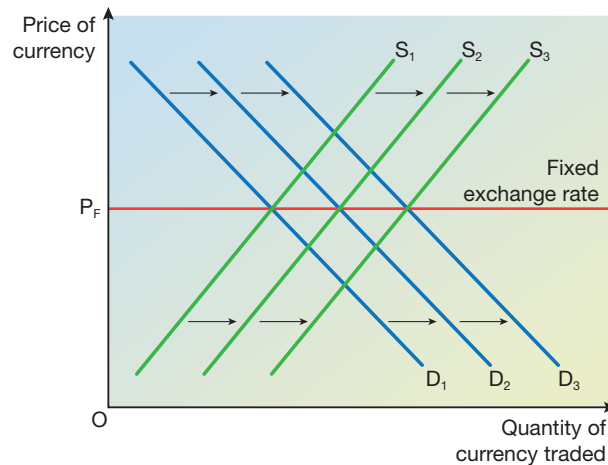
2 The need to revise fixed rates regularly

It is inevitable that, over time, the economic performance of countries will differ. Some will grow more quickly than others: for example, in 2015 China's economy grew by 6.8% and that of the USA by 2.1%, while Brazil's rate of economic growth was -4.5%. At the same time, rates of inflation can vary too: the three countries mentioned above had inflation rates of 1.6%, 0.7% and 10.7% respectively. In such circumstances a fixed exchange rate system would mean that some countries would accumulate permanent surpluses while others would run deficits in the long term. In order to avoid this, fixed rates need to be adjusted.

3 Pressure to revise rates from speculators

If the perception is that a fixed exchange rate is set at the wrong level, it is likely that speculators will begin to buy or sell the currency in the expectation of it being adjusted. For example, if speculators consider a fixed exchange rate currency to be overvalued, they may begin to sell the currency in the expectation of being able to buy it back more cheaply once its value has been adjusted. This poses a problem for the central bank concerned. It can act against the speculators in the market by buying its currency, thereby increasing demand for it and protecting its value. This is illustrated in Figure 53.5. The actions of the speculators are shifting the supply curve to the right, from S_1 through S_2 to S_3 . A central bank can combat this by buying the currency, increasing demand to D_2 and D_3 as necessary, if it has sufficient reserves of foreign currency and gold.

Figure 53.5 The effects of speculation against a currency within a fixed exchange rate system



A central bank or government may be unwilling to raise interest rates to the extent required to protect the currency. Furthermore, speculators are likely to be unwilling to purchase a financial asset that is expected to fall in value in the near future, whatever the returns. Hence it may prove impossible to maintain a fixed exchange rate, thereby losing many of its benefits.

Fixed exchange rates are generally more effective and successful if the correct rates are selected at the outset. There tends to be less pressure to adjust rates and the problems of overvaluation (lack of competitiveness in export markets) and undervaluation (imported inflation) are avoided. However, relatively few countries use fixed exchange rates nowadays because the demands arising from greater links

between economies in terms of trade and capital flows have been better served by the flexibility of floating exchange rates.

Arguments in favour of a floating exchange rate system

We saw earlier that floating exchange rates range from freely floating to managed exchange rates where governments intervene in the operation of the foreign exchange markets.

1 Automatic adjustment of exchange rates to reflect trading positions

In theory, freely floating exchange rates should adjust to eliminate balance of payments deficits and surpluses. A country with a large surplus should experience an appreciation of its currency as the popularity of its exports will provoke a large demand for the currency. Simultaneously, the competitiveness of its domestic industry will restrict purchases of imports, limiting the supply of the currency on the foreign exchange markets. The net effect will be to force up the exchange rate, reducing the price competitiveness of the country's goods and services and eliminating the surplus over time. Similarly, a deficit will be corrected by an automatic depreciation in the exchange rate.

2 Reduced possibility of conflicts between macroeconomic objectives

Governments are less likely to be faced with decisions on economic policy that create conflict between macroeconomic objectives. For example, a country with a shortage of reserve assets may be faced with a dilemma. It may need to raise interest rates to protect a fixed exchange rate, but this would have a deflationary impact on the economy, slowing growth and reducing employment and the standard of living.

A freely floating exchange rate removes the need for such challenging decisions as well as the requirement for a central bank to maintain large holdings of foreign currencies as reserve assets.

3 Greater flexibility

Under a managed version of a floating exchange rate, a government can enhance the price competitiveness of its exports by taking actions to provoke a fall in the value of its currency. For example, it may opt to lower interest rates or to sell its currency on world markets. Monetary policy decisions, such as the use of quantitative easing, which increases the supply of money, can also be used for such a purpose.

The flexibility of floating exchange rates has enabled the necessary adjustments in the value of different world currencies even during periods of financial crisis, such as that experienced over the period 2008–10.

Arguments against a floating exchange rate system

1 Floating rates do not adjust automatically to eliminate imbalances

Economic theory tells us that freely floating exchange rates adjust to eliminate imbalances in a country's balance of payments. However, the reality is different because only a minority of transactions on foreign exchange markets are associated with international trade in goods and services. Globalisation has led to a dramatic increase in capital flows between countries and there are enormous speculative flows

across the world. As a result, the necessary changes in the exchange rate to correct an imbalance do not always occur. For example, the UK's current account deficit widened significantly over the period 2012–14 from 3.5% to 5.9% of GDP. At the same time the value of the pound against the dollar remained around £1 = \$1.50.

2 Potential destabilisation of economies

Under fixed exchange rate systems, governments often limit the amount of currency that can be taken out of a country. Limiting outflows of currency in this way prevents the massive currency flows that occur under floating exchange rates. Freely floating rates can result in huge currency flows, with profound implications for the country's economy. For example, a large outflow of currency might increase inflationary pressures by increasing the price of imports. These inflationary pressures could be exacerbated if rising sales of increasingly price-competitive exports were to lead to sharp increases in aggregate demand.

Author tip

Do remember when discussing the advantages and disadvantages of fixed and floating exchange rates that less pure forms of the systems can avoid some of the disadvantages that might otherwise occur.

Currency unions

A simple currency union only involves the sharing of a currency by two or more countries. However, most currency unions, such as the 19 countries that use the euro and comprise the eurozone, are integrated to a greater degree, trading within a single market and without barriers to the free movement of resources such as labour and capital.

The eurozone

When the Treaty of Rome was signed in 1957, the founding countries focused on creating a 'common market' for trade. Later it became clear that closer economic and monetary cooperation was needed for the internal market to fulfil its potential, creating greater prosperity for Europeans.

In 1991 the Maastricht Treaty was signed, establishing the euro as the common currency for 11 EU member states. Since then, eight other EU member states have adopted the euro. In addition, two other countries have pegged their currencies to the value of the currency. Euro notes and coins came into circulation on 1 January 2002. The members of the eurozone are shown in Figure 53.6.

Eurozone members operate a common monetary policy, such as a single rate of interest, set by the European Central Bank (ECB) based in Frankfurt in Germany. The eurozone does not operate a common fiscal policy, though some cooperation occurs through the activities of the Eurogroup. However, the prospective members of the eurozone signed the Stability and Growth Pact (SGP) in 1997. The SGP established rules with the aim of maintaining the value of the euro by enforcing fiscal responsibility. Under the SGP rules, eurozone members must avoid budget deficits greater than 3% of GDP and must hold national debt at a figure that is below 60% of GDP.

Key term

A **currency union** exists when two or more countries use the same currency as well as a single monetary and foreign exchange policy.



Figure 53.6 The member states of the eurozone

Advantages of being a member of a currency union

1 A more competitive market

When a number of countries use the same currency, it is relatively simple to compare prices from different retailers or suppliers across the currency union. As a result, firms can use the cheapest source of supplies and consumers can obtain the best value for money. This helps to improve competitiveness in industries across the union. It would be realistic to expect prices across the union to converge over time towards the most competitive and lowest-cost producers.

Price transparency also helps to control inflation by increasing competition. For example, during the 1970s and 1980s many EU countries had very high inflation rates, some exceeding 20%. Inflation started to fall in the run up to the adoption of the euro and has been around 2% in the eurozone since then.

2 The benefits of removing exchange rate uncertainty

By trading within a currency union, firms can avoid the possibility of being affected by adverse movements in exchange rates. In effect, any transfers of money by households or firms are 'domestic' transactions. Companies no longer have to engage in strategies to avoid the effects of adverse exchange rate movements, such as hedging, which add to production costs.

In the eurozone, this removal of exchange rate risk was of particular value to countries such as Greece, Italy and Spain whose currencies had been more vulnerable to significant fluctuations. Research shows that creating the eurozone has more than doubled trade among its member states. The absence of exchange rate risk and restrictions to capital transfers has increased investment in physical capital within the eurozone as well as attracting increasing volumes of FDI.

Discussion point

Do you think that the benefits of a currency union are greater for smaller economies, such as Ireland, or for larger ones, such as Germany?

3 Reduced interest rates

The European Central Bank has achieved low inflation and stable prices within the eurozone. This success means that interest rates are also lower, which in turn means that consumer loans are cheaper and future repayments are more predictable, so ordinary citizens can borrow more easily and cheaply – for example, to pay for holidays or to buy a house. Mortgage rates have fallen from around 8–14% in the early 1980s to an average of less than 5%, increasing income available for consumer expenditure on other items. Low interest rates also encourage higher rates of business investment as more proposed projects become financially viable.

Disadvantages of being a member of a currency union

1 The loss of control over national economic policies

This is arguably the greatest potential problem for governments whose economies are part of a currency union. A currency union will have a single monetary policy administered by a common central bank and, of course, a single rate of interest. The chosen interest rate will be suitable for the currency union as a whole, but is most unlikely to be the optimal rate for every member country. Not all the member states' economic cycles will move in harmony and thus they may require a rate of interest that is substantially different from that set by the currency union's central bank.

At the time of writing the Greek economy is teetering on the brink of recession (its GDP fell by 0.9% in the last quarter) and its unemployment rate is above 24%. In contrast Ireland's economy is growing at over 7% per annum with unemployment around 8%. The ECB's chosen interest rate is unsuitable for both economies. The issue becomes more problematic when inflation is taken into account to calculate real interest rates. Real borrowing costs vary widely across the eurozone due to differing inflation rates. Real interest rates are higher in countries such as Greece and Spain, where prices are falling at the time of writing, than in Germany and the Netherlands. Yet Germany and the Netherlands are enjoying economic growth rates approaching 2% per annum and are more likely to require higher, rather than lower, interest rates. At the same time Greece and Spain would benefit from the reflationary effects of lower interest rates.

A consequence of the differences in economic performance and real interest rates is that capital and skilled labour tend to flow to the more prosperous members of the currency union. Unfortunately, in the eurozone the restrictions of the Stability and Growth Pact limit the ability of member states to use fiscal policy to address these issues. Thus it can be argued that a currency union can provoke differences in macroeconomic performance and remove the ability of individual governments to deal effectively with their diverse circumstances.

2 The inability to use the value of the currency as a tool of economic management

A country outside a currency union can allow the exchange rate of its currency to fall if it needs to boost aggregate demand or to correct a balance of payments deficit. This particular 'weapon' is not available to members of currency unions. The Greek economy has experienced huge financial difficulties in recent years. One dimension of its difficulties is that, as a member of the eurozone, it is operating with an exchange rate that many economists regard as too high. If Greece had not become a member of

the eurozone, its economy could have been much more competitive in global markets with a lower exchange rate.

Review questions

Total: 41 marks

- 1 Define the term 'multilateral exchange rate'. **(3 marks)**
- 2 Explain the difference between fixed and floating exchange rates. **(6 marks)**
- 3 Which of the following is likely to lead to an increase in the value of a freely floating currency, such as the Canadian dollar, *ceteris paribus*?
 - A A fall in sales of Canadian exports of goods and services
 - B A fall in purchases of imports by Canadians
 - C A fall in the rate of interest in Canada
 - D A fall in foreign FDI into Canada**(1 mark)**
- 4 Which of the following is likely to lead to a fall in the value of a freely floating currency, such as the Canadian dollar, *ceteris paribus*?
 - A Households and firms in economies in Europe and Asia increase imports from Canada
 - B Canadians sell investments in overseas businesses
 - C Canadians reduce their level of portfolio investment
 - D Canadian export sales fall substantially**(1 mark)**
- 5 Explain one way in which a government could support the value of a fixed exchange rate at a time when its value is falling on world markets. **(5 marks)**
- 6 Explain, with the aid of an example, what is meant by the term 'dirty floating'. **(5 marks)**
- 7 Explain why fixed exchange rates can help to make an economy attractive to international investors. **(6 marks)**
- 8 Explain why floating exchange rate systems may not automatically eliminate balance of payments surpluses and deficits. **(6 marks)**
- 9 Which of the following is *not* an advantage for an economy of being a member of a currency union?
 - A An increase in price transparency
 - B An increase in competitiveness throughout the currency union
 - C An increase in transaction costs
 - D An increase in trade between members of the currency union**(1 mark)**
- 10 Explain why being a member of a currency union may result in an economy having an interest rate that is unsuitable to its circumstances. **(7 marks)**

Economic growth & development

Key concepts from Year 1

There are no relevant concepts from Year 1 as this is an entirely new topic.

This chapter commences by explaining the difference between economic growth and economic development. It outlines the main characteristics of less developed economies as well as how development can be measured. We investigate the factors that affect growth and development including the barriers that may be encountered. Finally, we examine the policies that may be adopted to promote economic growth and development including the role that can be played by international aid.

Economic growth and development

We have discussed economic growth through this book and the Year 1 companion volume. It measures the extent of any growth in an economy's ability to produce goods and services as measured by its GDP. Economic growth rates can be easily compared between countries as well as over time. However, it has significant limitations as a measure in that it can only provide average data and does not measure the extent to which the available resources are used to meet the needs of all of those in society.

Development is a much broader term, but one which is not simple to define concisely. However, Michael Todaro, Professor of Economics at New York University and a specialist in development economics, has defined it as follows: 'The process of improving the quality of all human lives and capabilities by raising people's levels of living, self-esteem and freedom'.

He recognises that while economic development needs to measure the extent to which incomes rise in a country, it also needs to encapsulate other factors such as the extent to which poverty is alleviated over time. Reducing poverty enhances people's self-esteem as they become more self-reliant and it grants them greater freedom to make choices about their lives.

Poverty can be judged in two ways:

- **Absolute poverty:** This is a matter of acute deprivation, hunger, premature death and suffering. It is determined both by income levels and by access to services such as healthcare and the supply of clean water.
- **Relative poverty:** This form of poverty relates to the current place and time and it measures a person's wellbeing relative to others in society. It describes a level of deprivation that does not match the general living standards enjoyed by the majority of the population.

Economic development and the freeing of people from poverty is an important target

Key terms

Economic growth

occurs when, over time, an economy expands its capacity to produce goods and services.

Economic development

is a broader term encompassing judgements relating to factors such as standards of healthcare and education and not just income levels.

for many international organisations such as the United Nations (UN). The UN has identified 15 development goals that should be met if countries are to develop their economies (and not just raise average incomes) and to lift their populations out of poverty, thereby enhancing their self-esteem and granting greater freedom – for example, by offering more choices in how to make a living.

On 25 September 2015, developed countries adopted a set of goals set out in the 2030 agenda to end poverty, protect the planet and ensure prosperity for all as part of a new sustainable development agenda. Each goal has specific targets to be achieved by 2030.

We saw in Chapter 50 that globalisation has helped many economies to enjoy rapid rates of economic growth and thus to offer the potential for development. For example, China has removed more people from poverty than anywhere else in the world: its per capita income increased 25-fold between 1990 and 2010, from \$200 to \$5000. The rapid growth of international trade that has accompanied globalisation has been a powerful force in helping countries to lift their citizens out of poverty.

The main characteristics of less developed economies

The term ‘less developed economy’ is a broad one and capable of different interpretations. It may encompass economies at quite different stages of development. On the one hand, many underdeveloped economies in sub-Saharan Africa such as Burkina Faso, Liberia and Ethiopia have very low per capita incomes, severely limiting the ability of their governments to provide vital services including health and education. In contrast, relatively developed economies in South America such as Ecuador and Paraguay enjoy incomes per capita that are ten times higher than those in Liberia, providing considerable potential for a much higher level of development. Some less developed economies are rich in mineral resources; others have huge populations on which to draw.

Despite the term covering a range of economies, there are a number of factors that can be used to characterise and identify less developed economies.

1 Low income levels

Income is not the sole determinant of an economy’s degree of development, but it is used by many international organisations to classify countries as less developed or more developed. Most judgements concerning an economy’s level of development are made using income per head or capita.

Less developed economies can be classified as those having low or middle incomes, with a few exceptions. In 2016 the World Bank classified low-income economies as those with a national income per capita of \$1045 or lower in 2014. Middle-income economies have a national income per capita of more than \$1045 but below \$12,736.

Key terms

Income per head or capita is a country’s GDP divided by the size of its population.

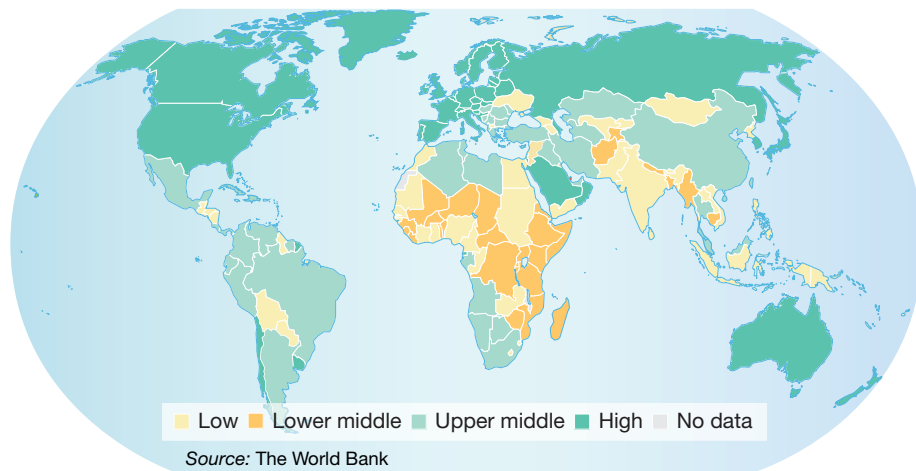
Purchasing power parity (PPP) is an exchange rate based on the cost of a common basket of goods and services in different countries.

The **primary sector of an economy** contains those businesses that make direct use of natural resources such as minerals and land.

Income levels are a limited means of identifying less developed economies. Some high-income economies may have higher than average incomes but rely on one or two industries (such as oil extraction and refining) and have large numbers of people in poverty due to extreme income inequality. Equally, some less developed nations can be exceptionally poor with very high levels of poverty, emphasising that this term encompasses vastly different economies. For example, the two African countries of Liberia and Burundi both had per capita income levels below £550 per annum in 2011. Figure 54.1 shows per capita incomes for most of the world's economies.

Finally, it is not always easy to compare the income of one country with that of another when there are different currencies. This is commonly done using official exchange rates, but therein lies a problem. The use of official exchange rates tends to exaggerate the income gap between developed and less developed economies. Official exchange rates ignore the differences in the domestic purchasing power of different currencies because they are determined by government actions and the activities of speculators rather than simply by relative prices in the economies. Economists have attempted to avoid this problem by using exchange rates based on purchasing power parity (PPP) instead of official exchange rates when calculating incomes per capita for different economies. PPP exchange rates are based on the cost of a common basket of goods and services in different countries and thus reflect relative purchasing powers more accurately. The data in Figure 54.1 are not based on PPP, but the exchange rates have been manipulated to reduce the impact of exchange rate fluctuations in the cross-country comparison of national incomes.

Figure 54.1 *The national income per capita of the world's economies in 2013*



There are other limitations in using income data to measure living standards and hence to provide an indicator of development. In some very poor economies, trade is based on barter (i.e. the exchange of goods and services without money). This approach to trade is likely to understate living standards as measured using GDP data. Furthermore, per capita income data provide an average measure and ignore the effects of extreme levels of income inequality on living standards.

2 Lower levels of expenditure on health and education

Less developed economies will tend to have lower levels of attainment in terms of the provision of healthcare and education, especially that given to children above primary

school age. Measures of this aspect of development might include the percentage of those of school age enrolled in secondary education. World Bank data show that for Afghanistan in 2013 the figure was 54% and for Burkina Faso it was only 28%. In contrast, developed economies such as Austria and Sweden achieved rates of 99% and 98% respectively. Low rates of education tend to hold back development by denying people the chance to utilise their abilities to their full extent and limiting opportunities to pursue different career choices.

Inadequate healthcare is likely to result in a greater incidence of diseases that would be treatable in other circumstances. Death rates are much higher, especially amongst the poorest groups in society. Poorer levels of healthcare can also result in higher birth rates as parents expect relatively high rates of infant mortality. This, along with limited access to techniques of birth control, may deny many women the opportunity to enter the workforce.

3 A shortage of capital

Less developed economies are sometimes referred to as ‘capital-poor economies’, reflecting their lack of this vital resource. It is common for them to have a relatively small stock of capital, exacerbated by low levels of investment. In many less developed economies, investment rates can be as low as 8% of GDP, and GDP levels are much lower. Data from the World Bank show that net investment in 2014 in Australia and France was \$392,762 billion and \$622,422 billion respectively. This is in contrast to the low levels of investment in Afghanistan (\$3607 billion) and South Sudan (\$1328 billion) in the same year.

Shortages of capital occur because the entire economy may not be monetised or because levels of savings are very low, providing banks with a limited supply of

REALWORLD ECONOMICS 54.1

Commodity prices in the global economy

It is possible to view the commodity price decline as the global economy

The Economist commodity-price index 2005 = 100



Figure 54.2 An index of global commodity prices, 2011–15

simply doing its job in the form of a supercycle. Rising prices bring forward new sources of supply as production at marginal sites becomes profitable; they also cause consumers to economise on raw material use. This brings prices down, so production falls sharply, allowing the whole cycle to begin again. Since it takes a lot of time to build new mines, develop oil fields etc., this whole process takes more than a decade.

Nevertheless, there seem to be broader signs of concern about the global economy, which the commodity price decline

appears to reinforce. Indices of manufacturing production in less developed economies have dipped slightly below 50, indicating a decline in output. Figures show that unemployment in emerging markets has risen from 5.2% to 5.7% since the start of 2015.

Source: Adapted from *The Economist*, 21.7.15

Exercise

- 1 Explain why fluctuations in commodity prices make it difficult for less developed economies to achieve satisfactory rates of economic development. (7 marks)

funds to lend to governments and businesses. In some less developed economies, the banking sector is not sufficiently advanced to be able to channel funds effectively into productive use. Economies with volatile exchange rates may struggle to attract FDI owing to the degree of risk attached.

4 High rates of population increase

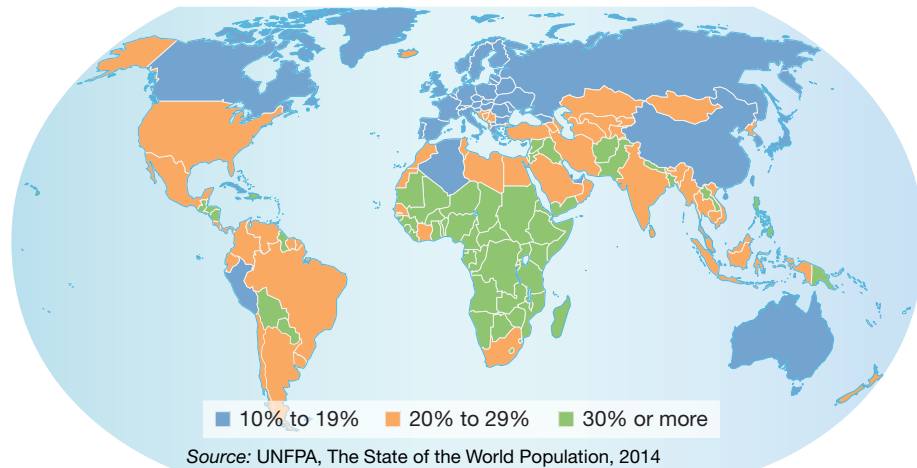
Many less developed countries continue to experience high rates of population growth. In part this can occur because of a lack of education linked to low levels of healthcare. The outcome is a lack of knowledge of, as well as limited access to, methods of birth control. As a consequence, birth rates are high in comparison to more developed countries. In some countries the rate of increase of the population is outstripping the ability of the government to provide education and healthcare, with negative consequences for development.

Countries with high rates of population growth tend to have high proportions of young people in the population. Figure 54.3 shows that many of the countries with the highest numbers of young people are in Africa where levels of economic development are generally low. In 2014 the five countries with the highest rates of population increase were all African: South Sudan – 4.02%; Malawi – 3.31%; Burundi – 3.27%; Niger – 3.25%; and Uganda – 3.24%.

Number crunching

In 2014 the population of South Sudan was 11.91 million, having grown by 4.02% since 2013. Calculate the size of South Sudan's population in 2013.

Figure 54.3 The percentage of 10–24 year olds in the population in countries throughout the world



Summary

Although less developed countries vary greatly, there are characteristics that they share. Table 54.1 shows how poverty and high rates of population growth are associated particularly with low-income countries. It is these countries that are likely to be the least developed. Some aspects of the data do indicate that development has taken place, but the fact that rates of population growth continue to be highest in the world's poorest countries is a significant statistic in relation to their prospects for development.

Indicator	High-income countries		Middle-income countries		Low-income countries	
	1990–2	2015	1990–2	2015	1990–2	2015
Percentage of population undernourished	0	0	22	12	41	27
Actual and forecast annual population growth rates (%)	2000–14	2014–25	2000–14	2014–25	2000–14	2014–25
	1	0	1	1	3	3
Life expectancy at birth	1990	2013	1990	2013	1990	2013
	74	79	64	70	50	61
Percentage of population having access to a source of safe water	1990	2015	1990	2015	1990	2015
	98	99	72	92	45	66

Source: World Bank Development Indicators

Table 54.1 Changes in a selection of development indicators for high-income, middle-income and low-income countries between 1990–92 and 2014–25

The main indicators of development

One measure of development is the level of per capita income in a country. It is relatively simple to compare the GDP per capita between different countries using official exchange rates or, preferably, exchange rates based on purchasing power parities. However, this measure of development has significant limitations.

- It ignores the distribution of income. In some countries the GDP per capita is comparatively high, but this disguises extreme differences in income levels between the highest-income and the lowest-income households. Some households can suffer poverty and deprivation.
- It ignores factors that cannot be easily quantified, such as negative externalities. Some countries are enjoying rapid rises in GDP per capita but are simultaneously blighted by the consequences of pollution. For example, many of China's cities are notorious for high levels of air pollution. Beijing, the Chinese capital, suffers from serious air pollution. This is the result of exhaust emissions from more than five million motor vehicles, the burning of coal in factories in nearby regions, and dust storms from the north and from local construction activities.

The limitations of GDP as a means of measuring the development of economies has resulted in the development of measures designed to provide a better indication of quality of life within an economy. We shall consider the Human Development Index (or HDI).

The Human Development Index

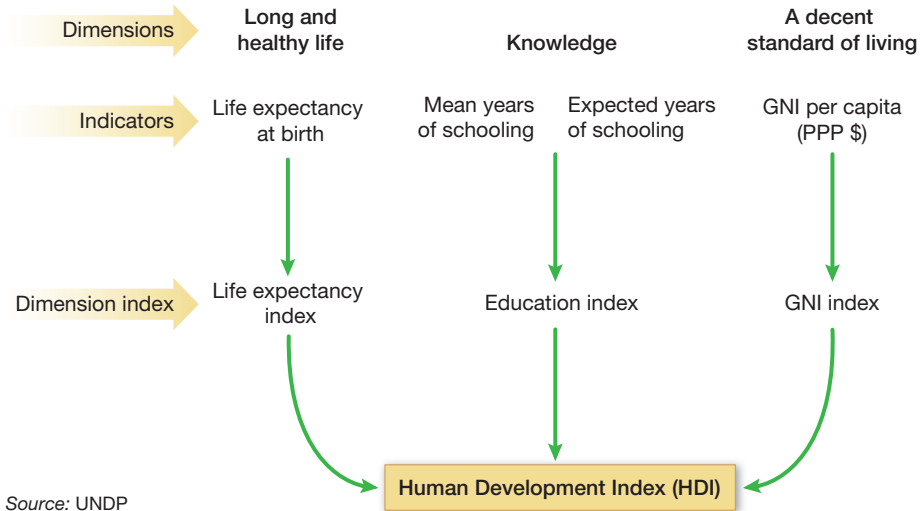
Pakistani economist Mahbub ul Haq created the Human Development Index (HDI) in 1990 on behalf of the United Nations Development Programme (UNDP) as a deliberate move towards putting people at the centre of any assessment of the degree of a country's economic development. It was designed to emphasise that people and their capabilities should be the ultimate criteria for assessing the development of a country rather than economic growth alone.

The HDI measures an economy's average achievement in three important dimensions of human development (Figure 54.4):

- *longevity*, as measured by life expectancy at birth – this is an indication of health;
- *education* – since 2010 this has been measured by a combination of mean years of schooling and expected years of schooling;

- *the average standard of living* – the HDI measures this using real GDP per capita based on purchasing power parity exchange rates.

Figure 54.4 The components of the Human Development Index, 2015



The HDI is the mean of the indices for each of the three elements; thus each element is given equal weighting. It is intended for use in charting the development of economies over time. To this end, the UNDP publishes an annual Human Development Report containing relevant data on development.

Interpreting HDI data

The HDI produces results expressed on a scale from 0 (no development) to 1 (complete development) and this allows economies to be classified and compared easily. The 2015 Human Development Report used the same bands for the four categories of human development that were introduced in 2014. These are shown below.

- **Very high human development:** Countries in this category have an HDI figure of 0.800 and above. In 2014, Norway was the world's most highly developed country with an HDI figure of 0.944.
- **High human development:** In this category the HDI figures extend from 0.700 to 0.799. In 2014, countries in this category ranged from Belarus (0.798) to Samoa (0.702).
- **Medium human development:** Countries in this classification achieve HDI scores between 0.550 and 0.699. In 2014 Botswana (0.698) was the highest and Cambodia (0.555) the lowest in this category.
- **Low human development:** This category comprises countries with HDI figures below 0.550. Niger with an HDI of 0.348 was the least developed country in 2014.

Discussion question

Do you think that income levels are given sufficient weight in the HDI calculation?

Table 54.2 illustrates several of the key features of the HDI and its global distribution. Income per capita is only one of the three factors and high-income countries do not always perform well in terms of HDI. For example, the USA has a higher per capita income than Australia, but a lower HDI. This is because it performs less well in terms of life expectancy and education. Similarly, despite a relatively high income per capita, Egypt has a comparatively low HDI. Education appears to be its weakness:

mean years of attendance are significantly below what is expected. Finally, many of the countries with low HDI figures are in Africa. The per capita income of people in the Central African Republic is very low, indicating extensive poverty.

Country and global rank	HDI	Life expectancy at birth (years)	Expected years of schooling	Mean years of schooling	Gross national income per capita 2011 (\$PPP)
<i>Very high human development</i>					
2. Australia	0.935	82.4	20.2	13.0	42,261
8. USA	0.915	79.1	16.5	12.9	52,947
14. UK	0.907	80.7	16.2	13.1	39,267
<i>High human development</i>					
52. Uruguay	0.793	77.2	15.5	8.5	19,283
66. Serbia	0.771	74.9	14.4	10.5	12,190
99. Jamaica	0.719	75.7	12.4	9.7	7415
<i>Medium human development</i>					
108. Egypt	0.690	71.1	13.5	6.6	10,512
116. Vietnam	0.666	75.8	11.9	7.5	5092
142. Bangladesh	0.570	71.6	10.0	5.1	3191
<i>Low human development</i>					
145. Kenya	0.548	61.6	11.0	6.3	2762
171. Afghanistan	0.465	60.4	9.3	3.2	1885
187. Central African Republic	0.350	50.7	7.2	4.2	581

Source: UNDP

Table 54.2 HDI data for a selection of countries

An assessment of HDI as a measure of development

The HDI is the most used measure of human development. There are a number of reasons for this. Its calculation provides a single figure, making comparisons between countries straightforward and transparent. It is based, to some extent, on relative income figures per head and they avoid the trap of using official exchange rates.

However, the index also has a number of shortcomings.

- It does not include any measure of the quality of the natural environment and the extent to which this is being damaged through pollution. This is a major influence on the life experiences of many people, especially in cities in developing countries and in mining areas.
- Another significant exclusion is technological developments and their ability to change lives and to contribute positively to the human experience. For example, access to the internet has the potential to change people's lives and to allow greater control and choice. It also ignores a range of other factors such as crime, gender disparity, income inequalities, human poverty or corruption.
- The HDI provides a single figure reflecting development across a country. Inevitably this is an average that may not represent the experiences of many people. In many countries, the level of development is greater in cities than in rural areas, especially those that are remote.

Factors affecting growth and development

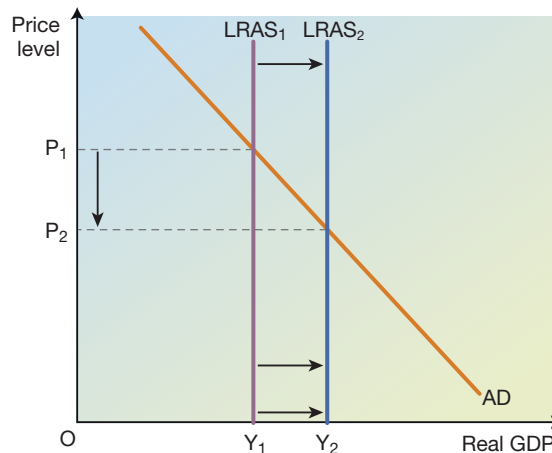
Investment in capital stock

An economy's capital stock includes its infrastructure (roads, telecommunication networks and electricity grid, for example) as well as items used by its firms to produce goods and services. Investment is expenditure on capital stock that can increase the productive capacity of an economy and can raise real GDP.

Investment is a component of aggregate demand and thus an increase in investment will help to boost aggregate demand and therefore result in economic growth. Investment can also initiate the multiplier effect. However, if an economy is close to full capacity, then investment may only serve to generate inflationary pressures.

Investment in capital stock can also increase productive capacity and aggregate supply, shifting the aggregate supply curve to the right, as illustrated in Figure 54.5. Investment in new technology and capital can increase the productive capacity of the economy. An increase in aggregate supply can increase long-term economic growth without adding to inflationary pressures. It reduces the price level from P_1 to P_2 .

Figure 54.5 An increase in long-run supply – the result of an increased level of investment



Investment in the research and development of new goods and services and processes of production can stimulate economic growth by increasing productive efficiency, allowing more goods and services to be supplied using given resources. It can also develop new products that are attractive to consumers and may stimulate demand – for example, in the form of export sales. Less developed economies spend much lower percentages of GDP on research and development. In 2013, Colombia spent 0.23% of its GDP on research and development, Moldova spent 0.35% and Puerto Rico 0.44%. In the same year, France spent 2.23% of its GDP on research and development and Germany spent 2.85%.

Key terms

Human capital is the knowledge and skills possessed by employees that contribute to their ability to create economic value.

Investment is the planned spending by firms that adds to an economy's capital stock (property, machinery, vehicles and other items used in production).

Capital stock is the value of all physical assets used in production in the economy that are still in use, such as machinery, property and vehicles.

Expenditure on training and education

Investing in education and training is a means of improving the abilities of employees that are available to firms and to economies more generally. This is often referred to as investing in human capital.

The concept of 'human capital'

The Institute for Fiscal Studies argues that there are three main components of 'human capital':

- early ability (whether acquired or innate);
- qualifications and knowledge acquired through formal education;
- skills, competencies and expertise acquired through training when in employment.

It is common for employers in developed and less developed economies to make a contribution to the training of their employees to improve their human capital. They hope to benefit in terms of having a more productive and more competitive workforce with the potential to deliver greater profits. This is a form of investment from the perspective of firms. Research by the National Institute of Economic and Social Research found that higher average levels of labour productivity in a range of European businesses were closely related to the greater skills and knowledge of their workforces.

The economic benefits of education and training

The benefits of education and training can spill over to others as well, so that the gains to the economy as a whole provide social benefits. Evidence provided by member states of the Organisation for Economic Cooperation and Development (OECD) indicates that countries investing more heavily in higher education enjoy higher rates of economic growth. The data from the OECD show that primary and secondary education skills are closely related to growth in less developed economies. The economies in sub-Saharan Africa with the poorest statistics relating to education have the lowest levels of development.

We saw in the previous section that investment in both R&D and in increasing the quantity and productivity of capital stock play a vital role in shaping rates of economic growth. It is possible to take a similar view of the importance of investment in human capital as a determinant of economic growth and development. It can also be argued that investment in human capital contributes directly to the degree of freedom experienced by individuals in that it gives them more choice and greater control over their lives. In this way education and training can be important in stimulating the process of economic development.

Barriers to growth and development

Less developed economies have to overcome a number of obstacles if they are to enjoy the benefits of economic growth and development. Governments in less developed countries need to create the right incentives for FDI and domestic investment. Private investment by both foreign and domestic firms contributes to the economic growth that is needed to reduce poverty in developing countries. However, the willingness of firms to invest depends on the business environment – the extent to which the laws, regulations and infrastructure within a country support or limit enterprising activities.

Institutional factors

The absence, or limited development, of a range of institutions within a less developed economy can act as a major barrier to growth and development.

Financial institutions

The development of effective financial markets in less developed economies can be difficult, especially in rural areas. For example, it is expensive to establish and maintain rural branches of banks to encourage savings and there is not an effective channel from savers to those who need funds to invest in businesses. Entrepreneurs find it difficult to access sufficient capital and this shortage of capital can stifle entrepreneurship and thus growth and development.

Similarly, granting loans to people in poor communities is risky. The loans are generally for comparatively small sums, offering banks limited opportunities to generate profits, and the likelihood of debtors defaulting is high. Asymmetric information often exists in financial markets in less developed economies. Financial institutions may have insufficient information on the creditworthiness of borrowers. This can result in higher interest rates for all borrowers, even those who are creditworthy, depressing the level of economic activity.

Ineffective systems of taxation

Taxation is often a significant problem in less developed economies and can represent a major obstacle to economic growth and development. Taxation revenue can be a vital element in promoting economic growth and development because it funds health and education programmes and improvements in an economy's infrastructure.

However, many governments experience difficulties in collecting taxes as they lack information as well as the resources required to collect them efficiently. As a consequence, very little tax revenue is collected in some countries, although many less developed economies do raise substantial sums from indirect taxes which are easier to collect and more difficult to evade. Despite some recent improvements in collecting taxes, about half of sub-Saharan African countries still raise less than 15% of their GDP from tax revenues; in Latin America it is 23%. The average in developed economies that are members of the Organisation for Economic Cooperation and Development (OECD) is approximately 35%. Such low yields from taxation make it difficult for a government to provide very basic services, let alone to build a better economic environment to foster economic growth.

Problems in raising tax revenue efficiently lead governments in less developed economies to use other sources of revenue to fund their expenditure programmes, such as taxes on exports. This reduces the incomes of poor communities in rural areas, hindering their development.

Key terms

Infrastructure refers to the basic facilities available to a society, such as transport and communication links which support economic activity.

Corruption is dishonest or fraudulent behaviour by those in authority and often takes the form of bribery or theft.

A **free rider** is someone who benefits from a good or a service without paying for it.

Lack of infrastructure

Infrastructure is important to help an economy to function effectively. Economic growth and development can be seriously hindered if an economy does not possess workable transport systems and methods of communication as well as supplies of utilities, such as electricity, to most areas. In an economy with poor infrastructure, many households will not be able to avail themselves of vital services provided by schools and hospitals, with damaging consequences for the level of development.

A lack of adequate infrastructure is an impediment to the development of the wider economy. A decent infrastructure is valuable in itself: clean water, for example, brings many health benefits. Investment in it also generates a multiplier effect, enabling further development and allowing governments to achieve social, economic and political aims. In an economy lacking an adequate infrastructure, producers and consumers will find it difficult to access markets either directly or remotely. This will make it less likely that the economy will attract high levels of foreign direct investment. This will limit its rate of economic growth, involvement in trade and the prosperity and wellbeing of its inhabitants.

Poor infrastructure has implications for employment levels. Research by the McKinsey Global Institute in 2013 indicated that an increase in infrastructure investment equal to 1% of GDP could add 3.4 million jobs in India, 1.3 million jobs in Brazil and 700,000 jobs in Indonesia.

Lack of enforceable property and land rights

Property rights are fundamental to the functioning of any market and thus to economic growth and development. However, markets are less efficient when property rights do not exist or are inadequate. The rights to intellectual property such as new designs for products can be stolen if effective laws do not exist.

In many less developed countries, particularly those with low incomes such as Rwanda, fewer than 10% of land holdings are formally surveyed and registered. Most of the land is administered by informal, customary rules and practices. A lack of property rights means that a farmer or other business owner does not have certainty that his or her rights will be protected, and they may not be able to maintain ownership and possession of the property. This, in turn, means that they are unlikely to invest in the business in order to create more efficient production methods because of the potential for free-rider problems. A lack of land and property rights hinders entrepreneurship and, as a consequence, economic growth and development.

Corruption

Corruption is dishonest or fraudulent behaviour by those in authority and often takes the form of bribery or theft. The scale of corruption in less developed economies can be huge. Research in 2014 conducted by the ONE Campaign (an international organisation that combats poverty and disease) estimates that \$1 trillion flows out of less developed economies annually as a consequence of corruption. It identified a range of forms of corruption including illegal tax evasion, money laundering and illegal natural resource deals. A report published in December 2015 by Global Financial Integrity confirmed the scale of corrupt and illicit activities in less developed economies. It found that less developed economies lost \$7.8 trillion in illicit financial flows from 2004 to 2013, with illicit outflows increasing at an average rate of 6.5%

per year – nearly twice as fast as global GDP. Research by PricewaterhouseCoopers indicates that Nigeria’s GDP in 2014 could have been 22% higher if corruption was at the same level as in nearby (and hardly corruption-free) Ghana.

Table 54.3 shows comparative data for different groups of countries with regard to a range of measures of corruption. The incidence of corruption is higher where economies are less developed.

Corruption matters because it diverts funds away from less developed economies in which they could be used to improve infrastructure, pay for healthcare and education or help to develop important institutions such as a fair and effective legal system. The loss of funds on such a large scale is a major barrier to economic growth and development.

Table 54.3 Comparative data on corruption

	Bribery incidence (% of firms experiencing at least one bribe payment request)	Bribery depth (% of public transactions where a gift or informal payment was requested)	Percentage of firms expected to give gifts in meetings with tax officials	Percentage of firms expected to give gifts to secure government contract	Value of gift expected to secure a government contract (% of contract value)
High-income economies (OECD members)	1.7	1.2	0.7	10.7	0.5
East Asia and Pacific	38.9	31.2	29.8	32.5	1.2
South Asia	24.8	21.0	19.6	45.5	2.9
Sub-Saharan Africa	24.0	18.3	17.4	31.1	2.1

Source: World Bank Group

Inadequate human capital

Earlier in this chapter we defined human capital as the knowledge and skills possessed by employees that contribute to their ability to create economic value. Governments and firms in less developed economies have to invest in education and training to develop their human capital. However, education is a merit good and the long-term social benefits are not realised, which can lead to underinvestment. Furthermore, governments may not have sufficient funds for this purpose. Similarly, firms may lack the resources and the willingness to invest in employees who may subsequently leave their employment – this is a form of market failure.

Investment in human capital improves rates of employment and levels of earnings. Employees with low levels of skills have a higher likelihood of poor health and they participate much less in social activities. Research by OECD shows that these results are consistent across a wide range of countries, confirming that skills have a profound relationship with economic and social outcomes. They are also important in tackling inequality and promoting social mobility. Investment in human capital is the single most effective way of not just promoting growth but also distributing its benefits more fairly.

Policies to promote economic growth and development

Governments in less developed economies can adopt two broad approaches to promoting economic growth and development. They can deploy market-based

strategies, such as a greater reliance on privatisation and deregulation. Alternatively, they may opt to intervene in markets through, for example, the imposition of tariffs.

Market-based and interventionist strategies

Market-based strategies are designed to promote economic growth and development by encouraging greater use of market forces to determine the goods and services to be supplied and by allowing markets to operate more efficiently.

Recently there has been an increased level of interest in market-based strategies for growth and development among international organisations such as the World Bank and among aid agencies such as the UK's Department for International Development (DFID). Aid agencies argue that the poor are dependent on market systems for their livelihoods. Therefore changing those market systems to work more effectively will help to reduce poverty.

There are differences within this approach, however. Some economists support the approach of allowing markets to operate entirely freely to enable the fastest possible rates of growth and development. However, the World Bank favours a market-friendly approach. This does advocate the use of free markets but recognises that there are many imperfections in the markets of many developing nations. The Bank therefore believes that some government intervention is necessary to overcome such imperfections. There is no dispute, though, that economic growth and development should broadly be based on market forces.

The alternative to this is to intervene in the operation of markets with the objective of promoting growth and development. This was seen in its most extreme form when economies such as the former Soviet Union developed a system of central planning, whereby the state rather than the market determined the goods and services to be produced. The poor economic performance of centrally planned economies means that most strategies for development entail at least some role for market forces. From 1978 onwards China, previously a centrally planned economy, began to implement a series of market reforms in two stages. Initially the Chinese government allowed individuals to own businesses and farms and it permitted foreign investment. The second stage involved a reduction in the role of the state in industry through privatisation as it moved towards a mixed economy.

There are many policies that can be termed interventionist which do not involve state ownership of industry: developing a country's infrastructure, encouraging expenditure on merit goods such as health and education, and protecting growing industries are all interventionist. These policies can play an important role in promoting economic growth and development.

Key terms

Comparative advantage is when one country can produce a good or service at a lower opportunity cost than another.

Globalisation is the trend for many markets to become worldwide in scope.

Foreign direct investment (FDI) is cross-border investment by a resident entity in one economy with the objective of obtaining a lasting interest in an enterprise resident in another economy.

The **Harrod–Domar model** of economic growth states that the rate of economic growth depends upon an economy's level of savings and the productivity of its investment.

Discussion point

Do you think that globalisation is the major factor influencing which strategy a country adopts to achieve increased rates of economic growth and development?

Import substitution

An import substitution strategy for growth and development substitutes domestically produced goods and services, especially basic necessities such as energy, food, and water, for imported ones. This strategy has been more popular with less developed South American economies and relies upon the use of barriers to trade such as tariffs and quotas to reduce the volume of imports. This represents intervention in the market as the local government is really subsidising relatively inefficient domestic producers, which results in domestic consumers paying prices that can be significantly above global levels.

Import substitution has not been a particularly successful policy. Brazil and Argentina both enjoyed short-term success with this policy, but it tended to end in failure for two main reasons:

- Domestic economies are rarely sufficiently large to allow industries to benefit fully from economies of scale. Globalisation has meant that many industries are worldwide in scope and are dominated by large-scale and highly efficient producers.
- Import substitution does not promote growth or development in the long term because it does not encourage highly efficient firms that can expect to expand over time. Furthermore, import substitution tends to focus on industrialisation and not to consider factors such as poverty or healthcare.

Export promotion

This type of strategy is often referred to as ‘export-led’ growth and has been a very successful model for Asian economies. Economies such as Taiwan and South Korea followed the example of Japan and achieved very high rates of economic growth based on booming exports. The precise approaches differed, however. South Korea, for example, protected domestic firms and provided cheap loans to exporters only. In contrast, China enjoyed rapid growth after opening its economy in 1978 and creating ‘special economic zones’ to attract foreign capital to build factories for export production. More recently, Malaysia, Thailand, Indonesia and Vietnam have all forged export-led paths to growth.

Export-led growth relies upon economies developing a comparative advantage in products for export, primarily to developed economies. Some less developed economies have attempted to develop strong exporting industries by processing primary products and using this as a route to developing competitive industries. More recently, countries such as China and Vietnam have built strong exporting sectors based on low-cost manufacturing using a resource in which they are well endowed – cheap labour. Globalisation has helped the causes of countries seeking export-led growth as it has increased international flows of trade and FDI.

Focus on one or more industries alone to export large volumes of products is unlikely to promote economic growth and development. Many economists argue that other conditions need to be in place as well. These include a stable macroeconomic environment and sufficient levels of capital (either domestic or from FDI) to finance the required level of investment. Governments should not be reluctant to implement other economic policies, such as stimulating continued growth and development by encouraging domestic demand.

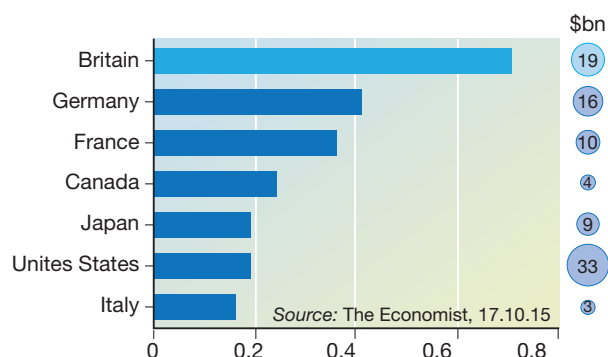
The role of aid and trade in promoting economic growth and development

Foreign aid

Since 1969, foreign aid has been officially termed ‘official development assistance’ (ODA). According to the OECD, it has ‘the promotion of the economic development and welfare of developing countries as its main objective’. Thus the main aim of foreign aid is to promote development rather than merely economic growth.

Foreign aid can be bilateral or multilateral. Bilateral aid is provided by one country to another. For example, the top three recipients of UK bilateral aid in 2013 were Pakistan (£338 million), Ethiopia (£329 million) and Bangladesh (£272 million). Multilateral aid is money from international organisations such as the IMF or the World Bank.

Foreign aid is normally provided on favourable terms. Indeed, some foreign aid is in the form of grants that do not have to be repaid, while other aid may be provided at very low rates of interest. Rich countries gave more than \$130 billion in ODA in 2014. Over the last five decades, western countries spent \$4.14 trillion – the equivalent of more than seven times the 2014 GDP of Nigeria. The UK is one of the world’s major providers of foreign aid, as shown in Figure 54.6.



Key terms

Foreign aid is the voluntary transfer of financial or physical resources from one country or organisation to a recipient country. It is also called official development assistance (ODA).

International trade is the exchange of goods and services across national frontiers.

Figure 54.6 Foreign aid as a percentage of gross national income, 2014 (\$ billion)

Aid from rich countries is supplemented by non-governmental organisations and other private charities. Yet in many of the developing countries that receive the aid, poverty still looms large and evidence of development is slight.

Does foreign aid promote development?

One view is that foreign aid does little to promote either growth or development. Economists cite the example of African countries such as Zambia and the Democratic Republic of Congo. In 2014 the real incomes of Zambia’s citizens were only about 80% of those they received in the late 1960s. In the Democratic Republic of Congo, real incomes have been falling for many years. While the typical African country has received about 5% of its national income as foreign aid, Zambia on average has received double that level. Of course, there are other factors that have shaped the poor performance of these economies in alleviating poverty. For example, the Democratic Republic of Congo has suffered a civil war for many years.

There is evidence that foreign aid does not promote development either, although this is incomplete. It does not appear to result in a more equal distribution of incomes in

poor countries, according to a study in 2012 by Dierk Herzer and Peter Nunnenkamp. This study compared 21 countries over 26 years. It found that countries that received more foreign aid developed a more uneven distribution of income. This supports the view that foreign aid mainly helps rich people in poor countries. While more work on poverty reduction is necessary, there is evidence that foreign aid may tend to end up in the hands of the wrong people.

Some economists hold very different views. The OECD has identified that ODA is directed towards countries coping with the most difficult and intractable development problems, including emergency situations arising from natural disasters, conflict and refugee influxes, and not countries with the highest potential investment returns. Many receivers of development assistance are therefore not fast-growing countries and they may not have the ability to increase incomes and remove people from poverty quickly.

The debate on the effectiveness of foreign aid will continue. The opinions of economists and the general public are diverse and it is an area on which more research is required.

Author tip

Chapter 51 looked in depth at the arguments for and against international trade. You may wish to review the relevant sections of this chapter alongside the text below.

International trade

Globalisation has been accompanied by a liberalisation of trade rules. Markets have become more global in scope, offering potential benefits to many countries. However, there are arguments that increased international trade does not benefit less developed nations to the same extent as it does developed economies. While globalisation has resulted in increased volumes of trade, the reduction of some tariffs has, some economists believe, adversely affected the industrial setups of the less developed and developing economies. Freer trade has resulted in the closure of some growing industries in these countries, since they have been unable to compete in global markets. These economists also believe that trade deals have favoured the richer nations. Certainly the latest of the World Trade Organisation's series of negotiations (the so-called Doha Round) has not delivered some of the outcomes desired by the poorer countries. Their access to the markets of certain rich countries remains limited. At the same time, developed economies continue to subsidise agricultural production within their economies. The US government paid \$6 billion (£4.2 billion) in subsidies to its cotton farmers in 2014, thereby undermining prices and markets for producers in West Africa. American farmers received \$41 billion (£28.5 billion) in subsidies in total in 2014, while those in Europe received \$86 billion (£59.7 billion).

Moreover, not all countries have enjoyed large rises in the volume of trade. In 2014, Africa accounted for less than 2% of world trade. Critics cited dumping by rich countries as a major problem for the continent's producers. However, African producers face a more fundamental problem. They lack the infrastructure that is necessary to exploit global market opportunities. Thus poverty remains, providing an insurmountable barrier to development.

The alternative view is that international trade can have a positive impact on economic growth in less developed economies. Sustained, strong economic growth over longer periods is associated with poverty reduction, while trade and growth are closely linked, as illustrated in Figure 54.7. According to this viewpoint, developing countries that have followed trade liberalisation policies have experienced all the favorable effects of globalisation and international trade. China and India are regarded as prime examples.

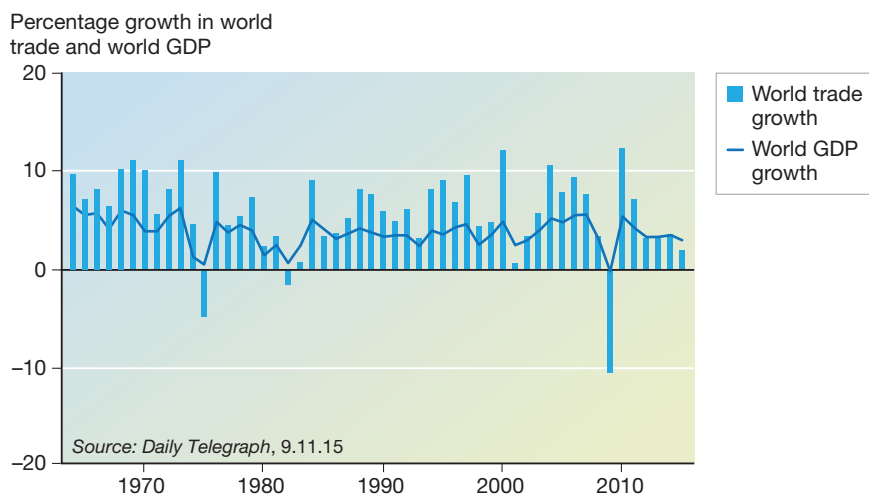


Figure 54.7 Growth in world trade and world GDP, 1964–2014

Christine Lagarde, the IMF’s Managing Director, believes strongly in the power of trade to promote economic growth. In a speech in Washington in 2015, she argued that reforms to increase trade can have indirect effects as well as more direct ones on economic growth: ‘Trade reforms can also have a powerful indirect effect on growth by igniting and amplifying other structural reforms. For example:

- Trade reforms can increase external competition in product and services markets.
- They can encourage key infrastructure investments – think of new ports and new roads.
- They can spur innovation through R&D and “learning by exporting”.
- And they can strengthen institutions by encouraging better governance and a better business environment.’

Trade has the potential to increase rates of economic growth and also to stimulate economic development, but only if it is conducted fairly. It can allow the most efficient and entrepreneurial producers in less developed economies to target global markets and to benefit from economies of scale. It can allow them access to the latest technology and techniques of management. It can also enhance efficiency and incomes.

However, trade alone will not promote economic growth. It has to be supported by a host of other factors including macroeconomic and political stability. There are further factors that determine whether economic growth will be translated into development. These include the income distribution within an economy, the degree of corruption and the extent to which the government is committed to reducing poverty.

Review questions

Total: 35 marks

- 1 Explain the difference between growth and development. (5 marks)
- 2 Distinguish between relative and absolute poverty. (4 marks)
- 3 Which of the following is a reason why it is difficult to compare living standards using GDP data based on official exchange rates?
 - A Official exchange rates are based solely on trade flows
 - B Official exchange rates tend to be stable over periods of time
 - C Official exchange rates are influenced by the activities of speculators
 - D Official exchange rates only reflect relative price levels in different countries(1 mark)
- 4 Explain why less developed economies may suffer from shortages of capital. (6 marks)
- 5 Which of the following is not a component used in the calculation of the Human Development Index (HDI)?
 - A Expenditure on healthcare
 - B Longevity
 - C Education
 - D Average standard of living(1 mark)
- 6 A country classified as having reached 'high human development' will have an HDI figure in which of the following ranges?
 - A Below 0.550
 - B Between 0.550 and 0.699
 - C from 0.700 to 0.799
 - D 0.800 and above(1 mark)
- 7 Explain what is meant by the term 'human capital'. (4 marks)
- 8 Explain why a lack of enforceable property rights might hinder economic development. (6 marks)
- 9 Distinguish between foreign aid and foreign direct investment (FDI). (6 marks)
- 10 A model of export promotion (or export-led growth) depends upon which of the following?
 - A The use of purchasing power parity (PPP) exchange rates
 - B The imposition of tariffs and quotas
 - C An increased use of central planning
 - D The development of comparative advantage(1 mark)

Topic 14 Exam-style questions

A-LEVEL PAPER 2

SECTION A – Context: International trade

Extract A International trade slows

Trade growth in recent years has been weak in advanced economies, particularly in the eurozone, and moderate in less developed economies. Export growth (as shown in Figure A) varied more across regions, but follows a pattern of stronger growth in less developed economies relative to advanced economies. Growth in African exports has been positive: for example, Ethiopia's exports reached a value of \$5.70 billion in 2014, compared with \$3.98 billion in 2010.

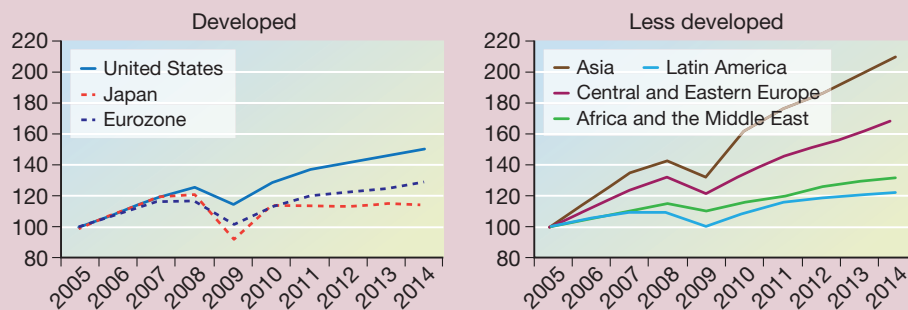


Figure A Export volumes for developed and less developed economies, 2005–14 (2005 = 100)

Source: Centre for Economic Policy Research

Extract B Deadlock in trade talks

Launched in 2001 and intended to deliver a bold new world trade order, the World Trade Organisation's Doha talks have stumbled from one deadlock to another. The collapse of the Doha Round is likely to prompt the USA and the EU to pursue bilateral trade agreements such as the Transatlantic Trade and Investment Partnership (TTIP).

There has been little or no progress in the Doha Round of talks since 2008, when developed and less developed countries found themselves arguing about whether poor countries should be allowed to increase tariffs to support their farmers if food imports rose sharply. The USA believes that since the last trade round was concluded more than 20 years ago, a number of developing countries have made such rapid progress that they should no longer be given the preferential treatment reserved for poor nations.

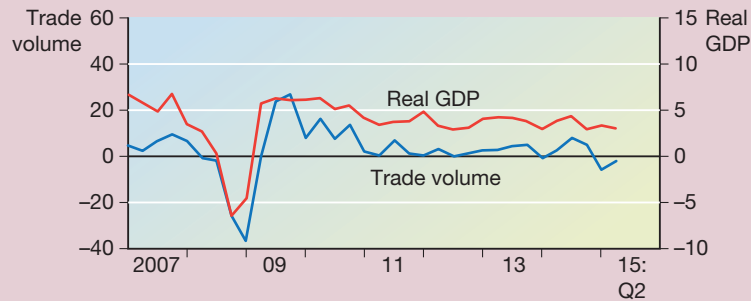
Extract C World trade grows less quickly than world GDP

Many economists and organisations such as the IMF see increasing trade as the key to ending extreme poverty and boosting shared prosperity. Evidence from the IMF shows that countries open to international trade tend to grow faster and provide more opportunities for their populations.

Trade has been a key driver of global growth, income convergence and poverty reduction. Both developing countries and emerging market economies have benefited from opportunities to transfer technology from abroad and to undergo domestic structural transformation via trade integration in the last decades. Yet, more recently, concerns have been raised over whether the current pace and direction of world trade will lead towards less development-boosting potential.

World trade suffered another disappointing year in 2015, experiencing a contraction in merchandise trade during the first half and only low growth during the second half. World trade volumes have lagged behind GDP growth, as illustrated in Figure B.

Figure B World trade volume and real GDP (annualised quarterly percentage change)



The period from the mid-1980s to the mid-2000s was peculiar in several respects. These decades featured economic reforms that aimed to remove barriers to trade, a multilateral trading system that reduced uncertainty for traders, and technological advances that reduced trade and communication costs. Combined, these trends ushered in years of sustained trade expansion.

Questions

Total: 40 marks

- Using the information in Extract A, calculate the percentage increase in the value of Ethiopia's exports between 2010 and 2014 to two decimal places. (2 marks)
- Explain one weakness of the data in Figure A in comparing the trade performances of the United States and Asia. (4 marks)
- Extract B states that '... developed and less developed countries found themselves arguing about whether poor countries should be allowed to increase tariffs to support their farmers if food imports rose sharply'. Explain the likely effects of the imposition of tariffs on the economic welfare of less developed countries. (9 marks)
- Using the data in the extracts and your economic knowledge, evaluate the importance of increased levels of trade to achieving higher levels of development in less developed economies. (25 marks)

SECTION B Essays

Total: 40 marks

Incomes in less developed countries have been converging with those of rich countries since the 1990s because growth has accelerated in developing economies, while in developed economies it has slowed down. The UK's balance of payments on current account reached a record level in 2014 – a deficit of around 5.6% of GDP.

- Explain the possible consequences of globalisation for a less developed economy. (15 marks)
- Evaluate the view that allowing a currency to float freely is the best way for governments to avoid imbalances on the current account of their balance of payments. (25 marks)

Exam-style questions

A-LEVEL PAPER 3

SECTION A

Total: 30 marks

Answer *all* questions in this section.

- 1 Which one of the following statements is a principle of behavioural economics?
- A People are not influenced by other people
 - B People are rational
 - C People base decision making on habit
 - D People enjoy taking risks
- 2 Choice architecture means:
- A Influencing choices but without removing freedom of choice
 - B Presenting options in a certain way in order to encourage a given response
 - C The different ways in which options can be presented
 - D Where individuals can only select from a limited range of options

(1 mark)

(1 mark)

- 3

Units of labour	Total returns	Average returns
1	10	10
2	24	12
3	42	14
4	54	13.5
5	?	12
6	58	9.7

 Table A

Units of labour	Total returns	Average returns
1	10	10
2	24	12
3	42	14
4	54	13.5
5	?	12
6	58	9.7

Based on the data shown in Table A, where labour is the only variable input, the marginal return from the fifth unit of labour is:

- A 2 units
 - B 3 units
 - C 6 units
 - D 12 units
- 4 Which pair of cost curves shows a decrease and then an increase as output rises?
- A ATC and MC
 - B AFC and TC
 - C ATC and AFC
 - D MC and TC

(1 mark)

(1 mark)

- 5  Figure A

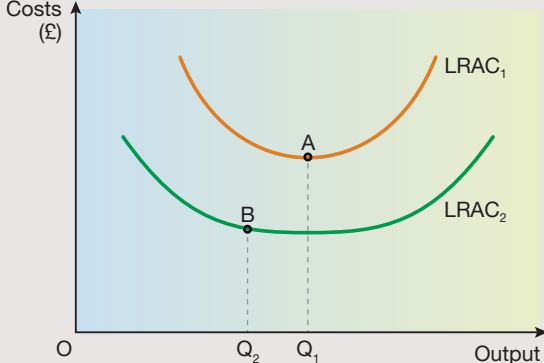
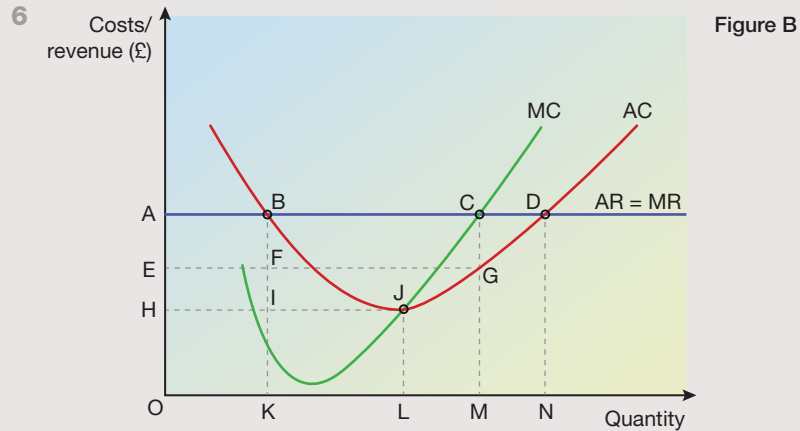


Figure A shows how technological change has led to a shift in a firm's long-run average costs from $LRAC_1$ to $LRAC_2$. This change shows:

- A Increased efficiency and an increase in the minimum efficient scale
- B Increased efficiency and a decrease in the minimum efficient scale
- C Decreased efficiency and an increase in the minimum efficient scale
- D Decreased efficiency and a decrease in the minimum efficient scale

(1 mark)



In Figure B, the firm's abnormal profit is shown by the area:

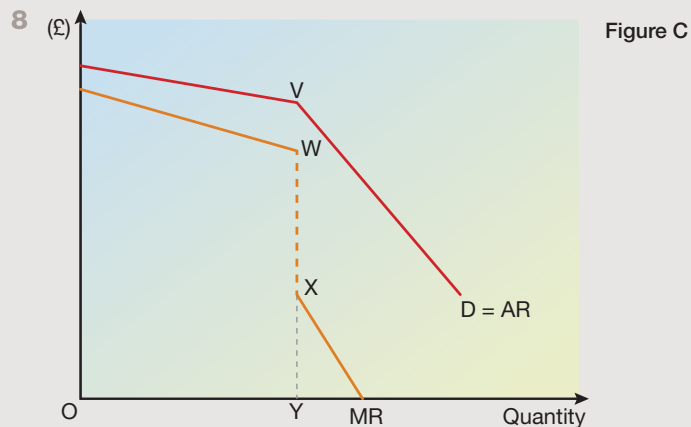
- A A C M O
- B E G M O
- C A C G E
- D B J D

(1 mark)

7 Which *one* of the following statements is true? Monopolistic competition:

- A Achieves productive efficiency in the short run
- B Achieves allocative efficiency in the long run
- C Achieves both productive and allocative efficiency in the long run
- D Does not achieve productive efficiency or allocative efficiency

(1 mark)



Based on Figure C, the firm should exit the industry if its average total costs (ATC) at OY units of output are greater than:

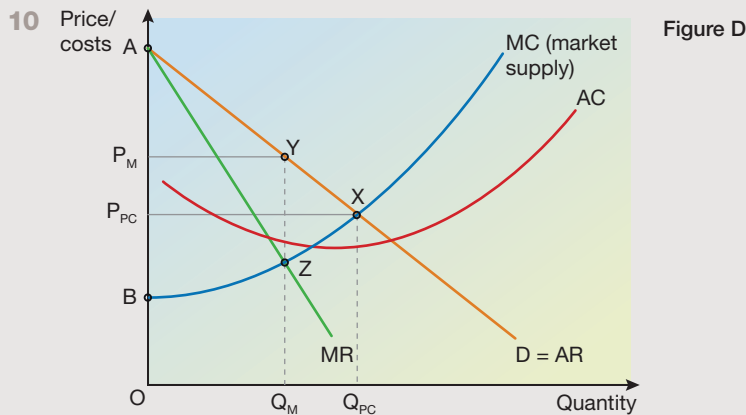
- A VY
- B WY
- C WX
- D XY

(1 mark)

9 Productive efficiency occurs where:

- A $MC = MR$
- B $ATC = MC$
- C $P = MC$
- D $ATC = AR$

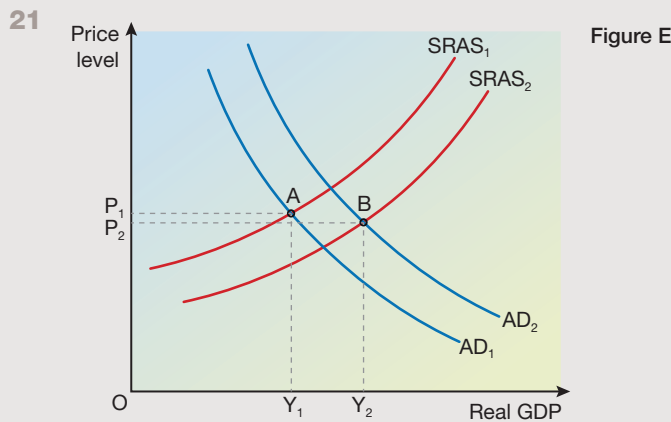
(1 mark)



Based on Figure D, which area shows the deadweight welfare loss under monopoly?

- A AYP_M B $P_{PC}XB$ C AXB D YXZ (1 mark)
- 11** Which one of the following will cause a movement along the demand curve for labour by a car manufacturer?
- A A change in consumers' incomes
 B A change in the price of cars
 C A change in the price of motorbikes
 D A change in the price of labour (wages) (1 mark)
- 12** The supply of labour for a particular job is price elastic. This is most likely to be the result of:
- A A long period of training required by workers
 B Employers demanding a specific qualification for employment
 C The job being unskilled
 D Very unpleasant working conditions (1 mark)
- 13** Bilateral monopoly occurs in a market when:
- A There are only two buyers
 B There are only two sellers
 C There is only one buyer and one seller
 D There are two buyers and two sellers (1 mark)
- 14** If individuals with different circumstances are treated differently, but with overall fairness, this is known as:
- A Horizontal equality
 B Vertical equality
 C Horizontal equity
 D Vertical equity (1 mark)
- 15** Which one of these examples of market failure leads to *overproduction* of goods?
- A Goods with negative externalities of consumption
 B Immobility of factors
 C Monopsony power
 D Public goods (1 mark)

- 16** Which of the following represents the difference between gross domestic product and gross national product?
 A Depreciation
 B Net property income from abroad
 C Income paid to owners overseas
 D Income generated by the shadow economy (1 mark)
- 17** Assume that an economy experiences a rise in income from £1000 billion to £1150 billion. As a consequence, consumption rises from £800 billion to £905 billion. Which of the following is the marginal propensity to consume for this rise in income?
 A 0.80 B 0.79 C 0.90 D 0.70 (1 mark)
- 18** Keynesian economists believe that the interaction of the accelerator and the multiplier models can result in which of the following?
 A The operation of the economic cycle
 B Large falls in net migration
 C The economy avoiding negative and positive output gaps
 D Steady rates of economic growth in the long term (1 mark)
- 19** The imposition of a minimum wage rate above the existing market wage rate is most likely to result in which of the following?
 A A rise in real wage unemployment
 B A fall in real wage unemployment
 C A rise in voluntary unemployment
 D A rise in frictional unemployment (1 mark)
- 20** Which of the following statements are true?
 (i) Commercial paper is sold on capital markets in the UK
 (ii) The coupon on a Treasury bill is zero
 (iii) A direct relationship exists between market interest rates and bond prices
 A (i) and (iii) B (i) and (ii) C (ii) only D (iii) only (1 mark)



An economy's aggregate demand and supply curves have shifted, changing the equilibrium from Point A (P_1, Y_1) to Point B (P_2, Y_2) in Figure E. Which of the following offers the best explanation of these changes?

- A A rise in oil prices and a rise in interest rates
- B A fall in wage rates and a rise in interest rates
- C A rise in wage rates and a fall the rate of income tax
- D A fall in commodity prices and a fall in interest rates

(1 mark)

22 Which of the following statements is true in relation to the balance sheet of a UK commercial bank?

- A The majority of the bank's liabilities are loans to UK households and firms
- B Wholesale funding represents a liability of the bank
- C A majority of the bank's assets will be highly liquid
- D Its assets must exceed its liabilities on every trading day

(1 mark)

23 A rise in the bank rate of interest in the UK is most likely to result in which of the following scenarios?

- A A fall in the price of bonds and a rise in the price of imports
- B A rise in the price of bonds and a fall in the price of exports
- C A rise in the price of bonds and a rise in the price of exports
- D A fall in the price of bonds and a rise in the price of exports

(1 mark)

24 Which term describes circumstances in which one person or organisation takes greater risks because third parties carry the burden of those risks?

- A Hysteresis
- B Systemic risk
- C Moral hazard
- D Hypothecation

(1 mark)

25 A government takes a decision to increase taxes and reduce the level of its spending for the next financial year. Which of the following best describes the outcome of this situation?

- A Its budget will be in surplus. The economy's national debt will be reduced.
- B Its budget balance will improve. The economy's national debt will rise.
- C Its budget balance will improve. The economy's national debt will be reduced.
- D Its budget balance will improve. The economy's national debt might increase or be reduced.

(1 mark)

26 Exetonia and Plymotho can only produce apples and pork. The maximum levels of production are shown below.

Country	Units of output using all resources to produce apples	Units of output using all resources to produce pork
Exetonia	10,000	16,000
Plymotho	9000	6000

Which of the following statements is true?

- A Trade is worthwhile because the opportunity cost ratios of producing the two products differ between Exetonia and Plymotho
- B There will be no benefit from trade because Exetonia has an absolute advantage in the production of both products
- C World output will be maximised if Exetonia produces pork and Plymotho produces apples
- D Exetonia will not benefit from trade, but Plymotho will

(1 mark)

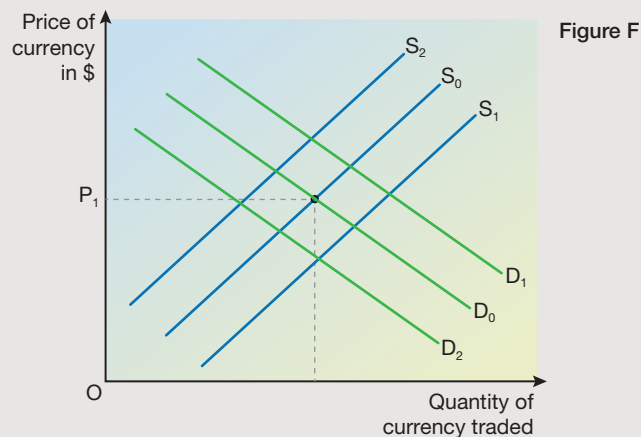
- 27** The Marshall-Lerner condition states that in which of the following situations would a depreciation of a currency improve the economy's balance of trade?

(1 mark)

	A	B	C	D
Price elasticity of demand for exports	-0.5	-0.4	-0.6	-0.3
Price elasticity of demand for imports	-0.3	-0.5	-0.5	-0.4

- 28** The imposition of a tariff on an import for which demand is strongly price elastic is likely to have which of the following effects in the medium term, ceteris paribus?
- A Improve the country's balance of trade, but weaken the government's budget balance
 - B Improve the country's balance of trade and improve the government's budget balance
 - C Weaken the country's balance of trade and weaken the government's budget balance
 - D Weaken the country's balance of trade, but improve the government's budget balance
- 29** An economy's floating exchange rate (P_1) is initially determined by the demand for and supply of its currency, as shown by D_0 and S_0 in Figure F. Assume that there is an inflow of hot money to the economy and that purchases of imports fall. There are no other changes.

(1 mark)



Which of the following combinations of demand and supply curves for currencies is most likely to be the outcome?

- A D_1 and S_1
 - B D_1 and S_2
 - C D_2 and S_1
 - D D_2 and S_2
- 30** Which of the following is the main reason for the use of purchasing power parity exchange rates for international comparisons of living standards?
- A Capital and speculative flows affect official exchange rates
 - B Not all countries use floating exchange rates
 - C It assumes more equal distributions of income
 - D It allows for economies being at different stages of development

(1 mark)

(1 mark)

SECTION B

Total: 50 marks

Case study **Development in Asia**Extract A **Banking concentration ratios in Vietnam**

Year	1999	2004	2009	2014
Five-firm concentration ratio (%)	93	85	65	50

Extract B **India's national minimum wage**

In 2016, the government of India is proposing an amendment to the Indian Minimum Wages Act of 1948. Its proposal is to set a legally binding national minimum wage, below which no worker can be paid, across all Indian states and industries.

Currently there is a national floor for the minimum wage – Rs 160 per day. However, it is only an advisory minimum wage and is not legally binding. Furthermore, it is much lower than the level advised by analysts of the labour market in India and by the government's own advisers.

The Indian government believes that a legally binding national minimum wage will lead to better compliance – at present only 61% of workers in India get paid at or above the level of the minimum wage.

The government's proposal is to set a monthly minimum wage of Rs 7100. However, at present the advisory minimum wage in India varies considerably between different places. For example, in Haryana, the minimum wage is currently Rs 7600 per month, while in Delhi it is Rs 9178 per month. The government's proposal will only help workers in states such as Uttar Pradesh, where the minimum wage is currently fixed at Rs 6814 per month.

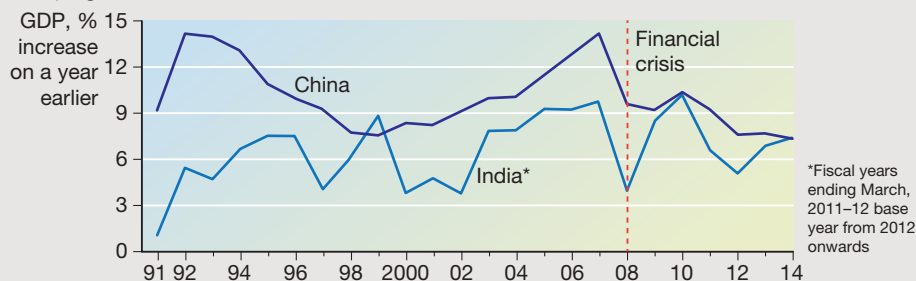
Extract C **Data relating to growth and development for India and Vietnam**

Indicator		India	Vietnam
GDP, 2014 (\$ billion)		2049.0	186.2
GDP growth rate, 2014		7.3%	6.0%
Per capita income (\$ PPP), 2011		5497	5092
HDI 2014		0.609	0.666
Annual percentage HDI growth:	1990–2014	1.48	1.41
	2010–14	0.97	0.47
GDP per capita (\$ PPP), 2011	Male	8656	5570
	Female	2116	4624
Gini coefficient, 2013		43.0	35.6
Percentage of population living on less than \$1.25 PPP in 2012		23.6	2.4
Average annual population growth rate (%)	2000–05	1.6	1.0
	2010–15	1.2	1.0
Percentage of GDP spent on health 2013		4.0	6.0
Percentage of GDP spent on education, annual average 2003–14		3.8	6.3
Percentage of population living on degraded land, 2010		8.0	9.6

Source: United Nations Development Programme, Human Development Reports

Extract D Chasing the Dragon?

Official statistics published recently revealed that India's GDP rose by 7.5% in 2014, a shade faster than China's economy managed over the same period. India is one of Asia's most rapidly developing economies.



India has been a rare bright spot among less developed economies. Foreign investors remain keen buyers of Indian assets, even as they opt not to invest in other less developed economies. The value of the Indian currency – the rupee – is firm. The central bank has even expanded its foreign exchange reserves to a record \$330 billion – thus keeping the rupee from rising by more.

The economy is likely to pick up further. The recent falls in commodity prices, which have hurt raw material exporters such as Brazil, Russia and South Africa, are a boon for India, which imports 80% of the oil it consumes. Rich economies may fret about the dangers of falling prices around the world. The diminishing threat from inflation has already prompted India's central bank to reduce interest rates in January, from 8% to 7.75%. More cuts in rates are expected to follow.

India has benefited from being an attractive location for multinational corporations, such as Microsoft, Nestlé, Sony and Ford, and it has recently been chosen as a location by emerging multinationals such as Huawei. Over the last year a new Indian government has implemented a number of business-friendly economic policies, such as making it easier to demonstrate property rights and to invest directly in India. Its education system produces hundreds of thousands of engineers and technical graduates annually. The country also has an advantage in low employment costs and widely available technical skills.

Source: Adapted from *The Economist*, 9.2.15

Answer *all* questions in this section.

You are an economist employed by an international organisation analysing the extent of economic growth and development in Asia. The international organisation has asked you to write a report addressing three questions. Refer to **Extracts A, B C and D** and then use these and your own economic knowledge to help you answer Questions 31–33.

- 31** Using Extract A and your own economic knowledge, explain how the changes in concentration ratios shown in Extract A are likely to have impacted on the banks and their customers. (10 marks)
- 32** Using Extract B and your own economic knowledge, analyse the possible consequences for the Indian economy of a legally binding national minimum wage throughout India. (15 marks)
- 33** India's economy has performed well recently and some economists believe that it is more developed than many of its Asian rivals. Use the information in Extracts C and D to analyse the view that India's economy is performing more strongly than that of Vietnam and make a judgement on whether you think this is the case. Justify your judgement. (25 marks)

Index

Definitions of key terms have been highlighted in **bold red** type.

A

accountants' wages 159

aggregate demand

- circular flow of income **262**, 263
- consumption function **262**, 263–264
- marginal propensity to consume **262**, 263, 264–266
- marginal propensity to withdraw **267**
- multiplier effect 265, **265**, 266–268
- reverse multiplier 268

aggregate supply **270**

- institutional structure impact on 271
 - banking system 271, 272
 - labour market 272–274
- long-run aggregate supply 270
 - Keynesian economics **274**, 275, 276–277
 - long-run aggregate supply curve 274–275

airline pricing 10, 38

airports

- long-run average cost curve 43

allocative efficiency 57, 82, 83, **125**, 125

see also **efficiency**

altruism **17**, 18

anchoring 16, **16**, 17

Apple iPhone 63–64

asymmetric information **11**, 12, 13, **365**

auctions 137

austerity **379**

automatic stabilisers **386**

availability bias 17, **17**

B

balance of payments 439, **439**

balancing 443

capital account 441

consequences of investment flows

- between countries 444–445

current account 439

balance of trade **449**

investment income 440–441

total current transfers 441

trade in goods 439–440

trade in services 440

current account imbalances

446–447

current account deficits **447**

current account surpluses **447**

expenditure-increasing policies 448

expenditure-reducing policies 447–448, **448**

expenditure-switching policies 448, **448**, 449–451

deflationary policy **448**

expenditure-switching policies 448, **448**, 451

import controls 449

reducing the value of the currency 449–451

financial account 441

flows of reserve assets 443

hot money 442, **442**

net foreign direct investment 442, **442**

net portfolio investment 442, **442**

other investment 442

implications of major economies with

global imbalances 451–452

eliminating a deficit 452

eliminating a surplus 452–453

'J-curve effect' 450, 451

marginal propensity to import **448**

Marshall–Lerner condition **449**, 450

Bank of England

central bank money **357**

functions of 351, 352

banker to other banks and the government 353

'lender of last resort' 352–353

managing the money supply 352

regulating the banking system 353, 361–362

resolution **362**, 362

setting the rate of interest 352

influence the supply of money

Funding for Lending Scheme 356–357, 358

interest rates 356

open market operations 357, **357**, 358

quantitative easing **358**, 358, 359

Monetary Policy Committee **351**, 352

see also **monetary policy**

banking industry

aggregate supply, and 271, 272

competition 139–140

divorce of ownership from control 74

foreign banks 338

investment banks **338**, 339, **340**

leverage ratio **271**

loss-bearing debt 348

regulation 236–7

systemic risk 339–340, **340**

see also **Bank of England;**

commercial banks; regulation of the financial system

barriers to entry

oligopoly 101

technology 61

barter **324**, 327

behavioural economics 14

altruism **17**, 18

bias in decision making 16, **16**, 17, **17**

bounded rationality 15, **15**

bounded self-control 15, **15**, 16

choice architecture **20**, 20–21, 22

default choice 21, **21**

framing 21, **21**

mandated choice **21**, 22

nudges 21, **21**, 22

principles of 14–15
 restricted choice 21, **21**
 sense of fairness 18, 19
 taxi drivers 18
see also **consumer behaviour**

bias in decision making 16, **16**, 17, **17**

'big data'
 price discrimination 114

Big Mac Index 257

bilateral monopoly **162**, 162, 169

Billingsgate Market 83–84

bonds **332**, 332, 333
 coupon rate **332**
 maturity **332**
 relationship between market interest rates and bond prices 333–334

bounded rationality 15, **15**

bounded self-control 15, **15**, 16

budget balance *see* **fiscal policy**

C

capital account 441
see also **balance of payments**

capital markets **329**, 332–334

car industry
 minimum efficient scale 43, 44
 productivity and wages 146

cartels 96, 97, **97**, 100, 101

central banks **351**
see also **Bank of England**

charities 214

China
 economic growth 286

choice architecture **20**, 20–21, 22

circular flow of income **262**, 263

collective bargaining 166

commercial banks 338, **338**, 339
 balance sheet 343, **343**
 assets **338**, 344
 capital **343**, 344, **365**
 liabilities 343, **343**, 344
 bank failure, causes of
 borrowing short and lending long 364–365

engaging in risky investment
 banking activities 365
 Royal Bank of Scotland 367
 trading with insufficient capital and liquidity 365

banks creating credit 345, 346
 credit multiplier **346**, 346
 limits on 346

base money **346**

capital requirements 366, 367–368

core tier one capital 367, 368, **368**

functions of
 accepting deposits 341, **341**, 342
 advancing loans 342
 financial services, provision of 342
 investing funds 342

global systemically important bank **368**, **368**

investment banking activities,
 separation of 341

liquidity requirements 366
 liquidity coverage ratio **344**, 347, 366
 net stable funding ratio 366
 'maturity transformation' 342

moral hazard **369**, 369–370

objectives of 347, **347**, 348
 conflict between objectives 349

reserve assets ratio 346
 retained profit **368**

commercial paper **330**, 331

commodities **308**
 fluctuations in commodity prices 471
 world commodity prices and domestic inflation 308–310

communications market 104

comparative advantage 418, **418**, 419–421, **481**

competition
 banking industry 139–140
 creative destruction **116**, 116, 117
 long-run benefits 116
 short-run benefits 115–116

see also **perfect competition**

competition policy **225**, **432**
 benefits of 227
 criticisms of 228–229
 EU policy 226–227
 UK policy 225–226

consumer behaviour
 economic incentives 7, 7
 rational economic decision making 6, 7
 utility theory 7–10
see also **behavioural economics**

consumer surplus **134**, 134, 135
see also **surplus**

consumption **262**

consumption function **262**, 263–264

corporation tax 383–384

costs of production
 airlines
 marginal cost pricing 38
 average costs 35, 36, **36**
 choice of factor inputs
 factor prices 39
 productivity 39
 factor prices, impact of 37, 38
 long run 37
 marginal costs 35, 36, **36**
 productivity and 38, 39
 returns to scale and average costs 37, 38
 short run 36, 37
 technological change, impact of 57–58
 total costs 35, 36, **36**, 37

creative destruction **59**, 59–60, 61, **116**, 116, 117

creditors **347**

crowding out **386**, 388

currency unions **464**, 464, 465–467
 eurozone 464–465

current account *see* **balance of payments**

customs unions 429, **429**, 430–431

cyclical unemployment **293**

D

deadweight welfare loss 136

debt 335, **335**

default choice 21, **21**

deindustrialisation **412**, 412–413

deregulation **234**, 234–235, **392**,
393

derived demand **142**, 142

discrimination

direct discrimination 177

impact of 179

perceptions of 181

types of 181

wage discrimination 177, **177**

causes of 178–179

conditions necessary for 178

‘equal work’ 177

gender pay gap 179–180

impact on workers, employers,
and the economy 180–181

see also **price discrimination**

distribution of income and wealth

disposable income 186

distribution of income **187**, **251**

factors influencing 187–188

UK 189, 190

distribution of wealth **187**

factors influencing 188–189

UK 191

equality 191, **191**

equity 191, **191**

horizontal equity 191–192

vertical equity 192

globalisation 414

income 186

inequality 187, 194

benefits of an unequal distribution

of income 194–195, 196

costs of an unequal distribution of
income 195–196

economic growth, and 286

Gini coefficient 193, **193**, 194

Lorenz curve **192**, 192–193

OECD countries 196–197

wealth 186

see also **poverty; redistribution of
income and wealth**

dividends **347**

donations 214, 215

double coincidence of wants **324**

duopoly 66

dynamic efficiency 58–59, 122, **122**,
123, 131–132

see also **efficiency**

Dyson 397, 406–407

E

‘earnings trap’ 205

economic cycle **287**

causes of changes in 289

asset price bubbles and

destabilising speculation **287**,
290–291

economic shocks 289–290

excessive growth in credit and

levels of debt 291

stages of 288–289

economic growth **248**, 282, **282**,

283

accelerator theory of investment **287**

benefits of 283–284

China 286

demand-side shocks **287**

disadvantages of **284**, 284–286

herding instinct **287**, 290–291

impact of 287, 288

long-run economic growth **282**

short-run economic growth **282**

supply-side shocks **287**

economic growth and economic

development 468–469

barriers to growth and development
477–478

corruption **478**, 479–480

free rider problems **478**, 479

inadequate human capital 480

ineffective financial markets 478

ineffective systems of taxation

478

infrastructure, lack of **478**, 479

lack of enforceable property and
land rights 479

characteristics of less developed

economies 469–473

economic development **409**, **468**

economic growth **468**

factors affecting growth and
development

investment in capital stock 476,
476

investment in education and
training 477

fluctuations in commodity prices 471

foreign aid 483, **483**, 484

Harrod–Domar model **481**

human capital **476**, 477, 480

international trade, role of **483**,
484–485

measures of development 473

Human Development Index 257,
473–475

policies to promote 480–481

export promotion 482

import substitution 482

interventionist strategies 481

market-based strategies 481

see also **standard of living**

economic incentives 7, **7**

economic man 6

economies of scale

globalisation **405**, 406–407

internal diseconomies of scale 40

internal economies of scale 40

long-run average cost curve 41–42,
43

minimum efficient scale **42**, 42, 43,
44

returns to scale, and 40, 41

efficiency **122**

allocative efficiency **125**, 125

dynamic efficiency 122, **122**, 123,
131–132

market contestability 119–120

market structures 130

- monopolistic competition 89, 127–128
 - monopoly 129–130
 - oligopoly 128–129
 - perfect competition 82–83, 125–127
 - National Lottery 131
 - productive efficiency **122**, 123–125
 - static efficiency 122, **122**, 123
 - technical efficiency **122**, 123
 - technological change 57, 58–59
 - see also **surplus**
 - electronic retailing 402**
 - employment and unemployment**
 - consequences of unemployment 299–301
 - cyclical unemployment **293**
 - economic growth
 - increased levels of employment 284
 - frictional unemployment **293**, 296, **396**
 - hysteresis **300**, 300
 - involuntary unemployment 293, **293**, 294
 - labour mobility **296**, 297–298
 - natural rate of unemployment **296**, 296–298, **314**, 396, **396**, 397–398
 - non-accelerating inflation rate of unemployment 298, 300, **314**, 316, 318
 - rate of unemployment **311**
 - real wage unemployment 294, **294**, 295–296
 - regional unemployment 422
 - seasonal unemployment **293**
 - structural unemployment **293**, 296, **396**, 422
 - unemployment **293**
 - voluntary unemployment 294, **294**
 - see also **Phillips curve**
 - energy market 48–49**
 - equality 191, 191**
 - see also **discrimination;**
 - distribution of income and wealth**
 - equity**
 - business capital 335, **335**
 - distribution of income and wealth 191–192, **191**
 - EU law**
 - precedence over UK law 227
 - European Commission 432**
 - European Union 431–432**
 - consequences of EU membership for the UK 433–435
 - Single European Market 432–433
 - European Union Emissions Trading System (EUETS) 241**
 - eurozone 423**, 464–465
 - exchange rates 455, 455**
 - Big Mac Index 257
 - bilateral rate 456
 - divergence in monetary policy
 - causing currency volatility 459
 - fixed exchange rate **455**
 - fixed exchange rate systems 456–457, 461–463
 - floating exchange rate **455**
 - floating exchange rate systems 457–458, 463–464
 - foreign exchange market 334, **334**, 335
 - freely floating exchange rates 459
 - government intervention
 - ‘dirty’ floating 460–461
 - managed floating exchange rates 459, 460
 - market exchange rate 255, **255**
 - multilateral rate 456
 - purchasing power parity exchange rates 255–256, **469**
 - spectrum of exchange rate systems 456
 - trade-weighted index **455**, 456
 - see also **currency unions**
 - externalities 219**
 - externalities in consumption 220–221
 - externalities in production 221–223
 - negative externalities 219, **251**, **284**, 284–285
 - noise pollution 223
 - positive externalities 219
- ## F
- fairness, sense of 18, 19**
 - financial account** see **balance of payments**
 - Financial Conduct Authority (FCA) 364**
 - financial markets 329, 330**
 - capital markets **329**, 332
 - capital market instruments 332–333
 - participants in 332
 - relationship between market
 - interest rates and bond prices 333–334
 - debt 335, **335**
 - equity 335, **335**
 - financial securities **329**
 - foreign exchange market 334, **334**
 - exchange rate **334**
 - participants in 334, 335
 - money markets **329**, 330
 - instruments traded on 331–332
 - participants in 330, 331
 - role in the economy 335–336
 - Financial Policy Committee (FPC) 363**
 - firms 76**
 - divorce of ownership from control 72–73, **73**, 74, 75
 - profit maximisation 69–72
 - satisficing principle **73**, 75
 - fiscal drag 205**
 - fiscal policy 376, 376**
 - austerity **379**
 - automatic stabilisers **386**
 - budget balance 384, **384**
 - budget deficits 384, **384**

budget surpluses **384**, 385
 consequences of cyclical deficits
 and surpluses 386
 consequences of structural
 deficits and surpluses 387
 cyclical deficits and surpluses
 385
 structural deficits and surpluses
 384, 385
 crowding out **386**, 388
 government expenditure **376**, 377
 consumption expenditure 377
 investment expenditure 377–378
 reasons for government
 expenditure 378–379
 transfer payments 378
 infrastructure **376**
 national debt **386**, 387
 ‘fiscal space’ 389
 implications of a high debt-to-
 GDP ratio 388, 389–390
 measurement of 387–388
 Office for Budget Responsibility 390
 taxation *see* **taxation**
Fisher’s equation of exchange 302,
 302, 303
foreign aid 483, **483**, 484
foreign banks 338
foreign direct investment (FDI) **408**,
 442, **442**, 446, **481**
foreign exchange market 334, **334**,
 335
framing 21, **21**
frictional unemployment **293**, 296,
 396
Funding for Lending Scheme (FLS)
 356–357, 358

G

gender pay gap 179–180
 see also **discrimination**
Gini coefficient 193, **193**, 194
global systemically important bank
 (G-SIB) 368, **368**

globalisation **380**, 402, **402**, 403,
 481
 causes of 403–404
 characteristics of 404–405
 economies of scale **405**,
 406–407
 greater specialisation 405,
 405
 increased international trade
 405, 406
 resource mobility 407–408
 consequences of 408
 less developed economies **409**,
 409–412
 more developed economies
 412–414
 global income distribution, and 414
 gross value added **405**
 multinational corporations **405**,
 414–415, 445
 World Trade Organisation **402**
government expenditure
 education and training 394, 395
 see also **fiscal policy**
government failure 239
government intervention 239
 extension of property rights 239–240
 pollution permits **240**, 240–242
Greece
 barter economy 327
gross domestic product (GDP) 249,
 250
gross national product (GNP)
 249–250, **250**

H

hairdressing 90
Harrod–Domar model **481**
health service workers
 wages 164
hedge funds **330**
higher education 61
Human Development Index (HDI) 257,
 473–475

I

income tax 382–383
industrial policy **394**, 395
industry **76**
inequality *see* **distribution of income**
 and wealth
inflation **302**
 commodities **308**
 world commodity prices and
 domestic inflation 308–310
 consequences of deflation and
 inflation 306–307
 deflation **302**
 disinflation **302**
 expectations 304–305, **305**
 Fisher’s equation of exchange 302,
 302, 303
 hyperinflation **306**, **324**
 monetarism 303, **303**, 304
 quantity theory of money 302, **302**,
 303, 304
 see also **Phillips curve**
information and knowledge
 asymmetric information **11**, 12, 13,
 365
 imperfect information 11, **11**, 162
 importance of information for
 decision making 11, 12, 13
 perfect information 81
 symmetric information **11**, 12
 technological change 60–61
infrastructure **376**, **478**
 lack of 479
innovation 55, **55**, **394**
 see also **technological change**
intellectual property **436**
International Monetary Fund **351**
international trade 417, **417**, 418, **483**
 absolute advantage **418**
 benefits of 418
 achieving macroeconomic
 objectives 421
 consumers, benefits to 422
 theory of comparative advantage
 418, **418**, 419–421

customs unions 429, **429**, 430
 common external tariff 429, **429**,
 430
 dynamic effects 431
 trade creation **429**, 430–431
 trade diversion **429**, 431
 drawbacks of 422
 dumping **426**, 427
 economic growth and economic
 development, impact on
 484–485
 export subsidies **426**, 429
 international merchandise trade **417**
 non-tariff barriers **426**
 opportunity cost **418**
 production possibility diagram **418**
 protectionist measures 426, **426**,
 427
 quotas **426**, 429
 specialisation **418**
 tariffs **426**, 427–429
 UK
 commodity pattern of trade 424,
 425–426
 geographical pattern of trade
 423–424, 425
 see also **European Union;**
 globalisation; World Trade
 Organisation

invention 55, **55**
 see also **technological change**
investment banks **338**, 339
 underwriting **340**
involuntary unemployment 293, **293**,
 294

K

knowledge *see* **information and**
knowledge

L

labour market
 aggregate supply, and 272

 declining role of trade unions
 272–273
 education and enterprise
 273–274
 increasing use of migrant labour
 273
 car industry 146
 demand
 derived demand **142**, 142
 determinants of 144–146, 147
 elasticity of demand 147
 marginal productivity theory **142**,
 142–143, 144
 wage rate and number of workers
 employed 143–144
 determination of wage rates and
 levels of employment
 imperfectly competitive labour
 market 161–165
 perfectly competitive labour
 market 155–157
 health service workers 164
 imperfect labour markets
 bilateral monopoly **162**, 162
 factors contributing to 161
 imperfect information 162
 monopoly power exerted by trade
 unions 162
 monopsony power amongst
 employers 161, **161**,
 162–165
 labour market flexibility **273**, 394
 labour market reforms 393–394
 labour mobility **296**, 297–298, **396**
 role of market forces in determining
 relative wage rates
 perfectly competitive labour
 market 157–158
 supply
 influences on the supply of labour
 to a particular occupation
 149, **149**, 150–151
 net advantage of the job 154
 non-monetary considerations
 153–154

 wage rates 151–152, 153
 training and wage levels 152
 wage differences within occupations
 159
 wage differentials 183, 184
 see also **discrimination;**
 employment and
 unemployment; national
 minimum wage; trade unions
liberalisation of markets *see*
deregulation
liquidity **324**, **343**, **365**
Lorenz curve **192**, 192–193

M

mandated choice **21**, 22
market contestability
 contestable markets **118**, 118, 119
 efficiency, and 119–20
 hit-and-run competition 121, **121**
 non-contestable markets **118**, 118,
 119
 sunk costs **120**, 120
 train operating companies 120
market failure
 externalities 219
 government intervention 239
 marginal social benefits and marginal
 social costs 220
 externalities in consumption
 220–221
 externalities in production
 221–223
 optimum allocation of resources
 219–220
 optimum output of merit and
 demerit goods 223
 pollution permits, use of **240**,
 240–242
 property rights **218**, **239**
 absence of 218
 extending 218–219, 239–240
 social benefits 219
 social costs 219

'tragedy of the commons' 217, 217–218

market structures 66

- distinguishing features of 66–68
- duopoly 66
- efficiency 125–130
- range of 66
- see also **monopolistic competition; monopoly; oligopoly; perfect competition**

Marshall–Lerner condition 449, 450

mergers 340

merit and demerit goods

- optimum output of 223

migration 273, 408, 413–414

minimum wage rate see **national minimum wage**

mobile phones, purchasing 12

monetary policy 351

- effectiveness of 356
- timing of the effects of 355–356
- transmission mechanism 353–354
 - effects on aggregate demand 354–355
 - effects on financial and asset markets 354
 - effects on inflation 355
- see also **Bank of England**

money 324, 324

- characteristics of 325–327
- functions of 325
 - medium of exchange 324, 325
 - standard of deferred payment 326
 - store of value 325–326
 - unit of account 325
- liquidity 324
- supply of money 327, 328
 - broad money 328, 328, 329
 - narrow money 327–328, 328

money illusion 314

money markets 329, 330–332

monopolistic competition 86

- characteristics of 86–87
- efficiency 89, 127–128

- hairdressing 90
- long run 88, 89
- non-price competition 89–90, 91
- short run 87–88, 89

monopoly

- advantages of 108–109
- bilateral monopoly 162, 162, 169
- disadvantages of 109, 110
- efficiency 129–130
- monopoly model 107–108
- monopoly power 107
- pure monopoly 107
- Royal Mail 109, 162

monopsony 161

- determination of wage rates and levels of employment 161, 162–165
- trade unions' influence 169–170

moral hazard 12–13, 369, 369–370

mortgages 342

multinational corporations 405, 414–415, 445

N

national debt see **fiscal policy**

national income 248, 248

- gross domestic product 249, 250
- gross national product 249–250, 250
- measurement of
 - measures of the size of an economy 249–250
 - methods of 249
 - reasons for measuring 250, 251
- net property income from abroad 249–250, 250
- standard of living 250, 251–256

national living wage (NLW) 173, 175–176

National Lottery 131

national minimum wage (NMW) 173, 173

- advantages of 174
- disadvantages of 174–175

- impact on labour markets 173–174, 175
- median wage, comparison with 175
- national living wage, and 173, 175–176
- wage 'stickiness' 296

'nationalisation see **public ownership**

natural rate of unemployment 296, 296–298, 314, 396, 396, 397–398

Netflix 114, 117, 382

noise pollution 223

non-accelerating inflation rate of unemployment (NAIRU) 298, 300, 314, 316, 318

nudges 21, 21, 22

O

OECD

- 'better life index' 257
- distribution of income and wealth in OECD countries 196–197

Office for Budget Responsibility (OBR) 390

oligopoly 92

- advantages of 103
- advertising 103
- barriers to entry 101
- cartels 96, 97, 97, 100, 101
- characteristics of 92–93
 - variation in 94
- collusive oligopoly 96–97, 97, 98
- communications market 104
- concentration ratios 95–96
- cooperation 98
- disadvantages of 104, 105
- efficiency 128–129
- examples of 94
- interdependence and uncertainty 103
- investment, factors influencing 102
- kinked demand curve 98–100
- market conduct 94, 95
- market structure 94
- non-collusive oligopoly 97, 98

non-price competition **100**
output, factors influencing 102
price agreements **100**, 101
price leadership 98, **100**, 101
price wars **100**, 101
prices, factors influencing 101–102
research and development 102–103
Royal Mail 109

Olympic Stadium 228

outsourcing 214, 215, 216

P

perfect competition 76

Billingsgate Market 83–84
efficiency 82–83, 125–127
freedom of entry and exit 79–81
identical products 79
large numbers of producers 78–79
long run 77–78
perfect information and knowledge 81
price takers 81, **81**
short run 76–77
sunk costs **80**, 81

Phillips curve **311**

development of 312–313
long-run Phillips curve 315–316
implications of 316–317
non-accelerating inflation rate of unemployment **314**, 316, 318
origins of 311–312
short-run Phillips curve 314
implications of 316
supply-side policies 317, **317**
value of, for shaping economic policy 317–319

pollution permits **240**, 240–242

poverty **199**

absolute poverty 199, **199**
case studies
UK 201–202
causes of 200–201, 202–203
effects of 203

‘income after housing costs’ 199, 200

‘income before housing costs’ 199

measures of poverty 199, 200

relative poverty 199, **199**

see also **economic growth and economic development; redistribution of income and wealth**

price agreements **100**, 101

price discrimination 111, **111**

advantages of 113
‘big data’, use of 114
conditions necessary for 112–113
consumer and producer surplus 136–137
disadvantages of 113
distinguishing factors 111
energy market 48–49

price leadership 98, **100**, 101

price mechanism

advantages of 212–213, 214
case studies
donating blood 215
disadvantages of 213, 214
use of 214
charities and donations 214
outsourcing 214, 215, 216
private sector provision of public sector goods and services 215–216

price takers 81, **81**

price wars **100**, 101

primary sector of an economy **469**

privatisation **232**, 232–233, **392**, 393

producer surplus **134**, 134, 135

see also **surplus**

production

average returns 29, **29**, 30
constant returns to scale **31**, 32
decreasing returns to scale **31**, 32
increasing returns to scale **31**, 32
law of diminishing returns 30–31
implications for costs of production 33

long run **28**, 28

marginal returns 29, **29**, 30

retailing 33

returns to scale 31, **31**, 32

implications for costs of production 33

short run **28**, 28, 30

total returns 29, **29**, 30

see also **costs of production; technological change**

productive efficiency 57, 82, 83, **122**, 123–125

see also **efficiency**

productivity **270**, **394**

costs of production, and 38, 39

profit **50**

abnormal profit **50**, 51
calculation of 50
normal profit 50, **50**
output levels 51, 52
profit maximisation **50**, 69, 70
marginal revenue and marginal cost **70**, 70–72
role of profit in a market economy 52–53, 54
supermarkets 53

progressive taxes **378**, 378

property rights **218**, 218–219, **239**, 239–240

Prudential Regulation Authority (PRA) 362, **362**

public finances

economic growth 284

public ownership **230**, 230–231

purchasing power parity (PPP)

exchange rates 255–256, **469**

Q

quantitative easing **358**, 358, 359

R

rational economic decision making

6, 7

- real economy** 370
- real wage unemployment** 294, 294, 295–296
- redistribution of income and wealth**
- Budgets 207
 - 'earnings trap' 205
 - fiscal drag 205
 - government policies 204–205
 - economic consequences of 206
 - impact on individual markets 207–208
 - see also* **fiscal policy**
- regional policy** 394
- regulation** 233, 233
- arguments against 234
 - arguments in favour of 233–234
 - banking industry 236–237
 - deregulation 234, 234–235, 392, 393
 - regulatory bodies 237
 - regulatory capture 229, 237, 237
- regulation of the financial system**
- macroprudential regulation 362
 - regulatory bodies
 - Bank of England 361–362
 - Financial Conduct Authority 364
 - Financial Policy Committee 363
 - Prudential Regulation Authority 362, 362
 - review of 361
 - stress testing the banking system 369, 371
 - systemic risk 362, 370, 371
- research and development** 394
- oligopolistic industries 102–103
 - subsidising expenditure on 395–396
- restricted choice** 21, 21
- retailing** 33, 53, 265–266, 402
- returns to scale** 31, 31, 32, 33
- revenue** 45
- average revenue 45, 45
 - energy market 48–49
 - imperfect market 46, 47, 48
 - marginal revenue 45, 45, 46
 - perfectly competitive market 46, 47
 - total revenue 45, 45
- Royal Bank of Scotland (RBS)** 367
- Royal Mail** 109, 161, 162
- rules of thumb** 16, 16
- ## S
- satisficing principle** 73, 75
- seasonal unemployment** 293
- secured loans** 347
- securitisation** 365
- shadow economy** 251
- shares** 332, 333
- skills gap** 397
- social norms** 17, 17
- sovereign wealth funds** 408, 408
- speculation** 287, 291
- standard of living** 250
- Big Mac Index 257
 - economic growth 283–284
 - Human Development Index 257
 - limitations of using national income data to measure 251, 252–253
 - OECD 'better life index' 257
 - purchasing power parity exchange rates 255–256, 469
 - using national income data to compare international standards 254–255
 - see also* **economic growth and economic development**
- static efficiency** 122, 122, 123
- see also* **efficiency**
- statistical analysis**
- groupings 190
- structural unemployment** 293, 296
- sunk costs** 80, 81, 120, 120
- supply-side policies** 392, 392
- categories of 393
 - free market supply-side policies 393
 - deregulation 392, 393
 - labour market flexibility 394
 - labour market reforms 393–394
 - privatisation 392, 393
 - tax cuts 393
 - interventionist supply-side policies 394
 - government spending on education and training 394, 395
 - industrial policy 394, 395
 - research and development, subsidising expenditure on 394, 395–396
 - microeconomic effects 398
 - natural rate of unemployment, and 396, 396, 397–398
 - regional policy 394
 - skills gap 397
- surplus**
- auctions 137
 - consumer surplus 134, 134, 135
 - price discrimination 136–137
 - producer surplus 134, 134, 135
 - welfare benefit 135–136
- symmetric information** 11, 12
- systemic risk** 339–340, 340, 362, 370, 371
- ## T
- takeovers** 340
- taxation** 379, 379
- corporation tax 382, 383–384
 - direct taxes 379
 - hypothecation 380
 - income tax 382–383
 - indirect taxes 379
 - ineffective systems of taxation 478
 - principles of 380, 380, 381
 - progressive taxes 378, 378
 - purpose of 379, 380
 - revenue from individual taxes 381
 - role and merits of the major taxes 382
 - tax cuts 393
 - value added tax 383
- taxi drivers** 18
- technical efficiency** 122, 123
- see also* **efficiency**

technological change 55, 56

- Apple iPhone 63–64
- competition and cooperation 62
- costs of production
 - impact of high-cost technology 57–58
 - impact of low-cost technology 58
- creative destruction 59, 59–60
 - higher education 61
- development of new markets 59
- development of new products 59
- efficiency
 - allocative efficiency 57
 - dynamic efficiency 58–59
 - productive efficiency 57
- innovation 55, 55
- invention 55, 55
- methods of production 56
- productivity 56–57
- structure of markets
 - barriers to entry 61
 - homogeneity and differentiation of goods and services 60
 - information and knowledge 60–61
 - number of firms 60

trade unions 162, 273, 294

- bilateral monopoly 169
- declining role of 272–273

- functions of 166
- influence on wages and levels of employment 167
- factors affecting 167–168
- monopsony labour market 169–170
- previously perfectly competitive labour market 168–169
- trade union wage gap 167, 171–172, 184, 294, 296
- involvement in labour markets
 - advantages and disadvantages of 170–171
 - membership 171
 - monopoly power exerted by 162
 - reducing the labour supply 298
- ‘tragedy of the commons’** 217, 217–218

train operating companies 120

transition economy 408

Treasury 330

Treasury bills 330, 331

U

unemployment *see* **employment and unemployment**

utility theory 7, 7

- equi-marginal utility 9
- law of diminishing marginal utility 8

- marginal utility 7, 8, 9
- total utility 7, 8
- utility maximisation 8, 9, 10

V

value added tax (VAT) 383

voluntary unemployment 294, 294

W

wage discrimination *see*

discrimination

wage rates

- UK real wages 276
- wage ‘stickiness’ 295–296
- see also* **labour market; national minimum wage**

wealth 186

- see also* **distribution of income and wealth**

working time 253

World Trade Organisation (WTO) 402,

- 435, 435–436
- dispute settlement 436
- trade agreements 436

Y

yield 334