STUDENT GUIDE

NEW EDITION

AQA A-LEVEL

Economics

Individuals, firms, markets and market failure

Ray Powell
James Powell
George Vlachonikolis

HODDER EDUCATION
LEARN MORE

Magazines

that give you the edge

From A-level biology to GCSE geography, Hodder Education Review magazines go beyond your textbooks to help you get your best grade.

Stretch your subject knowledge

with topical articles written by subject experts.

Stay up-to-date

with the latest case studies, discoveries and research to use in your exams.

Prepare for exam success

with grade-boosting advice, practice exam-style questions and analysis from senior examiners.

Develop your study skills

and get ready for higher education.

Also available as eMagazines via your school

Want to see a copy?

Just ask your teacher or librarian for more details.

www.hoddereducation.co.uk/magazines

HODDER EDUCATION
LEARN MORE

A-LEVEL
STUDENT GUIDE 1

AQA

Economics

Individuals, firms, markets and market failure

Ray Powell
James Powell
George Vlachonikolis

HODDER EDUCATION
AN HACHETTE UK COMPANY

This Guide has been written specifically to support students preparing for the AQA A-level Economics examinations. The content has been neither approved nor endorsed by AQA and remains the sole responsibility of the authors.

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

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

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

Orders: please contact Bookpoint Ltd, 130 Park Drive, Milton Park, Abingdon, Oxon OX14 4SE. Telephone: (44) 01235 827827. Fax: (44) 01235 400401. Email: education@bookpoint.co.uk. Lines are open from 9 a.m. to 5 p.m., Monday to Saturday, with a 24-hour message answering service. You can also order through our website: www.hoddereducation.co.uk.

ISBN 978-1-5104-7200-6

First printed 2020

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

www.hoddereducation.co.uk

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

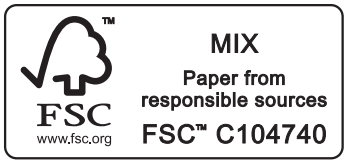
Year 2024 2023 2022 2021 2020

Cover photo: Vojtech Vlk/Adobe Stock

Typeset by Aptara, India

Printed in Italy

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



Contents

Content Guidance

Economic methodology and the economic problem

Individual economic decision making

Price determination in a competitive market

Production, costs and revenue

Perfect competition, monopoly and imperfectly competitive markets

The labour market

The distribution of income and wealth: poverty and inequality

The market mechanism, market failure and government intervention in markets

Questions & Answers

Getting the most from this book

Exam tips

Advice on key points in the text to help you learn and recall content, avoid pitfalls, and polish your exam technique in order to boost your grade.

Knowledge check

Rapid-fire questions throughout the Content Guidance section to check your understanding.

Knowledge check answers

1 Turn to the back of the book for the Knowledge check answers.

Summaries

- Each core topic is rounded off by a bullet-list summary for quick-check reference of what you need to know.

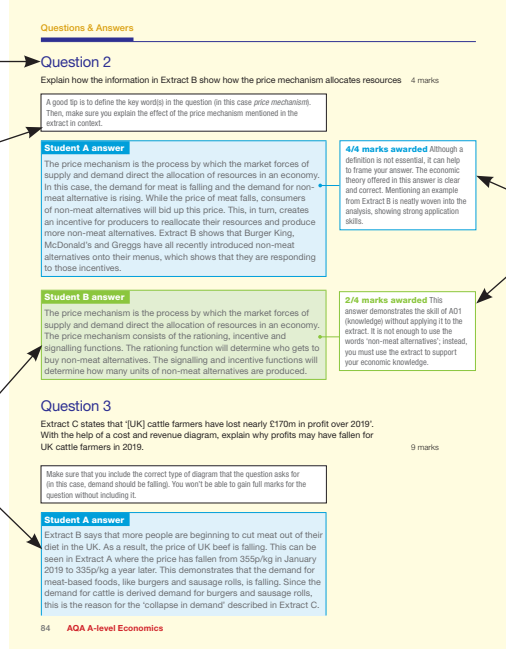
Exam-style questions

Commentary on the questions

Tips on what you need to do to gain full marks.

Sample student answers

Practise the questions, then look at the student answers that follow.



Questions & Answers

Question 2

Explain how the information in Extract B show how the price mechanism allocates resources 4 marks

A good tip is to define the key word(s) in the question (in this case *price mechanism*). Then, make sure you explain the effect of the price mechanism mentioned in the extract in context.

Student A answer

The price mechanism is the process by which the market forces of supply and demand direct the allocation of resources in an economy. In this case, the demand for meat is falling and the demand for non-meat alternative is rising. While the price of meat falls, consumers of non-meat alternatives will bid up this price. This, in turn, creates an incentive for producers to reallocate their resources and produce more non-meat alternatives. Extract B shows that Burger King, McDonald's and Greggs have all recently introduced non-meat alternatives onto their menus, which shows that they are responding to those incentives.

4/4 marks awarded Although a definition is not essential, it can help to frame your answer. The economic theory offered in this answer is clear and correct. Mentioning an example from Extract B is neatly woven into the analysis, showing strong application skills.

Student B answer

The price mechanism is the process by which the market forces of supply and demand direct the allocation of resources in an economy. The price mechanism consists of the rationing, incentive and signalling functions. The rationing function will determine who gets to buy non-meat alternatives. The signalling and incentive functions will determine how many units of non-meat alternatives are produced.

2/4 marks awarded This answer demonstrates the skill of AO1 (knowledge) without applying it to the extract. It is not enough to use the words 'non-meat alternatives'; instead, you must use the extract to support your economic knowledge.

Question 3

Extract C states that '[UK] cattle farmers have lost nearly £170m in profit over 2019'. With the help of a cost and revenue diagram, explain why profits may have fallen for UK cattle farmers in 2019. 9 marks

Make sure that you include the correct type of diagram that the question asks for (in this case, demand should be falling). You won't be able to gain full marks for the question without including it.

Student A answer

Extract B says that more people are beginning to cut meat out of their diet in the UK. As a result, the price of UK beef is falling. This can be seen in Extract A where the price has fallen from 355p/kg in January 2019 to 335p/kg a year later. This demonstrates that the demand for meat-based foods, like burgers and sausage rolls, is falling. Since the demand for cattle is derived demand for burgers and sausage rolls, this is the reason for the 'collapse in demand' described in Extract C.

84 AQA A-level Economics

Commentary on sample student answers

Read the comments showing how many marks each answer would be awarded in the exam and exactly where marks are gained or lost.

About this book

The aim of this guide is to prepare students for the AQA A-level Economics Paper 1 'Individuals, firms, markets and market failure' examination and for the microeconomic parts of AQA A-level Paper 3. All the topics explained in this book could be examined in the A-level Paper 3, which is a synoptic paper testing the whole of the A-level specification.

How to use this book

The **Content Guidance** section of the book covers eight microeconomic topics in the order in which they appear in the AQA A-level 'Individuals, firms, markets and market failure' specification, starting from 'Economic methodology and the economic problem' and finishing with 'The market mechanism, market failure and government intervention in markets'.

You should read the **Questions & Answers** section of the guide either after reading all eight specification topics in the Content Guidance section, or bit by bit, having revised a selected topic on a particular part of the specification. This second section of the guide includes examples of all the forms of assessment in the A-level economics examination. These are multiple-choice questions (MCQs), data-response questions (DRQs), essay questions (EQs) and finally an extended investigation/case study question (IQ).

This guide should be used as a supplement to other resources, such as class notes, the *AQA A-level Economics for A-level* textbook, *Economic Review* magazine and *A-level Economics My Revision Notes* (all published by Hodder Education). As this guide contains summaries rather than in-depth coverage of all the topics in the specification, you should not use the guide as your sole learning resource during the main part of the course. However, you may well decide to use the guide as the key resource in your revision programme. You are strongly advised to make full use of the Questions & Answers section, especially in the revision period when you should be concentrating on improving your examination skills.

Content Guidance

Economic methodology and the economic problem

Economic methodology

Economics as a social science

Like other social scientists in subjects such as psychology, economists start off by observing some aspect of human behaviour and then try to develop a theory from what they have observed. In the case of production theory, the starting point is observations of how firms react to changes in the prices of the goods and services they sell. Production theory then develops from establishing a tentative description, known as a **hypothesis**, of what has been observed. Predictions about human behaviour are deduced from the hypothesis, such as that the owner of a firm will always respond to the price of a good rising by supplying more of the good in question. This prediction is then tested against collected evidence about how firms behave in the market place. At this stage, the hypothesis becomes a **theory**. (A hypothesis is a proposed explanation for something, whereas a theory is when a hypothesis is tested and survives the test.) Nevertheless, a theory may not be true in all circumstances. All it says is that the hypothesis has survived the test or tests to which it has been exposed; it might not survive stronger tests, which may not yet have been devised. Scientific method is based on the possibility of falsification or refutation of a hypothesis.

Hypothesis A proposed explanation for something.

Theory When a hypothesis is tested and survives the test it becomes a theory.

Difference between social and natural sciences

Natural science theories such as those embodied in physical sciences, e.g. astronomy are usually much 'harder' than the theories associated with 'softer' social sciences such as economics. Economic theories often survive only through allowing a significant number of exceptions to their central predictions, which, according to critics, turns the theories into little more than generalisations.

Positive and normative statements

Economists often respond to the criticism that their subject is 'soft' by arguing that they are only concerned with '**positive economics**'. A positive statement can be scientifically tested to see whether it is incorrect. If a positive statement does not pass the test, it is falsified. However, a positive statement does not have to be true. For example, 'the earth is flat' is a positive statement. A few people may believe it, though obviously with the growth of scientific evidence the statement has been falsified.

In contrast, '**normative economics**', which is concerned with 'what should or ought to be', is about value judgements and opinions, and because people have

Positive economics is about statements that can be scientifically tested to see whether they can be falsified.

Normative economics is about value judgements and opinions that cannot be scientifically tested.

different opinions about what is right and wrong, normative statements cannot be scientifically tested: they are just opinions. Words such as 'ought', 'should', 'better', 'worse', 'good' and 'bad' (used as adjectives) often provide clues that a statement is normative.

How value judgements influence economic policy and decision making

A **value judgement** is about whether something is desirable or not — if we believe it is more desirable to study what *is* happening in the economy rather than what *ought* to happen, we have made a value judgement. Economics necessarily requires that government ministers make value-based judgements when deciding on economic policies. Despite this, economists often wrongly insist that the subject is value free.

Value judgement An opinion which cannot be proved.

Government ministers make decisions on issues such as where a new airport should be located or whether high-speed trains are worthwhile. Before making decisions on issues such as these, government ministers usually create the illusion that the decision-making process is completely scientific and objective. To do this, they hire independent 'experts' to provide advice. But the choice of expert in itself involves a value judgement. Whichever way you go, the so-called scientific processes used by the 'experts' to reach their conclusions may be riddled with value judgements. A classic case involved weighing up the costs and benefits of the location of a third London airport, which ultimately depended on putting money values on an hour of a business person's time and an hour of a holidaymaker's time. It was quickly found that when different values were put on these, the airport location recommended by the experts would have 'lost out' under different costing criteria.

The impact of moral and political judgements

Whatever decision is eventually made in the course of framing government economic policy, there will always be winners and losers who gain or suffer as a result of the decision. Governments often claim they have a moral right to make such decisions. They argue that their political manifesto published *before* the previous general election gives them the mandate, supported by the voters, to carry out their policies, regardless of the fact that among the electorate there will inevitably be some losers.

The nature and purpose of economic activity

The central purpose of economic activity

The central purpose of economic activity is the production of goods and services to satisfy needs and wants, with the ultimate objective of increasing people's happiness or **economic welfare**.

Economic welfare The economic wellbeing of an individual, a group within society or an economy.

Increased **production** enables economic welfare to improve, but only if the production of more goods and services leads to higher levels of **consumption**. Production and consumption often result in resource depletion (using up scarce

Production A process or set of processes that converts inputs into outputs.

Consumption The use of goods and services by households.

resources) and resource degradation (e.g. pollution and destruction of the natural environment).

As a general rule, consumption improves economic welfare and people's standard of living. Economists often use the word '**utility**' for the welfare that people enjoy when they consume goods and services. Goods, such as food bought for consumption, are known as '**consumer goods**'; by contrast, a good such as a machine bought by a firm in order to produce other goods is called a '**capital good**'. Goods that people produce for their own consumption, and activities such as contemplating the natural environment, contribute to people's utility or welfare, adding to the utility obtained from consuming goods bought in the market.

Utility Welfare people enjoy when consuming goods and services.

Consumer goods Goods, such as food, which yield utility consumption.

Capital good A good, such as a machine, used to produce other goods.

When discussing economic welfare, we need to distinguish between a **need** and a **want**. A need refers to something people have to have, something they cannot do without. Food is an example. If people starve, they will eventually die. By contrast, a want refers to something people would like to have but which is not essential for survival. It is not absolutely necessary, but it is a good thing to have.

Need A need is something that is necessary for human survival, such as food.

Want A want is something that is desirable, such as ice cream, but is not necessary for human survival.

There are also important elements of human happiness and welfare that have nothing to do with the consumption of material goods. These include quality of life factors, such as the pleasure gained from family and friends or from contemplating a beautiful view.

Key economic decisions

In large part, economics is the study of economising — the study of how people make choices about what to produce, how to produce and for whom to produce, in a world in which most resources are limited or scarce. How best can people make decisions on how scarce resources should be allocated among competing uses so as to improve and maximise human happiness and welfare? This is the economic problem, which is the main focus of this introductory topic.

Economic resources

For most people, most of the time, increased consumption of material goods is an important part of improving economic welfare. Most, if not all, of the goods we consume must first be produced. This requires the use of **economic resources**. These goods are scarce in relation to demand, which gives rise to the need for economising in their use.

Economic resources The inputs used to produce goods and services, often called the factors of production.

The inputs into the production process are often called the **factors of production**. Four factors of production are usually identified. These are land, labour, capital and enterprise, the last often being called the 'entrepreneurial input'.

Factors of production Land, labour, capital and enterprise (or the entrepreneurial input).

Entrepreneurs are different from the other factors of production. They are the people who address issues such as what to produce, how to produce it and for whom to produce it. An entrepreneur decides how much of the other factors of production, including labour, to employ. When making these decisions, the entrepreneur takes into account the financial risks involved. **Profit**, which is the entrepreneur's financial reward, results from successful decision making. Entrepreneurial profit is the profit left over after the cost of employing the other factors of production is deducted from

Profit Sales revenue minus costs of production.

the sales revenue gained from the sale of the goods and services the entrepreneur decides to produce.

The environment as a scarce resource

Environmental resources are part of the factor of production, land. Some environmental resources, such as the air we breathe and the water we drink, are often described as the 'free gifts of nature'. However, this view can be questioned. The need to get rid of the effects of pollution created by humankind means that clean air and water are scarce and not free. Production and consumption activities taking place in the economy affect and often damage the natural environment.

Environmental resources include:

- physical resources, such as soil, water, forests, fisheries and minerals
- gases, such as hydrogen and oxygen
- abstract resources, such as solar energy, wind energy, the beauty of the landscape, good air and clear water

Environmental resources can be further split into renewable and non-renewable resources, with the latter further divided into recyclable and non-recyclable resources.

Scarcity, choice and the allocation of resources

Scarcity as the fundamental economic problem

Economics is literally the study of economising — the study of how human beings make choices about what to produce, how to produce and for whom to produce, in a world in which most of the resources are limited. Because resources are limited in relation to people's infinite wants, scarcity is the fundamental **economic problem**.

Scarcity means that choices have to be made

The fundamental economic problem exists because both goods and the resources needed to produce goods are scarce. Scarcity also means that people (even the very rich) have limited incomes and face a budget constraint. If goods are scarce and incomes are limited, choices have to be made. And even when goods are free, time is scarce, so choices still have to be made. A need for **choice** also arises whenever an economic agent (for example, an individual, a household or a firm) has to choose between two or more alternatives that are mutually exclusive, in the sense that it is impossible or impractical to achieve both at the same time.

Choices have an opportunity cost

Whenever an individual, a household, a firm or the government has to choose between two or more alternatives which are mutually exclusive, in the sense that it is impossible or impractical to achieve both at the same time, an **opportunity cost** is involved. The opportunity cost of any choice, decision or course of action is measured in terms of the alternatives that have to be given up.

Economic problem There is only a limited amount of resources available to produce the unlimited quantity of goods and services people desire.

Choice The result of the fundamental economic problem of scarcity, which means that decisions have to made by firms, individuals and/or governments regarding which needs and wants to satisfy, and what types of products and services should be produced.

Knowledge check 1

What is meant by scarcity and economising?

Opportunity cost The cost of giving up the next best alternative.

Production possibility diagrams

Production possibility diagrams illustrate different features of the fundamental economic problem, such as: resource allocation, opportunity cost and trade-offs, unemployment of economic resources, economic growth.

A **production possibility diagram** illustrates the different combinations of two goods, or two sets of goods, that can be produced with a fixed quantity of resource, providing we assume that all available resources are being utilised to the full. The production possibility diagram in Figure 1 illustrates the different combinations of capital goods and consumer goods that the whole economy can produce when all the economy's resources are employed, with no spare capacity. To put it another way, the curve in the diagram shows what the economy can produce, assuming that all the labour, capital and land at the country's disposal are employed to the full, and assuming a given state of technical progress.

Suppose that initially the economy is at point *A* on the curve, producing K_1 capital goods and C_1 consumer goods. In the absence of economic growth (which moves the curve outwards), consumer good production can only increase to C_2 if the production of capital goods falls from to K_1 to K_2.

The fall in the production of capital goods when the production of consumer goods increases is an example of opportunity cost.

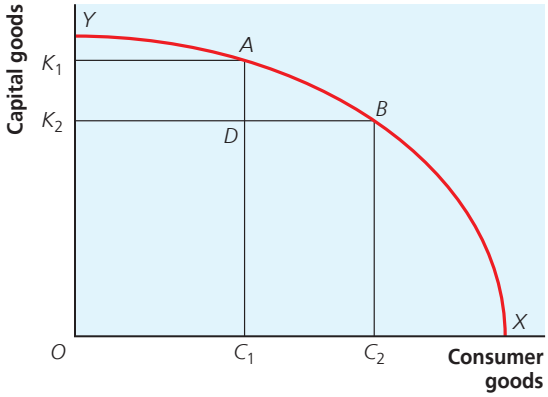


Figure 1 An economy's production possibility diagram

All points on an economy's production possibility diagram are **productively efficient** because they show the maximum possible levels of output that available inputs can produce. However, they are unlikely to be **allocatively efficient** because they say nothing about the economic welfare derived from consumption.

If the economy's production possibility curve is the line *YX* in Figure 1, short-run economic growth is illustrated by a movement from a point inside the curve such as *D* to a point on the curve. Long-run economic growth would be illustrated by an outward shift of the curve to a 'further out' position. The opportunity cost of increasing consumer good production from C_1 to C_2 is capital good production falling from K_1 to K_2. All points on the curve represent full employment of all available resources. By contrast, unemployed resources exist at all points inside the curve, such as *D*.

Production possibility diagram A production possibility diagram represents graphically alternative production possibilities open to an economy. Since resources are scarce, a choice has to be made between the alternative goods that can be produced.

Exam tip

You must learn to draw and interpret production possibility diagrams, which are as important in macroeconomics as in microeconomics.

Productive efficiency Maximising output from available inputs.

Allocative efficiency When it is impossible to improve overall economic welfare by reallocating resources between markets. In the whole economy, price must equal marginal cost ($P = MC$) in every market.

Knowledge check 2

Give an example of an opportunity cost, other than the capital goods given up when the production of consumer goods increases.

Summary

- Economics is the study of economising.
- Economics provides answers to questions on what to produce, how to produce and for whom to produce, in a world in which most of the resources are limited relative to wants and needs.
- Scarcity is the fundamental economic problem.
- The production of economic goods uses up scarce resources and people have to economise in their use.
- Scarcity of economic resources means opportunity costs exist and choices have to be made.
- A production possibility diagram shows the different combinations of goods that can be produced from available resources.
- A normative statement involves a value judgement, whereas a positive statement can be tested to see whether it is true or false.

Individual economic decision making

Consumer behaviour

Rational economic decision making and economic incentives

Economists generally assume that people behave rationally. **Rational behaviour** means people try to make decisions in their self-interest or to maximise their private benefit. When a choice has to be made, people always choose what they think at the time is the best alternative, which means that the second best or next best alternative is rejected. Providing people are rational, the opportunity cost of any decision or choice is the next best alternative sacrificed or foregone. For example, if you choose to spend half an hour reading an economics textbook, the opportunity cost is the lost opportunity to spend that time doing other things, for example watching a television programme.

Rational behaviour Acting in self-interest.

If individuals realise that wrong choices have been made, which fail to maximise personal economic welfare, they have incentives to change their economic behaviour, until welfare maximisation is achieved.

Utility theory: total and marginal utility, and the hypothesis of diminishing marginal utility

In economics, utility, defined as the pleasure or satisfaction obtained from consumption, is usually divided into total utility and **marginal utility**. To explain the difference between total and marginal utility, let us imagine a thirsty child who drinks six glasses of lemonade on a hot sunny afternoon, deriving successively 8, 6, 4, 2, 0 and –2 'units of utility' from each glass consumed. This information is shown in Table 1 and plotted in Figure 2. Note that marginal utility is plotted at 'halfway' points.

Marginal utility The additional satisfaction, welfare or pleasure gained from consuming one extra unit of a good or service.

Table 1 Total and marginal utility schedules for lemonade

Glasses of lemonade	Total utility (units of utility)	Marginal utility (units of utility)
0	0	
		8
1	8	
		6
2	14	
		4
3	18	
		2
4	20	
		0
5	20	
		–2
6	18	

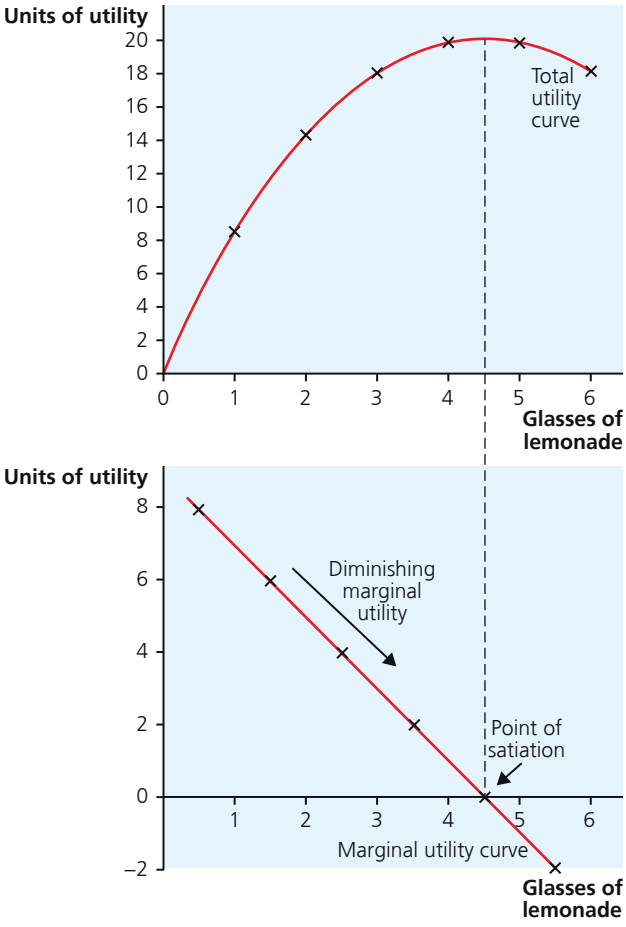


Figure 2 An example of total utility and marginal utility curves

The hypothesis or 'law' of diminishing marginal utility

The numerical examples in Table 1, and the graph in Figure 2, illustrate a famous economic hypothesis, which some would call an economic law: the **hypothesis of diminishing marginal utility**. This simply states that as a person increases consumption of a good — while keeping consumption of other products constant — there is a decline in the marginal utility derived from consuming each additional unit of the good.

The assumption of maximising behaviour by economic agents (consumers, workers, firms and the government) is central to orthodox or traditional economic theory. Economic agents decide their market plans so as to maximise a target objective or goal which is believed to be consistent with the pursuit of self-interest. In demand theory, the objective which households are assumed to wish to maximise is the utility, or satisfaction, obtained from the set of goods and services consumed.

Because of the problem of scarcity, consumers face a number of constraints which restrict the choices they make in the market place. The main constraints are: limited time available; limited income; the given set of prices they face; budget constraint. Taken together, limited income and the set of prices faced impose a budget constraint on consumers' freedom of action in the market place. As a general rule, a consumer

Exam tip

The relationships between marginal values and total values of an economic variable must be understood when studying production theory, cost theory and revenue theory, as well as when studying utility theory. With production theory, cost theory and revenue theory, you must also understand the relationships between marginal and average returns, marginal and average cost and marginal and average revenue.

Hypothesis of diminishing marginal utility For a single consumer, the marginal utility derived from a good or service diminishes for each additional unit consumed.

can only purchase more of one good by giving up consumption of some other good or service, which represents the opportunity cost of consumption.

The importance of the margin when making choices

Along with assumptions such as rational economic behaviour and opportunity cost, the 'margin' is one of the key concepts in traditional or orthodox economic theory. Given consistent tastes and preferences, rational consumers choose between available goods and services in such a way as to try to maximise total utility, welfare or satisfaction derived from consumption of the goods. Along with the relative prices that must be paid for each of the goods, the marginal utilities gained from the consumption of the last unit of each good determine the combination of goods the consumer must choose in order to maximise total utility.

In order to maximise a desired objective, an economic agent must undertake the activity involved up to the point at which the marginal private benefit received equals the marginal private cost incurred. For example, a utility-maximising consumer must choose to consume or demand a good up to the point at which $MU = P$. Marginal utility or MU is the marginal private benefit derived from consuming the last unit of the good, while the good's price, P, is its opportunity cost in consumption, at the margin.

Exam tip

The margin is one of the key concepts in A-level microeconomics. Make sure you understand and can apply the concept.

Knowledge check 3

What is the difference between total utility and marginal utility?

Imperfect information

The importance of information for decision making

When attempting to maximise total utility, more often than not consumers possess imperfect information. As a result, they make 'wrong' decisions. We shall explain on page 76 how consumers may choose to under-consume a merit good such as education and over-consume a demerit good such as tobacco because they possess imperfect information about the long-term consequences of their choices.

The significance of asymmetric information

Asymmetric information arises when either the buyer or the seller involved in a potential transaction knows something that is not observable to the other party. One of the ways in which asymmetric information can manifest itself is through the process known as 'adverse selection', which is a feature of many market transactions. For example, in the sale and purchase of a second-hand computer, the seller of the good knows more about the computer's defects than a potential purchaser. However, to avoid paying too high a price for an inferior product that has lots of defects, potential purchasers often offer low prices on all second-hand computers, regardless of the fact that some of the computers are in good working order.

Asymmetric information When one party to a market transaction possesses less information relevant to the exchange than the other.

Knowledge check 4

Explain the difference between symmetric and asymmetric information.

Aspects of behavioural economic theory

Behavioural economics has emerged in recent decades because of the dissatisfaction felt by some members of the economics profession with what they call traditional or orthodox economic theory. Traditional economic theories have been attacked by behavioural economists on the grounds that the simplifying assumptions on which the theories are built are unrealistic. Writing in November 2016 in *Economic Review* (also published by Hodder), Maria Kozlovskaya discussed the issue of whether people are rational. She argued that there are many situations in which consumers appear not to act rationally and that people sometimes take decisions that involve picking an option that makes them worse off. For example, people may pay for gym membership which they then do not use fully. Another phenomenon involves **loss aversion**. Experiments have shown that people's perception of probability is distorted, in that people may over-estimate the likelihood of extreme results. People's choices depend on the way in which options are framed to them and they often stick with the default option.

In recent years, behavioural economics has emerged to question many of the assumptions of traditional economic theory. Key concepts in behavioural economics include **bounded rationality**, bounded self-control, biases in decision-making processes and anchoring. Bounded rationality means that individuals, however high or low their intelligence, make decisions subject to three unavoidable constraints: imperfect information about possible alternatives and their consequences; limited mental processing ability; and a time constraint which limits the time available for making decisions. In complex choice situations, bounded rationality often results in satisficing rather than maximising choices.

Bounded rationality is closely linked to the related concept of bounded self-control. Traditional or orthodox economic theory implicitly assumes that when making choices, individuals have complete self-control. Behavioural economists, by contrast, believe that individuals have bounded (or limited) self-control. A good example is provided by New Year resolutions. Having put on weight during the Christmas festivities, people may decide to go for a daily jog early in the morning before going to work each day after 1 January. For many, this may work well for a few days, but the first bout of bad weather often leads to the resolution being broken.

Behavioural economics argues that the quick decisions (automatic thinking) people make when exercising choice are often heavily biased. This is because decisions are made on the basis of one's own likes, dislikes and past experiences. Psychologists use the term **cognitive bias** to describe this situation. A cognitive bias is a mistake in reasoning or other thought process, often occurring as a result of holding onto one's preferences and beliefs regardless of contrary information.

Many kinds of cognitive biases exist, such as confirmation bias which is the tendency to seek only information that matches what one already believes. The **availability bias** relates to people making judgements about the probability of events by how easy it is to recall examples of such events. **Anchoring**, which is an example of a predictable bias in individual decision making, is the tendency to rely too much on a single piece of information, frequently the first piece of information, when making decisions.

Behavioural economics A method of economic analysis that applies psychological insights into human behaviour to explain how individuals make choices and decisions.

Loss aversion The idea that losses generally have a much larger psychological impact than gains of the same size.

Bounded rationality When making decisions, an individual's rationality is limited by the information they have, the limitations of their mind, and the finite amount of time available in which to make decisions.

Knowledge check 5

Give an example of irrational behaviour.

Cognitive bias A mistake in reasoning occurring as a result of, for example, using rules of thumb or holding on to one's preferences and beliefs, regardless of contrary information.

Availability bias People making judgements about the probability of events by how easy it is to recall examples of such events.

Anchoring The tendency to rely too much on a single piece of information, frequently the first piece of information, when making decisions.

Behavioural economics and economic policy

In the context of the impact of behavioural economics on government economic policy making, you need to consider how behavioural economics might influence the design of a variety of government policies which aim to reduce or eliminate particular economic problems.

Behavioural economists do not deny that rational considerations play a part in people's decision making but they believe that traditional economic models provide an incomplete account of the factors that influence the choices people make. Behavioural economic analysis extends the traditional model in an attempt to provide a better, more realistic explanation of human decision making.

The Behavioural Insights Team (BIT), set up by the coalition government in 2010, suggested different ways in which behavioural insights can be applied to economic policy. These are to make it easy, attractive, social and timely for members of the general public to comply with the policy. In particular, the BIT focused on **default choices**, which in turn relate to **choice architecture**. Choice architecture is the format in which choices are presented to the general public, for example to try to get individuals to make the choice which the policy maker is seeking.

Governments can use choice architecture in an attempt to achieve what they perceive to be a more socially desirable outcome. When making choices, people have a strong tendency to go with the default or pre-set option set out before them. Making the default the option the policy makers wish to be chosen makes it more likely that members of the general public will choose that option.

Individuals have a strong tendency to stick with the 'default' option, which is the outcome that occurs if they do not choose otherwise. A good example is provided by organ donation. If the default option is allowing organs to be donated in the event of the donor's death, more organs end up being donated than if people have instead to 'opt in' to give permission for their organs to be donated. From 20 May 2020, all adults in England have been considered organ donors when they die unless they have recorded a decision not to donate or are in one of the excluded groups.

A variation of default choice is **mandated choice**. This is where people are required by law to make a decision, for example to have an eye test when applying for a driving licence.

More information does not always lead to the best outcome; restricted choice may lead to a better outcome.

The use of **nudges** is an important part of the choice architecture. Nudges are used in government policy to change people's behaviour in a predictable manner without removing their freedom of choice. Nudges aim to change behaviour in ways which comply with desirable **social norms**. The use of nudges is an alternative to using laws to restrict or ban certain activities. Nudges attempt to influence people's decisions, and thereby increase social welfare, without the use of coercion. Since people are making decisions on the basis of incomplete information, without being nudged into desirable decisions, they are likely to make sub-optimal choices. The well-thought-out use of nudges can improve people's wellbeing.

Default choice An option that is selected automatically unless an alternative is chosen instead.

Choice architecture A framework setting out different ways in which choices can be presented to consumers, and the impact of that presentation on consumer decision making.

Mandated choice When people are required by law to make a decision.

Nudges Factors which encourage people to think and act in particular ways. Nudges try to shift group and individual behaviour in ways which comply with desirable social norms.

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

Exam tip

If possible, read Thaler and Sunstein's *Nudge: Improving Decisions about Health, Wealth and Happiness*, and also some of the Behavioural Insight Team's publications, which can be accessed on the internet.

People are influenced by how information is presented. **Framing** is the tendency for people to be influenced by the context in which the choice is presented when making a decision. Advertisers have for many years presented consumers with choices in a manner that frames their products in a favourable light. A statement on a cola bottle that its content is '95% sugar-free' works better for the manufacturer than a label stating 'contains 5% sugar'.

The framing effect explains how we alter our decisions depending on how information is presented to us. We will react in a completely different way depending on whether a choice is presented to us in the context of a loss or a gain.

Framing How something is presented (the 'frame') influences the choices people make.

Knowledge check 6

Give an example of how framing affects your behaviour in an economics class.

Summary

- The starting point for understanding individual economic decision making is understanding the nature of demand, rationality and maximising behaviour.
- Economists have traditionally assumed that individuals wish to maximise utility, which can be thought of as satisfaction, pleasure or fulfilment of need.
- The hypothesis (or 'law') of diminishing marginal utility lies behind the derivation of an individual's demand curve. A market demand, which is explained on page 20, is the sum of all the demand curves of all the individuals in the market.
- Individual economic decision making is affected by imperfect and asymmetric information.
- In recent years, behavioural economics has emerged to question many of the assumptions of traditional economic theory.
- Choice architecture and nudge theory lie at the heart of the ways in which behavioural architecture can influence economic policy making.

Price determination in a competitive market

The determinants of the demand for goods and services

Market demand

The demand curve shows the relationship between price and quantity demanded.

Normally when economists refer to demand, they mean **market demand**. This is the quantity of a good or service that all the consumers in the market wish to, and are able to, buy at different prices. Market demand is shown by a downward-sloping demand curve such as the one shown in Figure 3. However, the market demand curve is simply the sum of the individual demand curves of each of the consumers in the market.

Market demand The quantity of a good or service that all the consumers in a market are willing and able to buy.

The law of demand states that as a good's price falls, more is demanded. There is thus an inverse relationship between price and quantity demanded. The market demand curve in Figure 3 illustrates the law of demand. If the price starts off high, for example at P_1, household demand is Q_1. But if the price falls to P_2, demand increases to Q_2.

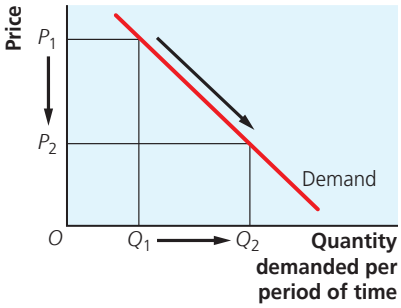


Figure 3 A market demand curve

Shifts of demand

It is easy to confuse the difference between a movement along (or adjustment along) a demand curve and a shift of a demand curve. When we draw a market demand curve to show how much of the good or service households plan to demand at various possible prices, we assume that all the other variables that may also influence demand are held unchanged or constant.

For example, an increase in income shifts demand curves rightwards — but only for normal goods. A **normal good** is defined as a good for which demand increases when income increases. By contrast, an **inferior good** is a good (such as poor-quality food) for which demand falls as income increases. If the good is inferior, an increase in income shifts the demand curve leftwards. Figure 4 shows a rightward shift of demand from D_1 to D_2, caused perhaps by a fall in the price of a good in

Normal good A good for which demand increases when income increases.

Inferior good A good for which demand falls as income increases.

joint demand (a complementary good) or by a successful advertising campaign for the product. Figure 4 shows that at price P_1, demand has increased from Q_1 to Q_2.

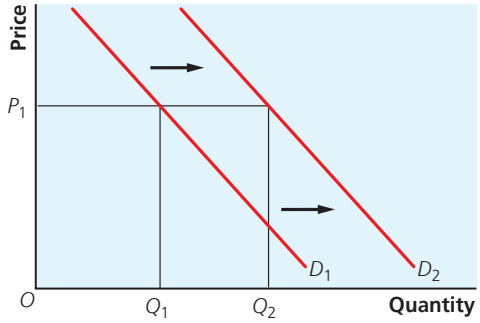


Figure 4 The effect of a rightward shift of demand in a market

Price, income and cross elasticities of demand

Whenever a change in one variable (such as a good's price) causes a change to occur in a second variable (such as the quantity of the good that consumers are prepared to demand), an **elasticity** can be calculated. Elasticity measures the proportionate response or change in a second variable following an initial change in a first variable. For example, if a 5% increase in price were to cause consumers to reduce their demand more than proportionately (say, by 10%), demand would be price elastic. If the response were less than proportionate (for example, a fall in demand of only 3%), demand would be price inelastic. And if the change in price were to induce an exactly proportionate change in demand, demand would be neither elastic nor inelastic with respect to the price change — this is called unit elasticity of demand.

The formulae for the three demand elasticities you need to know are:

$$\textbf{price elasticity of demand} = \frac{\text{proportionate change in quantity demanded}}{\text{proportionate change in price}}$$

$$\textbf{income elasticity of demand} = \frac{\text{proportionate change in quantity demanded}}{\text{proportionate change in income}}$$

$$\textbf{cross elasticity of demand for good A with respect to the price of good B} = \frac{\text{proportionate change in quantity of A demanded}}{\text{proportionate change in price of B}}$$

Price elasticity of demand

Price elasticity of demand measures consumers' responsiveness to a change in a good's price. The factors that influence price elasticity of demand are:

- Substitutability: this is the most important determinant of price elasticity of demand. When a substitute exists for a product, consumers can respond to a price rise by switching expenditure away from the good, buying instead the substitute whose price has not risen. Demand for necessities tends to be inelastic as they have few substitutes. A substitute good is a good that can be used instead of another good, e.g. apples as a substitute for pears.

Knowledge check 7

Explain the difference between a shift of demand and an adjustment of demand.

Exam tip

Exam questions often test whether you can distinguish between a shift of demand or supply and an adjustment of demand or supply along a demand or supply curve.

Elasticity A statistic which measures the proportional change of an economic variable in response to a change in another variable.

Price elasticity of demand Proportionate change in demand following an initial change in price.

Income elasticity of demand Proportionate change in demand following an initial change in income.

Cross elasticity of demand Proportionate change in demand for one good or service following an initial change in price of another good or service.

- Percentage of income: goods or services on which households spend a large proportion of their income tend to be in more elastic demand than small items on which only a fraction of income is spent.
- The 'width' of the market definition: the demand for Shell petrol is much more price elastic than the market demand for petrol as a generic product, i.e. the petrol produced by all the companies in the market for petrol (Shell, BP, Texaco etc.).
- Time: although there are exceptions, demand for many goods and services is more elastic in the long run than in the short run because it takes time to respond to a price change.

Income elasticity of demand

Income elasticity of demand — which measures how demand responds to a change in income — is always negative for an inferior good and positive for a normal good. The quantity demanded of an inferior good falls as income rises, whereas demand for a normal good rises with income. Normal goods are sometimes further subdivided into superior goods or luxuries, for which the income elasticity of demand is greater than unity, and basic goods, with an income elasticity of less than 1.

Cross elasticity of demand

Cross elasticity of demand measures the responsiveness of demand for one commodity to changes in the price of another good. The cross elasticity of demand between two goods or services indicates the nature of the demand relationship between the goods. There are three possibilities: joint demand (negative cross elasticity), competing demand or substitutes (positive cross elasticity) and an absence of any discernible demand relationship (zero cross elasticity).

The determinants of the supply of goods and services

Market supply

Market supply is the quantity of a good or service that all the firms or producers in the market plan to sell at different prices. The market supply curve in Figure 5 illustrates the law of supply, which states that as a good's price rises, more is supplied. If the price starts off low, for example at P_1, firms are willing to supply Q_1. But if the price rises to P_2, planned supply increases to Q_2.

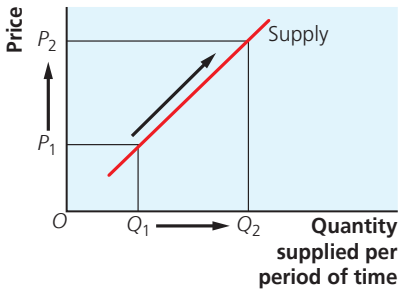


Figure 5 A market supply curve

Knowledge check 8

Why is the demand for Tesco's salt more price elastic than the demand for salt as a generic product?

Knowledge check 9

The income elasticity of demand for a good is +2.3. What does this statistic tell us about the good?

Knowledge check 10

The cross elasticity of demand for good A with respect to the price of good B is –0 8. What does this statistic tell us about the demand relationship between good A and good B?

Exam tip

Make sure you understand the different elasticities of demand and supply.

Market supply The quantity of a good or service that all the firms in a market are willing to sell at different prices.

A market supply curve shows the quantities of the good that all the firms in the market plan to supply at different possible prices, assuming the conditions of supply remain unchanged. If a condition of supply changes, the supply curve shifts to a new position. A rightward shift of supply is also known as an increase in supply, whereas a leftward shift is known as a decrease in supply.

The main conditions of supply are: costs of production (which include costs of wages, raw materials; energy and borrowing costs); technical progress; taxes imposed on firms, such as VAT, excise duties and the business rate; and government subsidies granted to firms.

Price elasticity of supply

In contrast to demand elasticities explained above, there is only one supply elasticity you need to know. This is price elasticity of supply, which measures how the supply of a good responds to an initial change in a good's price.

The formula for price elasticity of supply is:

$$\text{price elasticity of supply} = \frac{\text{proportionate change in quantity supplied}}{\text{proportionate change in price}}$$

The main factors determining price elasticity of supply are: the length of the production period; the availability of spare capacity; the ease of accumulating stocks, both of raw materials needed for production to take place and of finished goods waiting to be sold; the ease of switching between alternative methods of production; the number of firms in the market and the ease of entering the market; and the time period in question.

The determination of equilibrium market prices

Before explaining the determination of equilibrium market prices, we must define a **market**. A market, which is simply a voluntary meeting of buyers and sellers for the exchange of a good or service, is highly competitive when there is a very large number of buyers and sellers all passively accepting the ruling market price set, not by individual decisions but by the interaction of all those taking part in the market. Figure 6 illustrates the key features of a competitive market.

Market A voluntary meeting of buyers and sellers for the exchange of a good or service.

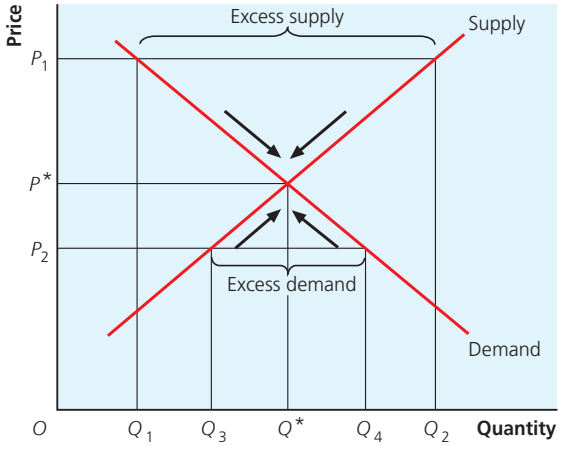


Figure 6 The price mechanism operating within a market

The demand curve shows the quantities of a good that households or consumers plan to purchase at different prices, and the supply curve shows how much firms or producers plan to supply at different prices. At all prices (except the **equilibrium price**) it is impossible for both households and firms simultaneously to fulfil their market plans. For example, at price P_1 firms would like to supply Q_2, but households are only willing to purchase Q_1. Planned supply is greater than planned demand, resulting in an **excess supply**. By contrast, at price P_2 households wish to buy Q_4 but firms restrict supply to Q_3 and **excess demand** results. At any price other than P^*, which is the equilibrium price, there will be either excess supply or excess demand, with either the firms or the households unable to fulfil their market plans. The market is in **disequilibrium** when there is excess supply or excess demand.

We now assume that firms respond to excess supply by reducing the price they are prepared to accept, while conversely households bid up the price to eliminate excess demand. The price falls or rises until equilibrium is achieved. The equilibrium price is the only price that satisfies both households and firms, which consequently have no reason to change their market plans. At P^* planned demand = planned supply.

Three conditions are necessary for a market to operate successfully:

- The individual buyers and sellers decide what, how, how much, where and when to trade or exchange.
- They do so with reference to their self-interest and to the alternative opportunities open to them. The exchange must be voluntary; if one party forces a transaction on the other, it is not a market transaction.
- Prices convey information to buyers and sellers about self-interest and opportunities. For a market to allocate resources among different types of activity and to coordinate economic activity throughout the economy, prices must respond to the forces of supply and demand.

Nearly 250 years ago, the great classical economist Adam Smith, who is often called the 'father of economics', described how the invisible or hidden hand of the market, operating in competitive markets and through the pursuit of self-interest, achieves an allocation of resources that is also in society's interest. This remains the central view of all **free-market economists**, i.e. those who believe in the virtues of a competitive market economy subject to minimum government intervention.

The interrelationship between markets

So far, we have looked at how the price mechanism operates in a competitive market. We have seen how shifts of either the demand curve for the good or the good's supply curve disturb market equilibrium and trigger an adjustment process to establish a new equilibrium.

Very often shifts of curves are caused by events taking place in other markets in the economy. They can be caused by a change of price of a good in **joint supply** or, on the demand side, by a change in the price of a good in **joint demand**, a **substitute good**, a good in **composite demand**, or a good in derived demand (see page 54).

Equilibrium price The price that clears the market, at which there is no excess demand or excess supply.

Excess supply When planned supply > planned demand.

Excess demand When planned supply < planned demand.

Disequilibrium When there is excess supply or excess demand.

Exam tip

Equilibrium and its opposite, disequilibrium, are two of the most important economic concepts you need to understand.

Free-market economists Those who believe that resource allocation should be left solely or mostly to the market mechanism (or price mechanism) and to private enterprise or free enterprise

Joint supply Production of one good also leads to the supply of a by-product.

Joint demand Two or more goods or services when they are demanded together. Jointly demanded goods are complementary, e.g. computers and keyboards.

Substitute goods Goods which a consumer sees as the same or similar to each other, e.g. olive oil and sunflower oil.

Joint supply occurs when production of one good also leads to the supply of a by-product. For example, a rise in the price of one of the two goods in joint supply leads to a shift of the supply curve of the other good. Competing supply, meanwhile, occurs when, because raw materials are used to produce one good, they cannot be used to produce another good.

When the supply curve of a good shifts, for example because of a joint supply or competing supply relationship with a good produced and sold in another market, the extent to which price and quantity respectively change in the market depends on price elasticity of demand for the good. The possibilities are shown in Figure 7.

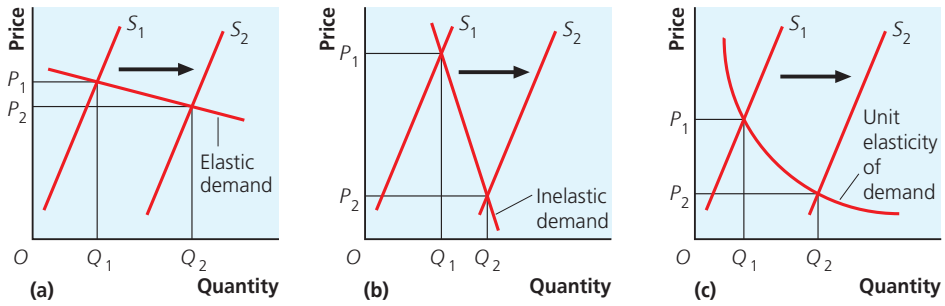


Figure 7 The extent to which price and quantity change following a shift of supply depends on price elasticity of demand

When demand is elastic (Figure 7a), the quantity bought and sold adjusts much more than price. The reverse is true when demand is inelastic (Figure 7b). Finally, when demand is unit elastic (depicted in the rectangular hyperbola in Figure 7c), price and quantity change by equal percentages.

An increase in the price of a good in joint demand (or a **complementary good**) has the opposite effect to an increase in the price of a substitute good (or a good in competing demand). Figure 8 shows a rightward shift of demand from D_1 to D_2, caused perhaps by a fall in the price of a good in joint demand (a complementary good) or by a rise in the price of a substitute good. Before the shift of demand, P_1 was the equilibrium price. Following the shift of demand, this is no longer the case. Planned demand is greater than planned supply and there is excess demand of Q_2 minus Q_1 in the market. To relieve the excess demand, the price rises to a new equilibrium at P_2.

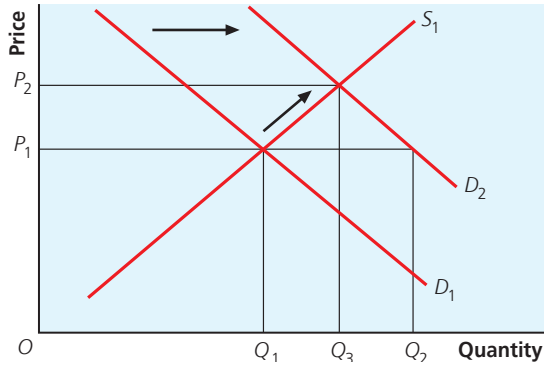


Figure 8 The effect of a rightward shift of demand in a market

Composite demand Demand for a good that has multiple different uses. If demand for one product that uses the commodity rises, the supply of other products using the commodity will fall. For example, steel is used in the production of cars and ships, so if more ships are built, less steel may be available for car production.

Complementary goods Goods that are in joint demand (they go together), such as cars and petrol.

It is easy to confuse competing demand, which occurs when two goods are substitutes, with composite demand and derived demand. Composite demand is demand for a good that has more than one use. This means that an increase in demand for one use of the good reduces the supply of the good for an alternative use. By contrast, derived demand for a good occurs when a good is necessary for the production of other goods. The demand for capital goods such as machinery and raw materials is derived from the demand for consumer goods or finished goods. If the demand for cars falls, so does the demand for car components such as tyres.

Knowledge check 11

What is meant by equilibrium in economics?

Knowledge check 12

Give another example of a good that is in composite demand.

Summary

- A market is a meeting of buyers and sellers for the purpose of exchanging goods or services.
- Two of the main features of a market are a demand curve and a supply curve.
- Market equilibrium occurs when demand equals supply.
- Disequilibrium in a market means there is either excess demand or excess supply in the market.
- In a competitive market, the market mechanism (or price mechanism) causes the price to rise or fall to get rid of excess demand or excess supply.
- Market equilibrium may be disturbed by a shift in either the demand curve or the supply curve.
- A change in a condition of demand, such as consumers' incomes, causes the demand curve to shift to a new position.
- Likewise, a change in a condition of supply, such as costs of production, causes the supply curve to shift to a new position.
- Joint demand is when two goods are demanded together; competing demand is when two goods are substitutes for each other.
- Composite demand is demand for a good that has more than one use, while derived demand means that the demand for one good is derived from the demand for another good.
- Joint supply occurs when the production of one good affects the production of a by-product produced from the same raw material.
- Elasticity measures the proportionate response or change in a second variable following an initial change in a first variable.
- The four elasticities you need to know are price elasticity of demand, income elasticity of demand, cross elasticity of demand and price elasticity of supply.
- When demand is price elastic, a change in price leads to a more than proportionate change in demand. In this situation, the price elasticity of demand is greater than 1 (or unity).
- When demand is price inelastic, a change in price leads to a less than proportionate change in demand. In this situation, the price elasticity of demand is less than 1 (or unity).
- When price elasticity of demand equals 1 (or unity), a change in price leads to an exactly proportionate change in demand. In this situation, the price elasticity of demand is neither elastic nor inelastic.
- The existence of substitutes is the main determinant of price elasticity of demand.
- With income elasticity of demand, the plus or minus sign of the elasticity statistic tells us whether the good is a normal good or an inferior good.
- With cross elasticity of demand, the plus or minus sign of the elasticity statistic tells us whether the two goods are substitutes or complementary goods in joint demand.

Production, costs and revenue

Production and productivity

Production is simply the process by which inputs are converted into outputs. The inputs into the production process are the four factors of production (land, labour, capital and enterprise) which are listed in Figure 9. It is easy to confuse **productivity** with production. While closely related, they do not have the same meaning.

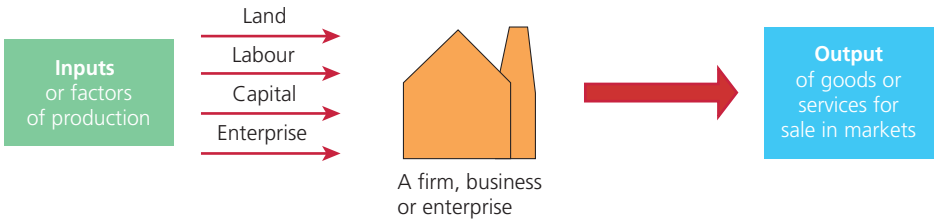


Figure 9 A firm undertaking production

For most purposes, productivity usually means labour productivity, which is output per worker. Productivity is a key concept in A-level economics, particularly in macroeconomics, where we look at the UK's **productivity gap**, which is the difference in productivity levels between the UK and competitor countries.

Specialisation, division of labour and exchange

If we assume that the capital and land employed by a firm are fixed and cannot be altered, at least in the short run, the only way a firm can increase production is by employing more factors of production such as labour. To start with, as more workers are employed, labour productivity (output per worker) may rise, as the firm benefits from **specialisation** and the **division of labour**, i.e. different workers specialising in different tasks. In the economy as a whole, different firms and industries also specialise in producing different goods and services. This, of course, necessitates **trade** and **exchange**, which take place in the economy's markets.

Nearly 250 years ago, Adam Smith explained one of the most fundamental of all economic principles: the benefits of specialisation or the division of labour. According to Smith, there are three main reasons why a factory's total output can be increased if workers perform specialist tasks rather than if each worker attempts all the tasks themselves:

- A worker will not need to switch between tasks, so time will be saved.
- More and better machinery or capital can be employed.
- The 'practice makes perfect' argument, i.e. the more time spent on specialist tasks, the more efficient (productive) workers become.

On the downside, however, workers might become deskilled and find their jobs uninteresting if all that is involved is constant repetition of a single task, such as attaching wheels to cars passing by on an assembly line. This is a disadvantage of the division of labour.

Productivity Output per unit of factor input.

Productivity gap The difference in productivity levels in one country and in competitor countries.

Specialisation Workers performing different tasks in the production process.

Division of labour Production divided into different tasks, undertaken by workers with different skills.

Trade Selling or buying goods or services.

Exchange Selling goods or services in return for others.

For specialisation to be economically worthwhile for those taking part in the division of labour, a system of trade and exchange is necessary. This is because workers who completely specialise cannot enjoy a reasonable standard of living if forced to consume only what they produce. The solution is to produce more than what the worker actually needs and then to trade the surplus for that produced by others. Specialisation, the division of labour, trade and exchange are central to the functioning of a modern economy.

The law of diminishing returns and returns to scale

To understand the law of diminishing returns and returns to scale, we must first appreciate the difference between the short run and the long run. In microeconomic theory, the **short run** is defined as the time period in which, in the course of production, at least one of the factors of production is fixed and cannot be varied. The only way in which a firm can increase output in the short run is by adding more variable factors of production, such as labour, to fixed factors of production such as capital. As this happens, production is affected by the **law of diminishing returns**. In the **long run**, by contrast, all the factors of production are variable and none is fixed.

Explaining short-run production theory

Suppose a small manufacturing firm decides to employ only one worker. The worker must be a jack-of-all-trades, doing all tasks involved in production. But if more workers are hired, output can rise at a faster rate than the number of workers employed. This is because the workers benefit from specialisation and the division of labour, as production tasks are divided between the workers. In this situation, the **marginal return of labour** will increase. Marginal return of labour (marginal product) is the increase in output that results from adding an extra worker to the labour force. By contrast, the **average return of labour** is total output divided by the number of workers employed and the **total return of labour** is the total output of the labour force.

The law of diminishing returns

However, eventually, as more and more workers are combined with the firm's fixed capital, the benefits of further specialisation and division of labour come to an end. The law of diminishing returns (also known as 'the law of diminishing marginal productivity') sets in when the marginal product of labour starts to fall — that is, when one more worker adds less to total output than the previous worker who joined the labour force.

Exam tip

Make sure you don't confuse the two words 'returns' and 'revenue'. One way of avoiding confusion is to prefer the term 'marginal product' to the term 'marginal return'.

Short run The time period during which at least one of the factors of production is fixed and cannot be varied.

Law of diminishing returns A short-term law which states that as a variable factor of production is added to a fixed factor of production, eventually both the marginal and average returns to the variable factor will begin to fall.

Long run The time period during which all the factors of production are variable and none is fixed.

Marginal return of labour The change in the quantity of total output resulting from the employment of one more worker, holding all the other factors of production fixed.

Average return of labour Total output divided by the total number of workers employed.

Total return of labour Total output produced by all the workers employed by a firm.

Returns to scale

The law of diminishing returns is a short-run law that does not operate in the long run, when a firm can increase the scale of all its inputs or factors of production. You must not confuse the short-run law of diminishing returns with **returns to scale**, which occur only in the long run.

Returns to scale The rate by which output changes if the scale of all the factors of production is changed.

The difference between increasing, constant and decreasing returns to scale

With returns to scale, there are three possibilities:

- **Increasing returns to scale.** An increase in the scale of all the factors of production causes a more than proportionate increase in output.
- **Decreasing returns to scale.** An increase in the scale of all the factors of production causes a less than proportionate increase in output.
- **Constant returns to scale.** An increase in the scale of all the factors of production causes an exactly proportionate increase in output.

Exam tip

'Increasing returns to scale' and 'economies of scale' are often treated as interchangeable terms, although, strictly speaking, returns to scale are part of long-run production theory whereas economies of scale are part of long-term cost theory. You must understand the relationship between returns to scale and economies or diseconomies of scale.

Knowledge check 13

What is the difference between the law of diminishing marginal productivity and decreasing returns to scale?

Costs of production

As well as confusing production and productivity, economics students often confuse production and **costs**. As previously explained, production simply converts inputs into outputs. It does not consider the money cost of using inputs such as capital and labour.

Costs The money a firm has to pay out when hiring the services of the factors of production.

The difference between fixed and variable costs

As we have mentioned, in the short run, defined as the time period in which at least one factor of production is held fixed, costs of production divide into fixed and variable costs. Fixed costs are the costs a firm incurs when hiring or paying for the fixed factors of production. Capital is usually assumed to be a fixed factor of production. Variable costs, such as the costs of hiring many types of labour and buying raw materials, change as the firm's level of output changes.

Knowledge check 14

Give two examples of fixed costs and two examples of variable costs, other than those mentioned above.

At any level of output, a firm's total costs of production can be calculated by adding up the cost of producing each extra unit of output. Average costs, by contrast, are total cost divided by total output. Likewise, average fixed cost is total fixed cost divided by the level of output.

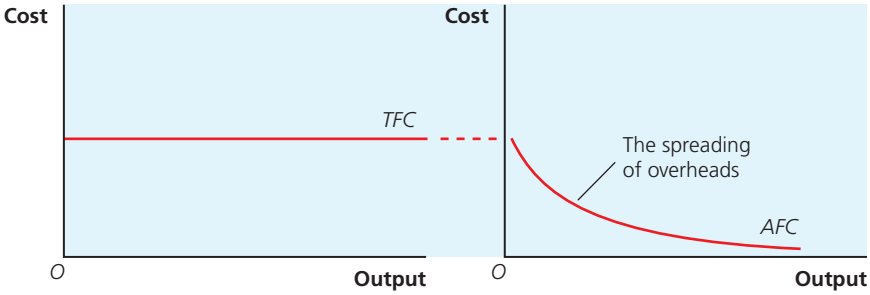


Figure 10 Total fixed cost and average fixed cost

The left-hand panel of Figure 10 shows that total fixed cost (*TFC*) remains unchanged when output increases. Total fixed cost can be thought of as the cost of overheads, such as the rent a firm pays for leasing its buildings. However, the right-hand panel of the figure shows **average fixed cost** (*AFC*) per unit of output falling as output increases, since overheads are spread over a larger output.

The difference between marginal, average and total cost

Figure 11 shows an **average variable cost** (*AVC*) curve, with a **marginal cost** (*MC*) curve rising and cutting through the lowest point on the *AVC* curve. Marginal cost is the extra cost of producing one more unit of output. The shape of the *MC* curve is explained by marginal productivity theory. As long as the marginal productivity of labour is increasing, then, assuming all workers are paid the same wage rate, the cost of producing an extra unit of output falls. Hence marginal costs fall at low levels of output. But as soon as the law of diminishing returns sets in, an extra worker hired adds less to total output than the previous worker taken on. Total costs of production rise faster than output, leading to rising marginal costs.

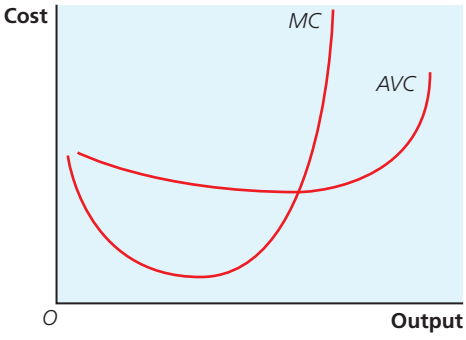


Figure 11 Average variable cost and marginal cost

The left-hand panel of Figure 12 shows how the firm's **average total cost** (*ATC*) curve is arrived at by adding up the *AFC* and *AVC* curves. The right-hand panel of Figure 12 shows the *ATC* curve without its two 'building blocks' (*AFC* and *AVC*). The *ATC* curve is *U*-shaped, showing that average total costs per unit of output first fall and later rise as output is increased. *ATC* must eventually rise because, at high levels of output, any further spreading of fixed costs is insufficient to offset the impact of diminishing returns upon variable costs of production. Eventually, rising marginal costs (which, as we have explained, result from diminishing marginal returns) must cut through and pull up the *ATC* curve.

Average fixed cost
Total cost of employing the fixed factors of production to produce a particular level of output, divided by the size of output: *AFC* = *TFC* ÷ output.

Average variable cost
Total cost of employing the variable factors of production to produce a particular level of output, divided by the size of output: *AVC* = *TVC* ÷ output.

Marginal cost Addition to total cost resulting from producing one additional unit of output.

Exam tip

Make sure you understand the relationship between the marginal returns curve and the short-run marginal cost curve.

Average total cost
Total cost of producing a particular level of output, divided by the size of output; often called 'average cost': *ATC* = *AFC* + *AVC*.

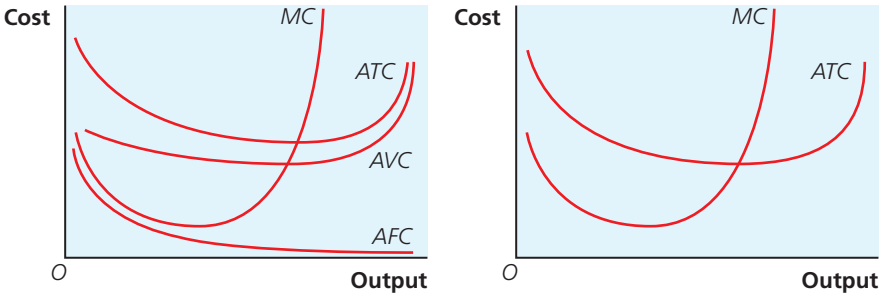


Figure 12 The average total cost curve results from adding *AVC* to *AFC*

The difference between short-run and long-run costs

So far, we have looked at short-run costs of production, which are the costs incurred when at least one factor of production is held fixed. We now move on to long-run costs of production. These are the costs incurred when all the factors of production are variable, in which case the scale or size of the firm can be changed.

Just as the short-run law of diminishing returns explains rising marginal costs and (eventually, when marginal costs cut through average total costs) rising average total costs, so we shall now use long-run production theory concepts to explain the firm's **long-run average cost** (*LRAC*) curve. This is illustrated in Figure 13. If, as the firm increases the size or scale of all its factors of production, it benefits from increasing returns to scale, the *LRAC* curve falls. Falling long-run average costs result when a firm benefits from an **economy of scale**. Conversely, a **diseconomy of scale** causes long-run average costs to rise. Note that a number of short-run average total cost (*SRATC*) curves, labelled $SRATC_1$, $SRATC_2$, $SRATC_3$ and $SRATC_4$, have been drawn in Figure 13, and that the *LRAC* curve touches (or is tangential to) each *SRATC* curve. Each *SRATC* curve represents a particular short-run size of firm.

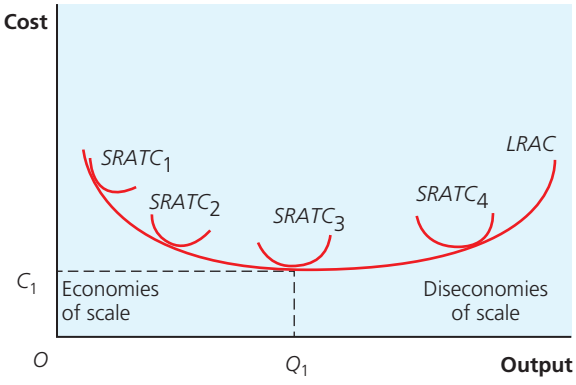


Figure 13 A long-run average cost curve

Figure 13 shows a *U*-shaped *LRAC* curve in which economies of scale and falling long-run average costs give way beyond $SRATC_3$ to diseconomies of scale. $SRATC_3$ represents the lowest unit cost and most **productively efficient size of firm**. This is also sometimes called the optimum size of firm. However, other shapes of *LRAC* curve are also possible.

Knowledge check 15

Why are a firm's average variable cost (*AVC*) curve and its average total cost (*ATC*) curve U-shaped in the economic short run?

Long-run average cost Total cost of producing a particular level of output divided by the size of output when all the factors of production are variable.

Economy of scale When a firm has falling long-run average costs when it increases in size or scale.

Diseconomy of scale When a firm has rising long-run average costs when it increases in size or scale.

Knowledge check 16

How might an increase in wage rates, compared with the price of capital, affect a firm's costs of production and its demand for different factors of production?

Exam tip

Make sure you can illustrate economies and diseconomies of scale on a cost curve diagram.

Productively efficient size of firm The size or scale of firm which minimises long-run average costs.

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

Figures 11 and 12 illustrate the important relationship between any marginal curve and the average curve plotted from the same data:

- When the marginal > the average, the average rises.
- When the marginal < the average, the average falls.
- When the marginal = the average, the average is constant, neither rising nor falling.

The relationship between marginal and average curves has several economic applications: marginal and average product curves (in production theory); marginal and average cost curves (illustrated in Figures 11 and 12); and, as we shall explain shortly, marginal and average revenue curves. You must understand this relationship. It does not state that an average curve will rise when the related marginal curve is rising, or that the average curve must fall when the related marginal curve falls. Look again at Figure 11. After diminishing returns set in, the *MC* curve starts to rise, but the *AVC* curve continues to fall as long as marginal costs are below average variable costs. Eventually, however, the *MC* curve rises through the *AVC* curve, causing the *AVC* curve also to rise. As a result, the *AVC* curve is U-shaped, with the *MC* curve cutting through the curve at its lowest point.

At any level of output, total cost is the addition of the marginal costs of producing each successive unit of output. When marginal costs fall, the rate of increase of total cost slows down. But when marginal costs begin to rise, the rate of increase of total cost speeds up or accelerates.

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

Factor prices such as wages and the price of capital goods are important business costs, of course. If they increase, costs of production increase and cost curves shift upwards. And if labour becomes more expensive than capital, over time firms alter the nature of production, in the long run employing less labour and more capital. Conversely, if wages fall relative to the price of capital, firms will employ more labour and less capital.

To remind you, labour productivity is output per worker. If labour productivity increases, the real cost of labour falls, so more workers may be employed. Two points to note, however, are, first, increasing labour productivity mainly results from a firm employing more and better capital goods, and second, with increased labour productivity, fewer workers are needed to produce the same level of output.

Economies and diseconomies of scale

Economies and diseconomies of scale are illustrated in Figure 14, which shows how a firm's average costs of production may change as it increases the size or scale of its operations. Economies of scale occur when a firm experiences falling average costs as it increases its size and scale and produces more output. If beyond a certain size of firm (point *X* in Figure 14) average costs begin to rise, diseconomies of scale set in.

Exam tip

Make sure you don't confuse the economic short run and the economic long run, and short-run cost curves and long-run cost curves.

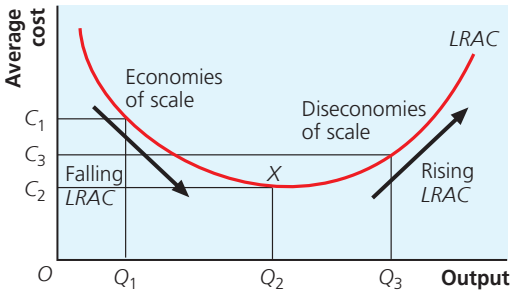


Figure 14 A firm's long-run average cost curve and economies and diseconomies of scale

There are various types of economy of scale, such as **technical economies of scale** and **managerial economies of scale**.

- An example of a technical economy of scale is a volume economy. When a warehouse is doubled in dimension, its storage capacity increases eightfold. As a result, average storage costs fall as the size or scale of storage capacity increases. However, diseconomies of scale could also set in as storage capacity increases. Stored materials may become mislaid or lost more easily, and the firm owning the warehouse may be tempted to store unnecessary stocks of raw materials or goods awaiting sale.
- Managerial economies of scale occur because large firms can employ specialist and more highly skilled managers. However, too many managers and a breakdown in communication between them may result in managerial diseconomies of scale.

The difference between internal and external economies of scale

The types of economy of scale described above are all **internal economies of scale**, which result from the growth of the firm itself. By contrast, **external economies of scale** result from the growth of the whole industry and market in which the firm exists. A similar distinction is made between **internal diseconomies of scale** and **external diseconomies of scale**.

Reasons for diseconomies of scale

Among the diseconomies of scale which firms suffer, and which raise long-run average production costs, are poor communication and coordination problems. With regard to the former, firms find it difficult to maintain an effective flow of information between the different parts of a large firm. With regard to the latter, large firms often find it much harder to coordinate different activities undertaken by a large labour force than a small firm does.

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

To remind you, returns to scale are a part of long-run production theory. Economies and diseconomies of scale, by contrast, are a part of long-run cost theory. The two are related. For example, increasing returns to scale are a cause of technical economies of scale.

Technical economies of scale Large firms can afford to invest in expensive and specialist capital machinery. For example, a supermarket chain such as Asda can invest in technology that improves stock control.

Managerial economies of scale These occur when large firms can afford specialist managers with particular skills.

Internal economies of scale Falling long-run average costs resulting from the growth of the firm itself.

External economies of scale Falling long-run average costs resulting from the growth of the whole industry in which the firm exists.

Internal diseconomies of scale Rising long-run average costs resulting from the growth of the firm itself.

External diseconomies of scale Rising long-run average costs resulting from the growth of the whole industry in which the firm exists.

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

Again, to remind you, if you refer back to Figure 14, you will see that economies of scale are shown by falling long-run average costs, while diseconomies of scale are depicted by rising long-run average costs. Figure 14 shows the 'textbook' *LRAC* curve, but other shapes are possible. One of these is shown in Figure 15.

The L-shaped long-run average cost curve

Figure 15 shows an L-shaped *LRAC* curve. The size of firm represented by $SRATC_3$ is sited at the point on the *LRAC* curve beyond which no more economies of scale are possible. But there are no diseconomies of scale, so all sizes of firms beyond this firm size are equally productively efficient.

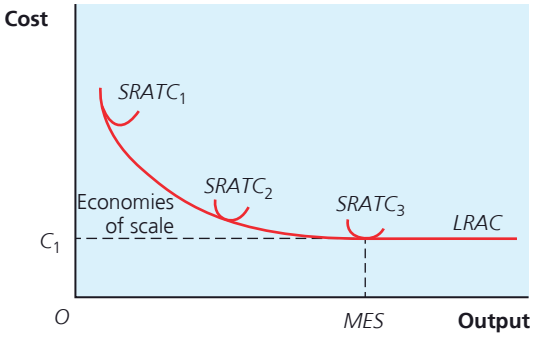


Figure 15 The 'L'-shaped *LRAC* curve

The concept of the minimum efficient scale of production

The lowest point on the *LRAC* curve shows the **minimum efficient scale (*MES*)** firm. As noted in the previous section, beyond this size of firm, no more economies of scale are possible. In Figure 15, the *LRAC* curve is flat to the right of the *MES* size of firm. By contrast, with the U-shaped *LRAC* curve shown in Figure 14, diseconomies of scale set in to the right of *MES*.

Marginal, average and total revenue

The difference between marginal, average and total revenue

Revenue is the money a firm receives from selling its output. **Marginal revenue** is the addition to **total revenue** received when an extra unit of output is sold. At any level of output, **average revenue** is total revenue divided by output, while total revenue is the addition of the marginal revenues received from each unit of output.

Why the average revenue curve is the firm's demand curve

To explain why its average revenue curve is the demand curve for a firm's output, we shall take the example of a perfectly competitive firm. (Perfect competition and other forms of market structure are explained in the next topic, starting on page 38.) A perfectly competitive firm's revenue curves are derived from the assumptions that

Minimum efficient scale (*MES*) Achieved after economies of scale have been benefited from to the full.

Revenue The money a firm receives from selling its output.

Marginal revenue Addition to total revenue resulting from the sale of one more unit of the product.

Total revenue The sum of all the marginal revenues received by a firm.

Average revenue Total revenue divided by output.

the firm can sell whatever quantity it wishes at the ruling market price, but that it cannot influence the ruling market price by its own action.

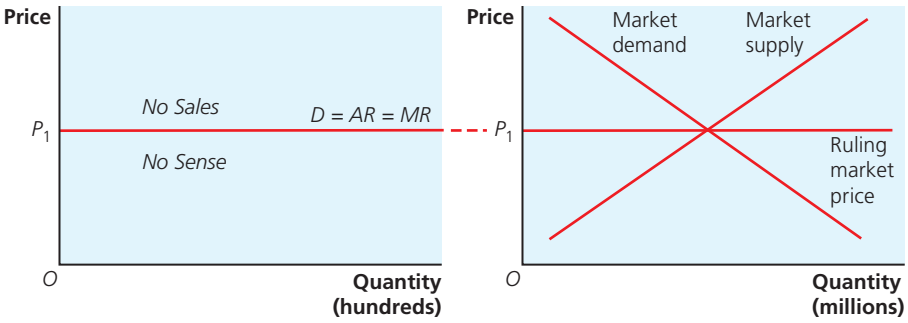


Figure 16 Deriving the *AR* and *MR* curves of a perfectly competitive firm

The right-hand panel of Figure 16 shows the whole of a perfectly competitive market, whereas the left-hand panel shows the situation facing a single firm within the market. The ruling market price P_1 is determined in the right-hand panel, where market demand equals market supply. In the left-hand panel, each firm faces an infinitely elastic or perfectly elastic demand curve located at P_1, the ruling price set by market forces in the whole market.

Consider also the two slogans, '*No Sales*' and '*No Sense*', which are respectively above and below P_1. Suppose, first, that the firm tries to set a price above P_1. Possessing perfect market information, the firm's customers immediately stop buying, deciding instead to buy the identical products (which are perfect substitutes) available at P_1 which are produced by other firms in the market — hence '*No Sales*'. But if the firm can sell as much as it wishes at the ruling price, there is no point in reducing the price below P_1. No extra sales are gained, but the firm loses sales revenue (and profit) — hence '*No Sense*'. We can conclude that a perfectly competitive firm is a **price-taker**, passively accepting, but unable to influence, the ruling market price.

Price-taker A firm that is so small that it has to accept the ruling market price. If the firm raises its price, it loses all its sales; if it cuts its price, it gains no advantage.

As well as being the perfectly elastic demand curve for the firm's output, the horizontal line drawn through P_1 is the perfectly competitive firm's average revenue (*AR*) curve and its marginal revenue (*MR*) curve. Every time it sells one more unit of output, total sales revenue rises by the price at which the extra unit is sold (P_1). Thus, marginal revenue is P_1. And because revenue per unit sold is always the same however much is sold, average revenue is P_1 at all levels of output and sales.

The relationship between average and marginal revenue

The relationship between average and marginal revenue is explained by the following mathematical rule:

- When the marginal > the average, the average rises.
- When the marginal < the average, the average falls.
- When the marginal = the average, the average is constant, neither rising nor falling.

The rule applies to revenue and returns curves as well as to cost curves.

Figure 16 shows that in perfect competition, when *MR* = *AR*, the average is constant, neither rising nor falling.

Knowledge check 17

Distinguish between marginal revenue and marginal returns.

The relationship between marginal revenue and total revenue

Since total revenue is the sum of the marginal revenues received from each unit of sales, the total revenue (*TR*) curve rises at a constant rate in perfect competition because *MR* is the same for each unit of sales. In monopoly, by contrast, *MR* falls as sales increase, with the result that the *TR* curve rises at a slower rate than output. Eventually, *MR* may become negative, in which case *TR* falls.

Exam tip

An exam question may ask you to plot cost and revenue curves from data presented in the question.

Profit

Students often confuse **profit** and revenue, mistakenly believing that the two terms have the same meaning. However, profit and revenue are different. Profit is the difference between the sales revenue a firm receives when selling the goods or services it produces and the costs of producing the goods.

Normal and abnormal profit

In the next topic on market structures, we shall see how economists distinguish between two theoretical profit concepts: **normal profit** and **abnormal profit**. Normal profit is the minimum level of profit necessary to keep existing firms in production, while being insufficient to attract new firms into the market. Because a firm must make normal profit to stay in production, economists treat normal profit as a cost of production, including it in a firm's average cost curve. In the long run, firms unable to make normal profit leave the market.

Abnormal profit, which is also known as 'supernormal profit', is any extra profit over and above normal profit. In the long run, and in the absence of entry barriers, abnormal profit performs the important economic function of attracting new firms into the market. However, in highly competitive markets, which in the real world means those that approximate to perfect competition, the entry of new firms into the market brings down the market price and whittles away abnormal profit, until in long-run or 'true' equilibrium, surviving firms make normal profit only.

The role of profit in a market economy

Profit performs a number of roles in a market economy. These include the creation of business, worker and shareholder incentives. Profit also influences the allocation of resources; it is an efficiency indicator; and it is a reward for innovation and for risk taking. Finally, profits provide an important source of business finance.

Technological change

Technology is knowledge put to practical use to solve problems facing human societies. **Technological change**, by contrast, involves improving existing technologies and the development of completely new technologies, both to improve existing products and the processes involved in making the products, and to develop completely new products and processes.

Knowledge check 18

When the *MR* curve lies below the *AR* curve, what must happen to the *AR* curve?

Profit The difference between sales revenue and costs of production.

Normal profit The minimum profit a firm must make to stay in business, which is, however, insufficient to attract new firms into the market.

Abnormal profit Profit over and above normal profit. Also known as 'supernormal profit'.

Technological change The overall effect of invention, innovation and the diffusion or spread of technology in an economy.

The difference between invention and innovation

Invention refers to advancements in pure science, whereas **innovation** is the application of the new knowledge created by invention to production.

Through its diffusion into the economy, technological change affects methods of production, productivity, efficiency and firms' costs of production. With regards to methods of production, much recent technological progress has centred on **mechanisation** giving way to **automation**. And since this involves machines such as robots, rather than human beings, operating other machines, labour productivity (output per worker) obviously increases. The reorganisation of methods of production associated with automation causes average costs to fall, which leads to improvements in productive efficiency.

The effects of technological change

In the economic sphere, technological change leads to the development of completely new markets, to changes in market structure, and also to the destruction of existing markets. It can affect methods of production, productivity, efficiency and firms' costs of production. Technological change can also influence the structure of markets and generally improves economic efficiency. By increasing productivity, over time technological changes shift both short-run and long-run cost curves downwards. Firms create new products that satisfy people's needs and wants. Over time, their actions improve social welfare.

Technological change is also part of a process known as **creative destruction**, in which economic growth occurs in the economy as a result of new innovations creating more economic value than that being destroyed by the decline of the technologies that the new innovations replace. Over time, societies that allow creative destruction to operate grow more productive and richer; their citizens benefit from new and better products and higher living standards.

Invention Making something that did not exist before at all.

Innovation improves on or makes a significant contribution to something that has already been invented, thereby turning the results of invention into a product.

Mechanisation Process of moving from a labour-intensive to a more capital-intensive method of production, employing more machines and fewer workers.

Automation Automatic control, where machines such as robots operate other machines.

Creative destruction The economy evolving and renewing itself over time through new technologies and innovations replacing older technologies and innovations.

Summary

- Production and cost theory divide into short-run and long-run theory.
- The key concept in short-run production theory is the law of diminishing returns, also known as 'the law of diminishing marginal productivity'.
- In the short run, the marginal cost curve and the average variable cost curve are derived from the law of diminishing marginal (and average) returns.
- Assuming that the variable factors of production experience diminishing returns, the average variable cost (*AVC*) and the short-run average total cost (*SRATC*) curves are U-shaped.
- The key concept in long-run production theory is returns to scale.
- The key concepts in long-run cost theory are economies and diseconomies of scale.
- The long-run average cost (*LRAC*) curve may be U-shaped, but other shapes are possible.
- Profit is total sales revenue minus total costs of production. Normal profit is just sufficient to keep incumbent firms in the market but is insufficient to attract new firms into the market.
- Abnormal, or supernormal, profit is any profit over and above normal profit.
- Diseconomies of scale are rising average costs of production when the size or scale of the firm increases.
- Technological change involves improving existing technologies and the development of completely new technologies, both to improve existing products and the processes involved in making the products, and to develop completely new products and processes.

Perfect competition, monopoly and imperfectly competitive markets

Market structures

The spectrum of competition

Whereas cost curves derive from production theory and the cost of hiring the factors of production, a firm's revenue curves depend on the market structure in which it sells its output. Figure 17 shows the main market structures recognised by economists.

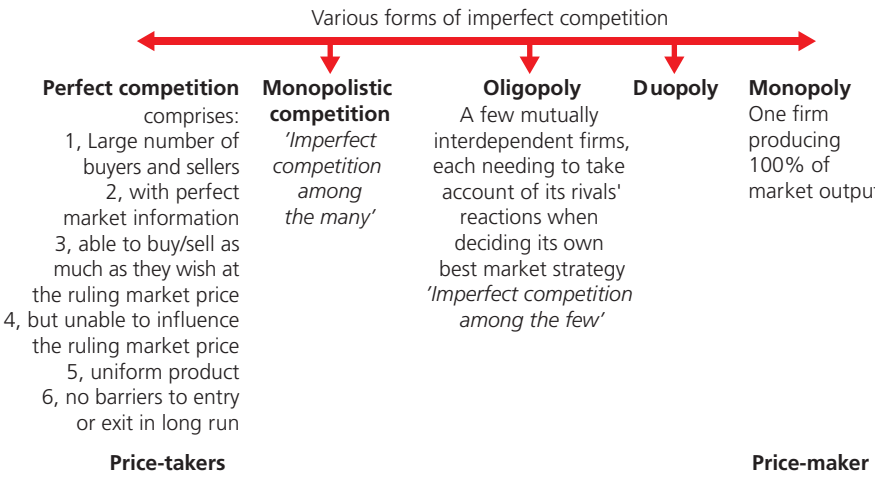


Figure 17 The main market structures ranging from perfect competition through imperfect competition to monopoly

Perfect competition and **monopoly** are at opposite ends of the spectrum shown in Figure 17. In a perfectly competitive market there is a large number of firms. By contrast, in monopoly (or, strictly, pure monopoly), a single firm produces the whole of the output of a market or industry. A pure monopolist faces no competition at all, since there are no other firms to compete against. Monopolists do, however, usually face some competitive pressures, both from substitute products and sometimes also from outside firms trying to enter the market to destroy their monopoly position. Pure monopoly is exceedingly rare. Often the word 'monopoly' is used in a looser sense to refer to any highly concentrated market, in which one firm is dominant.

Every market structure between the extremes of perfect competition and monopoly illustrates a form of imperfect competition. There are two main forms of **imperfect competition**: **monopolistic competition** and **oligopoly**. Because there is a large number of firms in monopolistic competition, the market structure is often called 'imperfect competition among the many'. By contrast, an oligopoly is a market dominated by a few large interdependent firms. Interdependence means that an

Perfect competition Best defined by the six conditions or characteristics of the market structure listed in Figure 17.

Monopoly A firm producing 100% of the market output.

Imperfect competition A competitive market lying between the extremes of perfect competition and monopoly.

Monopolistic competition A market structure closer to perfect competition than to monopoly.

Oligopoly A market or industry in which a few large firms dominate.

oligopolist has to take account of the likely reactions of the other firms when deciding price and output. The market structure is often called 'imperfect competition among the few'. **Duopoly** is a special case of oligopoly in which there are just two dominant firms.

Duopoly Extreme case of oligopoly with only two dominant firms in the market.

Perfect competition is an abstract economic model that does not actually exist in any real-world market. This is because the conditions listed in Figure 17 which define perfect competition are too demanding and never occur together simultaneously. Competitive markets in the real world are examples of imperfect competition rather than perfect competition, though some highly competitive markets, such as commodity and financial markets, possess some of the features of perfect competition.

Distinguishing between different market structures

There are a number of factors which are used to distinguish between different market structures. The factors include the number of firms in the market, the degree of **product differentiation** and ease of entry. While there are many small firms in the UK economy today, relatively few produce just a single good or service. Most firms — particularly large and medium-sized business enterprises, but also small businesses — undertake varying degrees of product differentiation. Firms often produce a range of relatively similar products, some of which compete with each other, but others of which are aimed at differentiated market segments. Mobile phones provide examples of both. Apple is well known for introducing two new smartphones roughly every year: a 'high-end' and a slightly more basic (and cheaper) model. However, Apple continues to manufacture and sell 'last year's model'. As a result, the latest models — at the time of writing, the iPhone 11 models — compete with earlier models which Apple still sells, such as the iPhone 6. In 2020, Apple also launched a cheaper and lower specification iPhone SE. Samsung, which is Apple's main rival in the smartphone market, differentiates its phones in a similar way, but launches its new phones more frequently and in a wider variety of options than its American competitor.

Product differentiation The production and marketing of generally similar products with minor variations.

The objectives of firms

Economists usually assume that firms have a single business objective: to maximise profit. We shall explain profit maximisation in the context of perfect competition in the next topic. However, real-world firms may have other objectives, such as to maximise sales revenue, maximise the growth of the business or maximise managerial objectives. The last is significant when there is a divorce of ownership from control in a business, which occurs in large firms organised as public limited companies (plcs). Plcs are owned by thousands of shareholders who employ managers or executives to run the business. As business scandals in the early twenty-first century in firms such as Royal Bank of Scotland (RBS) have shown, the managers may pursue their own agendas, maximising their personal pay and making decisions that are not in the interests of the business's owners.

Firms may also be profit 'satisficers' rather than profit 'maximisers'. Under this assumption, decision makers in firms, be they the owners of small corner shops or the chief executives of huge plcs, may be content with a satisfactory outcome, say satisfactory profit, rather than the best possible outcome. They may be happy with an easy life.

Perfect competition

Perfect competition short-run equilibrium

In the previous topic on production, costs and revenue, we explained how a perfectly competitive firm's revenue curves are derived from the assumptions that the firm can sell whatever quantity it wishes at the ruling market price, but that it cannot influence the ruling market price by its own action.

Figure 18 shows the equilibrium level of output produced by a perfectly competitive firm in the short run. The firm (shown in the left-hand panel of the diagram) has to accept the ruling price determined by market supply and demand (shown in the right-hand panel). As explained in the previous topic, the ruling price is also the firm's average revenue (AR) curve and its marginal revenue (MR) curve.

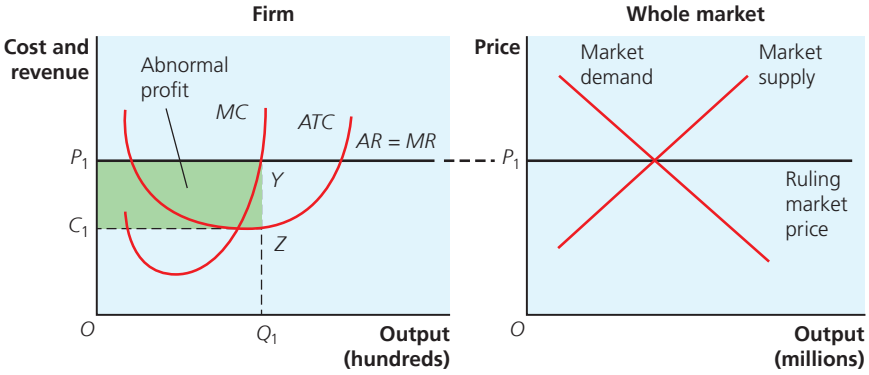


Figure 18 The short-run equilibrium level of output of a perfectly competitive firm

At this point in the analysis, we further explain the condition that must be met for a firm in any market structure — including perfect competition — to maximise profit. The firm must produce the level of output at which the marginal revenue it earns exactly equals the marginal cost incurred when producing this level of output: $MR = MC$. If the firm produces below this level of output (in which case $MR > MC$), then, by stepping up output, profit increases. Conversely, if the firm produces beyond the profit-maximising level of output (in which case $MR < MC$), then, by cutting back output, profit increases.

Using the $MR = MC$ condition, the firm's profit-maximising or equilibrium output is Q_1. At Q_1, total sales revenue (quantity sold times price) is shown by the area OP_1YQ_1. Likewise, total cost (quantity sold times average cost) is shown by the area OC_1ZQ_1. This means that the shaded area C_1P_1YZ shows abnormal profit (total revenue minus total cost). Abnormal profit can, of course, be made at levels of output other than Q_1 — indeed, at all levels of output at which price is above average cost. But at these levels of output profit is less than at Q_1. Only by producing and selling Q_1 can the firm make the largest possible abnormal profit.

Perfect competition long-run equilibrium

The short-run equilibrium shown in Figure 18 is a temporary equilibrium rather than a true equilibrium. In the short run, new firms cannot enter the market, so incumbent

Knowledge check 19

List three industries or markets that approximate to perfect competition (i.e. industries or markets that exhibit most but not all of the characteristics of perfect competition).

firms (i.e. firms already in the market) continue to make abnormal profit. However, in the long run, when there are no **entry barriers** or **exit barriers** and firms can enter or leave the market freely, abnormal profit (shown by the shaded area in the left-hand panel of Figure 18) acts as a magnet, attracting new firms into the market. The entry of new firms shifts the market supply curve rightwards from S_1 to S_2 in the right-hand panel in Figure 19. This causes the ruling market price to fall until it settles at P_2. Market and firm are now both in long-run or true equilibrium.

Entry barriers High start-up costs or other obstacles that prevent new competitors from easily entering a market or industry.

Exit barriers Obstacles that stop or prevent the exit of a firm from a market or industry.

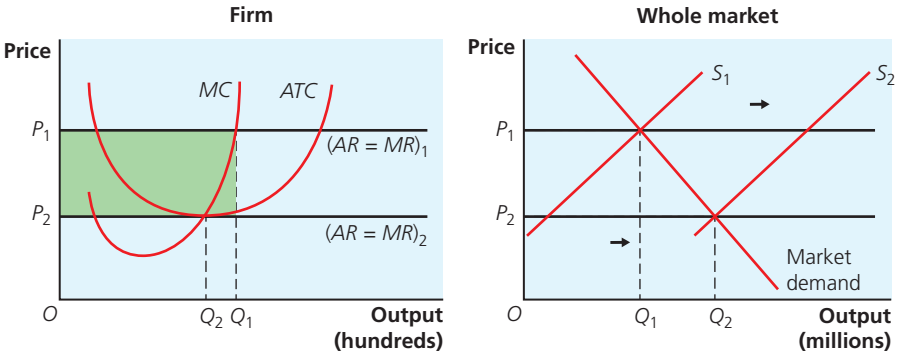


Figure 19 How long-run equilibrium is achieved in perfect competition

Figure 20 shows more clearly a perfectly competitive firm in long-run equilibrium. The price line just touches the lowest point of the firm's *ATC* curve, so no abnormal profit is made. Because the profit made by surviving firms is restricted to normal profit, the incentive for new firms to enter the market no longer exists.

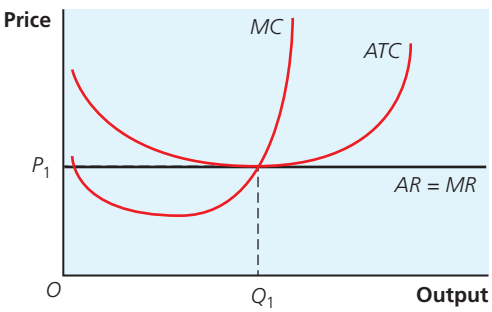


Figure 20 A perfectly competitive firm in long-run equilibrium

Knowledge check 20

Why is it impossible to find real-world examples of perfect competition?

Perfect competition and economic efficiency

Given certain assumptions, perfect competition results in an efficient allocation of resources. In the long run, perfect competition is productively efficient, minimising average costs and allocatively efficient ($P = MC$). However, this conclusion only holds if we pretend there are no economies of scale and no externalities. It is a bit like assuming away the real world in which we live.

Exam tip

Remember that perfect competition is a completely unrealistic market structure.

Monopoly and monopoly power

Before we explain how a monopoly maximises profit, we must first explain a monopolist's average revenue (*AR*) and marginal revenue (*MR*) curves. Monopoly revenue curves differ from those facing a firm in a perfectly competitive market. Because there is only one firm in the market, the **market demand curve** is the demand curve for the monopolist's output. This means that the monopolist faces a downward-sloping demand curve, which can affect the monopolist in one of two ways:

- If we regard the monopolist as a **price-maker**, then whenever it sets the price, the demand curve determines how much it can sell. If the monopolist tries to raise the price, it must accept a fall in sales.
- Alternatively, if the monopolist decides to act as a quantity setter, the demand curve dictates the maximum price at which any chosen quantity can be sold. Thus, the downward-sloping demand curve means that the monopolist faces a trade-off.

A monopoly cannot set price and quantity independently of each other.

Because the demand curve shows the price that the monopolist charges at each level of output, the demand curve is the monopolist's average revenue curve. Unlike perfect competition, however, marginal revenue and average revenue in monopoly are not the same. Because the average revenue curve falls, the marginal revenue curve must be below it. This is illustrated in the left-hand panel of Figure 21. Note that the *MR* curve is twice as steep as the *AR* curve. This is always the case whenever the *AR* curve is both downward sloping and a straight line.

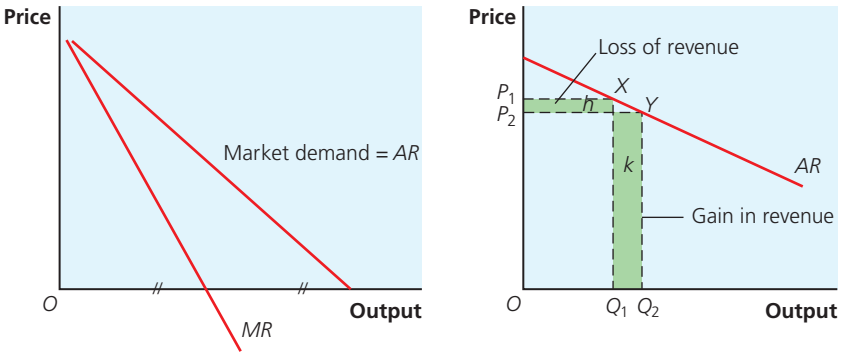


Figure 21 Monopoly average revenue and marginal revenue curves

The relationship between *AR* and *MR* in monopoly can also be explained in another way. This is illustrated in the right-hand panel of Figure 21. Because the demand curve (*AR* curve) is downward sloping, the monopolist can sell an extra unit of output only by reducing the price (and average revenue) of all units of output sold. In this situation, total sales revenue increases by the shaded area *k* on the diagram but decreases by the shaded area *h*. Area *k* shows a revenue gain, namely the extra unit sold multiplied by its price. By contrast, area *h* shows a revenue loss. The revenue loss results from the fact that, in order to sell one more unit of output, the price has to be reduced for all units of output, not just the extra unit sold. In monopoly, marginal revenue = the revenue gain *minus* the revenue loss, which must be less than price or average revenue.

Market demand curve Shows how much of a good or service all the consumers in the market plan to demand at different possible prices.

Price-maker A monopoly possessing sufficient market power to choose the price at which it sells its output.

Monopoly equilibrium

Just like a perfectly competitive firm, a monopoly maximises profit by producing the level of output at which $MR = MC$. In Figure 22, point A locates the profit-maximising level of output (Q_1). However, the price charged by the monopoly is located at point D on the demand curve (and AR curve), immediately above point A. Abnormal profit is shown by the shaded area C_1P_1DB. Unlike in perfect competition, Figure 22 does not distinguish between short-run and long-run equilibrium. This is because, in monopoly, entry barriers prevent new firms joining the market, thus enabling the monopoly to make abnormal profit in the long run as well as the short run.

In contrast to perfect competition, where abnormal profit is temporary, a monopoly makes abnormal profit as long as entry barriers protect it. Indeed, in monopoly, abnormal profit is often called monopoly profit, indicating the monopolist's power to preserve profit by preventing competition.

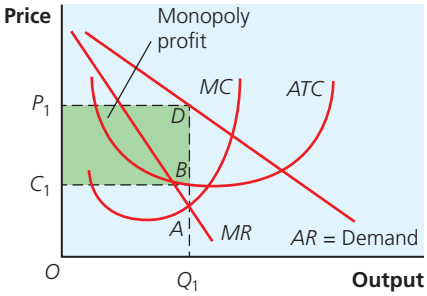


Figure 22 Monopoly equilibrium

Causes of monopoly and sources of monopoly power

Monopoly power stems from a firm's ability to exclude rivals from the market by imposing entry barriers. A pure monopoly obviously possesses monopoly power, but firms in imperfectly competitive markets such as oligopolies can also exercise monopoly power to a greater or lesser extent. Whereas perfect competition is characterised by **consumer sovereignty** (in the sense that firms respond to the wishes of consumers exercised through their pounds spent in the market), monopolies exercise and exploit **producer sovereignty**. Consumers cannot go elsewhere to buy the good and are presented with a 'take it or leave it' choice. Enjoying producer sovereignty, a firm with monopoly power exploits consumers by restricting output and raising price, by limiting consumer choice and by making permanent excess profit.

But even when a firm is a monopoly producer of a particular good or service, monopoly power is weak if close substitutes exist, produced by other firms in other industries. Monopoly power is greatest when the firm produces an essential good for which there are no substitutes. Factors that give rise to monopoly power include: advantages of geographical location; control over raw material supply or market outlets; economies of scale; use of advertising, branding and product differentiation as entry barriers; and laws such as patent legislation, which protect innovations and intellectual property from being copied.

Monopoly power The US Supreme Court defines monopoly power as 'the power to control prices or exclude competition'.

Consumer sovereignty The power of consumers to determine what goods and services are produced.

Producer sovereignty When firms have the power and ability to influence consumer decisions.

Advantages and disadvantages of monopoly

The possible advantages of monopoly result from two sources: economies of scale and **dynamic efficiency**. When substantial economies of scale are possible in an industry, monopoly may lead to a better outcome than competition. A further advantage of monopoly is that, compared to perfect competition, a monopoly can use its abnormal profit to fund research and development (R&D), which then leads to better ways of making existing products and to the development of completely new products.

Dynamic efficiency
Improvements in productive and allocative efficiency taking place over time.

The main disadvantage of monopoly is that it may lead to productive inefficiency and allocative inefficiency. The latter results in resource misallocation.

Monopolistic competition

Monopolistic competition, which, as we have mentioned, is often called 'imperfect competition among the many', resembles perfect competition in the following ways:

- As in perfect competition, there is a large number of firms in the market.
- In the long run there are no barriers to entry or exit.
- As a result, the entry of new firms, attracted by short-run abnormal profits, brings down the price each firm can charge until only normal profits are made in the long run.

Exam tip

Beware of confusing monopolistic competition with monopoly.

However, monopolistic competition resembles monopoly in two other ways:

- Each firm faces a downward-sloping demand curve. This results from the fact that each firm produces a slightly different product — differentiated by such features of modern production and marketing as style, design, packaging, branding and advertising. The goods produced by the various firms provide partial but not perfect substitutes for each other. The resulting product differentiation in the market means that each firm possesses a degree of monopoly power over its product. Unlike in perfect competition, if a firm raises its price, it does not lose all its customers because there is brand loyalty.
- Each firm's marginal revenue (*MR*) curve is below its average revenue (*AR*) curve, which of course is the demand curve for the firm's output.

Short-run profit maximisation in monopolistic competition

Short-run profit maximisation in monopolistic competition is illustrated in Figure 23, which is very similar to profit maximisation in monopoly, illustrated in Figure 22. However, in monopolistic competition the demand or average revenue curve represents demand for the goods produced by just one firm within the market rather than demand for the output of the whole market. And because the other firms within the market produce partial though not perfect substitutes, the demand curve facing the firm is likely to be rather more elastic at the prices each firm may decide to set than would be the case in pure monopoly.

The profit-maximising level of output, Q_1, is located below point A in Figure 23, where $MR = MC$, and the abnormal profits made by the firm in the short run are shown by the rectangular area C_1P_1BD.

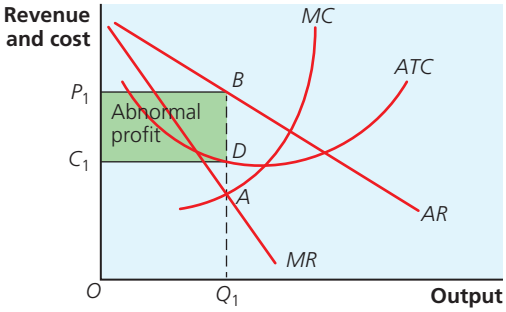


Figure 23 Short-run profit maximisation in monopolistic competition

Long-run profit maximisation in monopolistic competition

The absence of barriers to entry or exit in the long run is of great importance in the theory of monopolistic competition. Long-run profit maximisation in monopolistic competition is different in an important respect from profit maximisation in monopoly. In the long run, the entry of new firms causes the demand curve or AR curve facing an established firm to shift leftwards or inwards. The leftward shift may result from the introduction of new substitute products, attracting some customers away from the existing firms. Long-run profit maximisation is achieved when the AR curve forms a tangent to the firm's ATC curve, thereby removing the firm's abnormal profit.

This is shown in Figure 24, at point B immediately above level of output Q_1. Since only normal profit is made, total sales revenue and total costs of production are both shown by the rectangle OP_1BQ_1. Note also that the point of tangency between the AR and ATC curves occurs immediately above point A, which determines the profit-maximising level of output Q_1 where $MR = MC$.

Exam tip

This is a difficult diagram to draw, so study it carefully and practise drawing it.

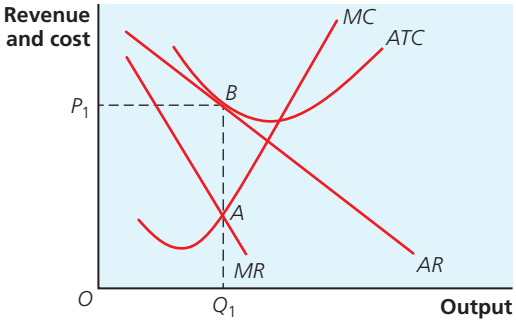


Figure 24 Long-run profit maximisation in monopolistic competition

Monopolistically competitive markets will be subject to non-price competition, which we explain in the next section on oligopoly.

Oligopoly

Oligopoly is a market structure in which a few large firms dominate the market. This means there is a high degree of market concentration, which can be measured by a **concentration ratio**. For example, a five-firm concentration ratio of 80% means that the five largest firms produce 80% of market output.

However, oligopoly is best defined by the behaviour or **market conduct** of the firms within the market, rather than by market structure. Oligopolists are interdependent rather than independent, in the sense that they need to take account of the likely reactions of their rivals, the other oligopolists, when making price and output decisions.

Consider, for example, an oligopolist who is thinking of raising the price charged in order to increase profit. Whether the price rise succeeds in increasing profit depends on the likely reactions of the other firms. Will rival firms follow suit and match the price rise, or will they hold their prices steady, hoping to gain sales at the expense of the firm that raised the price? Clearly, when deciding whether to raise or lower its price, an oligopolist must make assumptions about the likely response of the other firms.

Concentration ratio A measure of the total output produced in an industry by a given number of firms in the industry.

Market conduct The way that firms behave in a particular market structure: for example, whether they compete, collude, set prices and levels of output, innovate or indulge in anti-competitive restrictive practices.

Competitive oligopoly

As noted, in competitive oligopoly a firm has to take account of the reactions of its rivals when forming its market strategy, but it does so without cooperating or colluding with the other firms. Uncertainty is a characteristic of competitive oligopoly — a firm can never be completely certain as to how rivals will react to its marketing strategy. Will they or will they not follow suit?

Knowledge check 21

What is the size of the industry concentration ratio for a pure monopoly?

Collusive oligopoly

Uncertainty can be reduced by the rivals colluding to fix prices or output, or even by allocating customers to particular members of the oligopoly. For example, by forming a **cartel** agreement or price ring, oligopolists can achieve a better outcome for them all, in terms of joint profit maximisation and an easier life, than by remaining a competitive oligopoly.

However, collusion may not be good for the consumer, resulting in the disadvantages of monopoly, such as high prices and restriction of choice, without any of the benefits, such as economies of scale. For this reason, collusive oligopolistic arrangements, such as cartel agreements, are normally illegal, regarded by governments as against the public interest. In any case, it is seldom possible to eliminate uncertainty completely. Members of a cartel may cheat or renege on an agreement, secretly selling extra output at a price that undercuts the cartel's agreed price.

Cooperation, as distinct from collusion, may be good for consumers as well as the oligopolists. Cooperation can improve health and safety within the industry and ensure that product and labour standards are maintained.

Cartel A group of colluding firms that fixes prices (forming a price ring) and perhaps also divides up the market by allocating maximum levels of output for each firm.

Knowledge check 22

Why may competitive oligopoly give rise to collusive oligopoly?

The kinked demand curve theory of competitive oligopoly

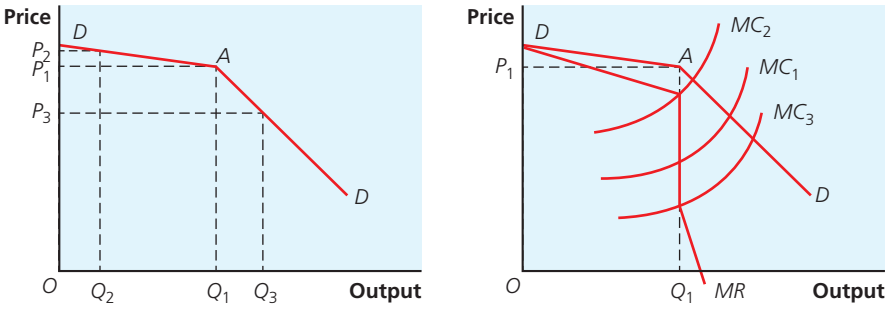


Figure 25 The kinked demand curve theory

The kinked demand curve theory, which is illustrated in Figure 25, can be used to explain a number of features of competitive oligopoly, such as interdependence, uncertainty and a preference for avoiding price wars. The theory was originally developed to explain price rigidity and the absence of price wars in oligopolistic markets. Suppose an oligopolist sells output Q_1 at price P_1 as shown in the left-hand panel of Figure 25. Because oligopolists lack accurate information about the demand and revenue curves they face, particularly at outputs other than those they are currently producing, each firm has to guess what will happen to demand if it changes its price.

The demand curve *DD* in Figure 25 represents an oligopolist's estimate of how demand will change with respect to either a price rise or a price fall. *DD* has been drawn on the assumption that the firm expects demand for its product to be relatively elastic in response to a price rise because rivals are expected to react by keeping their prices stable in the hope of gaining profits and market share. But the oligopolist expects demand to be relatively inelastic when the price is cut. This is because the oligopolist expects rivals to react to a price cut by decreasing their prices by the same amount. Few, if any, customers are likely to be lured away from rival firms.

The oligopolist therefore expects rival firms to react asymmetrically when price is raised compared with when price is lowered. The oligopolist's initial price and output, P_1 and Q_1, are located at the junction of two demand curves with different elasticities, each curve reflecting a different assumption about how rivals are expected to react to a change in price. The oligopolist expects profit to be lost whether price is raised or cut. On these assumptions, the best policy is to leave price unchanged.

Developing the kinked demand curve theory

The right-hand panel of Figure 25 illustrates a way in which the kinked demand curve theory can be developed further. As in all market structures, the demand curve facing a firm is also its average revenue (*AR*) curve. But as we saw for monopoly, when *AR* falls, marginal revenue (*MR*) is below *AR*. You should note that the *MR* curve in the right-hand panel of Figure 25 has three sections.

- The uppermost section relates to the more elastic section of the *AR* curve to the left of the kink below point *A*.

- The lowermost section relates to the less elastic section of the *AR* curve below and to the right of the kink.
- The mid-section of the *MR* curve is the vertical line joining the upper and lower sections of the *MR* curve below point *A* at the output level Q_1.

Suppose the *MC* curve is initially MC_1. Since $MR = MC$ at this level of output, P_1 must be the profit-maximising price. However, if marginal cost rises or falls between MC_2 and MC_3, the profit-maximising output and price continue respectively to be Q_1 and P_1. The oligopolist's selling price remains stable despite quite significant changes in costs of production.

Weaknesses of the kinked demand theory

Although at first sight attractive as an explanation of price stability in conditions of oligopoly, the kinked demand theory has two significant weaknesses. First, it is an incomplete theory, since it does not explain how and why a firm chooses to be at point *A* in the first place. Second, evidence provided by the pricing decisions of real-world firms gives little support to the theory. Rival firms seldom respond to price changes in the manner assumed in the kinked demand curve theory, and it is also reasonable to expect that an oligopolist will test the market, that is, raise or lower the selling price to see whether rivals react in the manner expected. If the rivals do not, then the oligopolist will surely revise its estimate of demand for its product. Evidence conclusively shows that oligopoly prices tend to be stable or sticky when demand conditions change in a predictable or cyclical way, and that oligopolists usually raise or lower prices quickly and by significant amounts, both when production costs change substantially and when unexpected shifts in demand occur.

Non-price competition

The kinked demand curve theory suggests that oligopolists are reluctant to use **price competition** to gain sales and market share, although there is plenty of evidence that oligopolists do on occasion engage in **price wars**, even though, according to the kinked demand curve theory, such wars are self-defeating. Nevertheless, oligopolists also engage in many forms of **non-price competition**, such as marketing competition (for example, obtaining exclusive outlets such as tied public houses and petrol stations through which breweries and oil companies can sell their products), the use of persuasive advertising, product differentiation, brand imaging and packaging, and quality competition, including the provision of after-sales service.

The advantages and disadvantages of oligopoly

These are much the same as the advantages and disadvantages of monopoly, which we covered earlier. See page 44.

Price discrimination

Oligopolists (and monopolists) sometimes use price discrimination to increase their profits. Price discrimination occurs when firms charge different prices to different customers based on differences in the customers' ability and willingness to pay. Those customers who are prepared to pay more are charged a higher price than those who are only willing to pay a lower price. It is important to understand that discriminatory

Exam tip

For the most part, examination questions on oligopoly can be fully answered using only the 'simple' version of the kinked demand theory, which is illustrated in the left-hand panel of Figure 25. Use of the more complicated right-hand panel often leads to students including errors in their diagrams and perhaps also straying into irrelevance.

Price competition When firms compete on the basis of low prices and value for money.

Price wars When rival firms compete by reducing prices. A price cut by one firm provokes retaliation by its competitors.

Non-price competition When firms compete on the basis of quality, advertising, after-sales service and other non-price factors.

prices are for the most part based on differences in demand conditions rather than on differences in costs of production or on differences in the quality of the good or service.

Knowledge check 23

Can all the differences in the prices charged by oligopolists be explained by the theory of price discrimination?

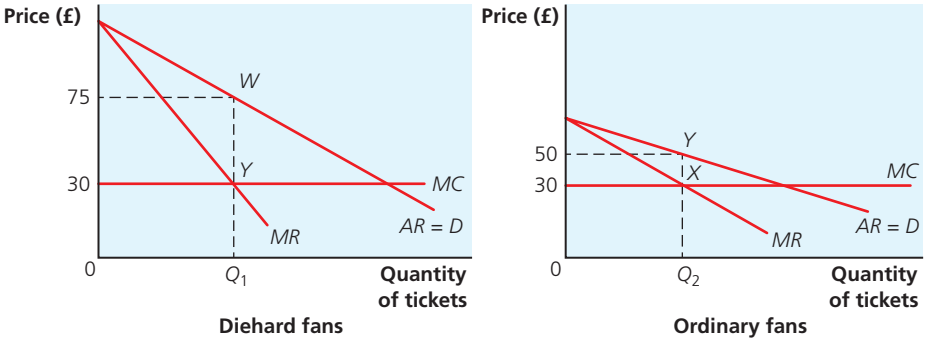


Figure 26 Price discrimination at a rock concert

The theory of price discrimination explains why different prices are charged for different tickets at events such as rock concerts. We shall assume that 'diehard' fans of the rock group playing at the concert are prepared to pay £75 for a ticket, whereas 'ordinary' fans are only prepared to pay £50. This means that the two different groups of fans have different price elasticities of demand and different demand curves, which are shown in Figure 26.

At all the prices that could be charged for seeing the rock concert, diehard demand is more inelastic than ordinary demand, indicating that the latter fans are less enthusiastic about the group performing at the rock concert. For both groups of fans, the downward-sloping demand curves in Figure 26 show average revenue (*AR*), but not marginal revenue (*MR*). In each case, the *MR* curve is twice as steep as the *AR* curve. The diagrams also assume (perhaps unrealistically) that the marginal cost (*MC*) of selling an extra ticket is always the same, namely £30. This explains the horizontal *MC* curve in both panels of the diagram.

To maximise profit, *MR* must equal *MC* in both sub-markets. As the diagrams show, this means that diehard fans pay the higher price of £75 for admission, with ordinary fans paying the lower entry price of £50. With the different prices being charged, Q_1 diehard fans and Q_2 ordinary fans watch the concert. The different prices charged result from the different price elasticities of demand. Profit is maximised when more price-sensitive ordinary fans pay less than the less price-sensitive diehard fans.

The dynamics of competition and competitive market processes

The traditional view in economics has been that price competition is the main form of competition in markets in which there is a large number of firms, but that as markets become less competitive, at least in the form of price competition, large firms become more dominant and market concentration increases.

However, many economists now argue that all market structures in a capitalist economy can be highly competitive, although business decisions which on first sight

may appear to be competitive are in fact 'anti-competitive' in the sense that they aim to increase the market power of already dominant firms. This view of competition is closely linked to the process of creative destruction, defined on page 37, an idea developed by the Austrian–American economist Joseph Schumpeter to explain the dynamic process through which capitalist economies change over time. Indeed, creative destruction has become the centrepiece for modern thinking on how economies evolve. Back in 1942, Schumpeter wrote:

> **The opening up of new markets, foreign or domestic, incessantly revolutionises the economic structure from within, incessantly destroying the old one, incessantly creating a new one. This process of creative destruction is the essential fact about capitalism.**

Contestable and non-contestable markets

In **contestable market** theory, monopoly power is not dependent on the number of firms in the market or on concentration ratios, but rather on the ease or difficulty with which new firms may enter the market. Industrial concentration is not a problem, provided that an absence of barriers to entry and exit creates the ability for new firms to enter and contest the market. Actual competition in a market is not essential. According to this view, the threat of entry by new firms or potential competition is quite enough to ensure efficient and non-exploitative behaviour by existing firms within the market.

In recent years, contestable market theory has had a major impact on UK monopoly policy. Instead of interfering with firms' pricing and output policies, governments should restrict the role of monopoly policy to discovering which industries and markets are potentially contestable. Deregulatory policies should be used to develop conditions in which barriers to entry and exit are minimised. In this way, non-contestable markets can be made contestable, though the existence of **sunk costs** may inhibit this process.

Market structure, static efficiency, dynamic efficiency and resource allocation

The specification requires you to be able to apply efficiency concepts when comparing the performance of firms in different market structures. In addition to understanding market structures, you must understand how **conduct indicators** and **performance indicators** can be used to compare market outcomes.

The left-hand and right-hand panels of Figure 27 respectively show a perfectly competitive firm and a monopoly in equilibrium — assuming that firms in both markets have similar *ATC* curves, which means there are no economies of scale. The figure shows that the perfectly competitive firm is productively efficient (producing where *ATC* is lowest), but that the monopoly is productively inefficient (producing above minimum *ATC*). Likewise, the perfectly competitive firm is allocatively

Contestable market A market in which the potential exists for new firms to enter the market. A perfectly contestable market has no entry or exit barriers and no sunk costs, and both incumbent firms and new entrants have access to the same level of technology.

Sunk costs Costs that have already been incurred and cannot be recovered.

Conduct indicators Information about a firm's conduct, e.g. its price-setting behaviour, which helps to indicate whether the firm's actions are competitive or anti-competitive.

Performance indicators Information about a firm's performance, e.g. its profit margins and investment record, which helps to indicate the extent to which the firm is protected from competition.

efficient (as $P = MC$), whereas the monopoly is allocatively inefficient (since $P > MC$). Compared with perfect competition, the monopoly's price is too high and its output is too low.

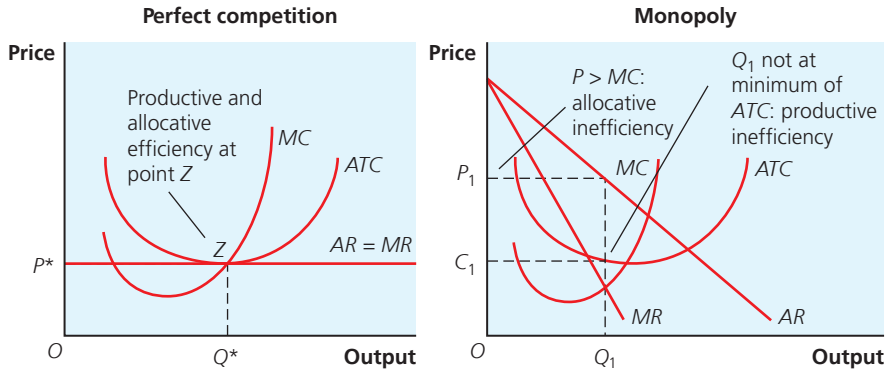


Figure 27 Evaluating perfect competition and monopoly in terms of economic efficiency

But is perfect competition necessarily more efficient than monopoly?

The conclusion made in the previous paragraphs that perfect competition is productively more efficient than monopoly depends partly on an assumption that there are no economies of scale. When substantial economies of scale are possible in an industry, monopoly may be productively more efficient than competition.

Figure 28 illustrates a **natural monopoly** where, because of limited market size, there is no room in the market for more than one firm benefiting from full economies of scale. Producing on the short-run average cost curve $SRATC_N$, the monopoly may be producing above the lowest point on this particular cost curve, hence exhibiting a degree of productive inefficiency. However, all points drawn on $SRATC_N$ incur lower unit costs — and are productively more efficient — than any point on $SRATC_1$, which is the relevant cost curve for each firm were the monopoly to be broken up and transformed into a number of smaller competitive enterprises.

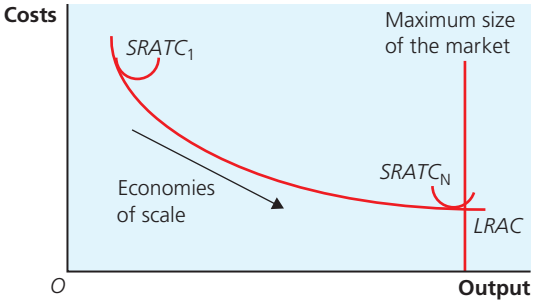


Figure 28 The justification of monopoly when economies of scale are possible

So far we have considered only **static efficiency** but not dynamic efficiency. As soon as we introduce dynamic efficiency into the analysis, monopolies can be justified on the grounds that overall they may be more efficient than perfectly competitive firms.

Natural monopoly The term has two meanings: first, when a country or firm has complete control of a natural resource; second, when there is room in a market for only one firm benefiting to the full from economies of scale.

Static efficiency Productive and allocative efficiency at a particular point in time.

Protected by entry barriers, a monopoly can earn monopoly profit without facing the threat that the profit will be whittled away as new firms enter the market. This allows an innovating monopoly to enjoy the fruits of successful R&D and product development in the form of monopoly profit. In perfect competition, by contrast, there is little or no incentive to innovate because other firms can 'free-ride' and gain costless access to the results of any successful research. This argument is used to justify patent legislation, which grants a firm the right to exploit the monopoly position created by innovation for a number of years before the patent expires.

However, there is a counter-argument, namely that monopoly reduces rather than promotes innovation and dynamic efficiency. Protected from competitive pressures, a monopoly may profit satisfice rather than profit maximise, content with satisfactory profit and an easy life.

Consumer and producer surplus

To explain how market structures affect economic welfare, it is first necessary to introduce the concepts of consumer surplus and producer surplus as measures of welfare. **Consumer surplus** and **producer surplus** are illustrated in Figure 29.

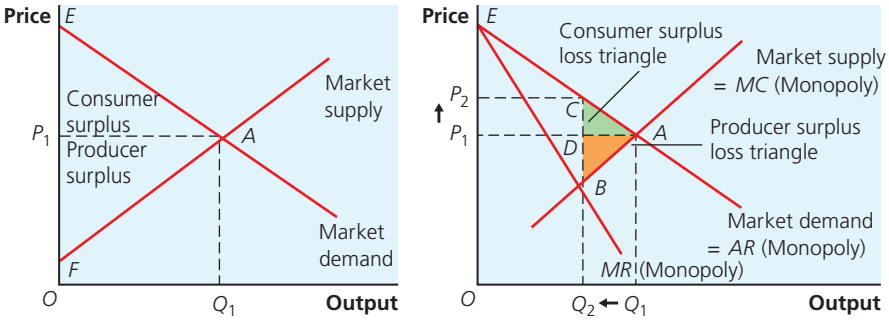


Figure 29 Market structure and economic welfare

Consumer surplus is the difference between the maximum price a consumer is prepared to pay and the actual price they need to pay. In a competitive market such as the left-hand side of Figure 29, the total consumer surplus enjoyed by all the consumers in the market is measured by the triangular area P_1EA. Consumer welfare increases whenever consumer surplus increases — for example, when market prices fall. Conversely, however, higher prices reduce consumer surplus and welfare.

Producer surplus, which is a measure of producers' welfare, is the difference between the minimum price a firm is prepared to charge for a good and the actual price charged. In the left-hand side of Figure 29, the producer surplus enjoyed by all the firms in the market is measured by the triangular area FP_1A.

What may happen to economic welfare when monopoly replaces perfect competition?

The right-hand side of Figure 29 illustrates what can happen to economic welfare when monopoly replaces perfect competition (assuming there are no economies of scale). Market equilibrium in perfect competition is determined at point A: output

Consumer surplus
A measure of consumer satisfaction or welfare, being the difference between what consumers are willing to pay for a good or service and the market price they actually need to pay.

Producer surplus A measure of producer welfare, being the difference between what producers are willing to accept in payment for a good or service and the market price they actually receive.

is Q_1 and price is P_1. Monopoly equilibrium, by contrast, is determined at point B, where $MR = MC$. (Note that the marginal cost curve in monopoly is the same curve as market supply in perfect competition.) The figure illustrates the standard case against monopoly, namely that compared with perfect competition, monopoly restricts output (to Q_2) and raises price (to P_2).

But we can now take the analysis one stage further and investigate how consumer surplus and producer surplus (and hence economic welfare) are affected. Raising the price from P_1 to P_2 transfers consumer surplus away from consumers and to the monopoly. The transfer is shown by the rectangle bounded by the points P_1, P_2, C and D. Producer surplus (in the form of monopoly profit) increases at the expense of consumer surplus. Over and above this transfer, there is also a net loss of economic welfare caused by the fact that the amount bought and sold falls to Q_2. The welfare loss (**deadweight loss**) is shown by the two shaded triangular areas in the right-hand panel of Figure 29. The upper triangle shows a loss of consumer surplus and the lower triangle shows a similar loss of producer surplus.

Deadweight loss The cost to society created by market inefficiency or the inefficient allocation of resources.

Knowledge check 24

How may monopoly lead to market failure?

Summary

- Perfect competition and monopoly are the extreme forms of market structure separated by imperfect competition.
- In the short run, perfectly competitive firms can make abnormal profit, but in the long run, market forces eliminate abnormal profit.
- True or long-run equilibrium in perfect competition occurs when surviving firms make normal profit only.
- Because a monopoly is protected by barriers to market entry, it can make abnormal profit (monopoly profit) in the long run.
- Monopolistic competition must not be confused with monopoly.
- Oligopoly is a market structure dominated by a few firms.
- In terms of market conduct or behaviour, the extent to which a market is oligopolistic can be judged by the extent to which the firms are interdependent.
- Price discrimination involves charging prices on the basis of how much consumers are prepared to pay.
- A high degree of interdependence leads to uncertainty about how rival oligopolists will react to a firm's market strategy, which in turn creates an incentive to collude and to form cartels.
- The theory of the kinked demand curve can be used to model oligopoly behaviour.
- If there is no scope for economies of scale, in the long run perfect competition is productively and allocatively efficient; by contrast, monopoly is productively and allocatively inefficient.
- In the long run, however, economies of scale may result in monopoly being more productively efficient than perfect competition. Likewise, a monopoly may be more dynamically efficient.
- In terms of economic welfare, when monopoly replaces perfect competition, the monopoly gains at the expense of consumers (i.e. consumer surplus is replaced by producer surplus).

The labour market

Labour market theory is really just the price theory that you have studied in the goods market or product market of the economy, but operating in the factor market, which is the market for the services of factors of production. Households and firms function simultaneously in both sets of markets, but their roles are reversed. In the labour market, which is part of the factor market, firms demand labour services that households supply. As with the goods market or product market, we shall start by explaining the demand for labour, then switch to the supply of labour, before bringing demand and supply together in the context of a perfectly competitive labour market.

The demand for labour: marginal productivity theory

Firms demand labour because they believe profit can be made by selling the goods produced by their workers. This means that the demand for labour is a **derived demand**. Just as the market supply curve of labour in a perfectly competitive labour market is the sum of the supply curves of the individual workers in the labour market, so the market demand curve for labour is the sum of the demand curves for labour of each firm in the market.

Marginal revenue product and marginal physical product

Each firm's demand curve is the **marginal revenue product (MRP) of labour** curve facing the firm in the labour market. The marginal revenue product curve shown in panel (c) of Figure 30 is obtained by multiplying the **marginal physical product (MPP) of labour**, shown in panel (a), by marginal revenue (*MR*), shown in panel (b). The *MPP* of labour is just another name for the marginal returns (marginal product) of labour, which you first came across in Production, costs and revenue on page 28. Because of the law of diminishing returns, the marginal product of labour falls as additional workers are hired. As its name indicates, the *MPP* curve shows only the physical output produced by one extra worker — measured in whatever goods the firm produces. To convert this into a money value, the *MPP* of labour must be multiplied by marginal revenue. The end result is the *MRP* curve:

marginal physical product × marginal revenue = marginal revenue product

Figure 30 Deriving a firm's demand curve for labour (the MRP curve) from the MPP curve

Derived demand Demand for a good or service derived from the fact that the good or service is needed in the production of another good or service. An increase in the demand for the latter causes the demand for the input to increase.

Knowledge check 25

Test your synoptic knowledge by explaining the difference between derived demand and composite demand.

Marginal revenue product (MRP) of labour The monetary value of the addition to a firm's total output brought about by employing one more worker.

Marginal physical product (MPP) of labour The addition to a firm's total output brought about by employing one more worker.

Shifts in the demand curve for labour

Factors that can shift the demand curve for labour include: a change in the quantity demanded of the product that the labour produces; and a change in the production process that uses more or less labour.

The elasticity of demand for labour

The determinants of elasticity of demand for labour are:

1. The proportion of labour costs in total costs: if labour costs form a large proportion of total costs, a change in wages will have a significant impact on costs and hence demand will be elastic.
2. The ease with which labour can be substituted by capital: if it is easy to replace workers with machines, demand will again be elastic.
3. The elasticity of demand for the product produced.

Knowledge check 26

Write out the formula for the elasticity of demand for labour with respect to a change in the wage rate.

Influences on the supply of labour to different markets

Monetary and non-monetary considerations

The supply of labour to a particular occupation is influenced by both monetary and non-monetary considerations. Workers obviously gain utility from the wages they earn. However, they also gain utility from the job satisfaction they enjoy (or suffer a utility loss from any job dissatisfaction). Different types of work yield different amounts of job satisfaction and dissatisfaction. When a worker enjoys the job, the **net advantage** of work is greater than the welfare yielded by the wage. In this situation, the worker is willing to work for a money wage lower than the wage that would be acceptable if there were no satisfaction from the work itself. But for some workers, work such as routine assembly-line work in factories and heavy manual labour is unpleasant, yielding job dissatisfaction. The supply of labour for this type of employment reflects the fact that the hourly wage rate must be high enough to compensate for the unpleasantness (or sometimes the danger) of the job.

Net advantage The sum of the monetary and non-monetary benefits of working,

The market supply curve of labour

The market supply curve of labour is obtained by adding together the individual supply curves of all the workers in the market. A worker's decision to supply one more hour of labour time must also mean that they sacrifice an hour of leisure time. For the worker to supply more labour, the hourly wage rate must rise to compensate for the fact that, as more money is earned, an extra pound means less and less, but an extra hour of leisure time sacrificed means more and more. (In economic terminology, the marginal utility of money falls and the marginal utility of leisure time rises as the worker supplies more labour, which eats into leisure time.)

The resulting upward-sloping supply curve of labour is shown in Figure 31(a). At the going hourly wage rate, a worker will not wish to supply labour beyond the point at which marginal utility of the wage = the marginal utility of leisure, other things remaining equal. At this point, the wage received from the last hour

worked yields the same utility as the last hour of leisure time enjoyed. To make it worthwhile for a worker to supply labour beyond this point, the hourly wage rate must rise — for example, from W_1 to W_2 in Figure 31(a). It is possible, however, that the supply curve may bend backwards above a certain wage rate (W_2 in Figure 31b), showing that as the wage rate rises above a critical level, the worker chooses to work fewer hours.

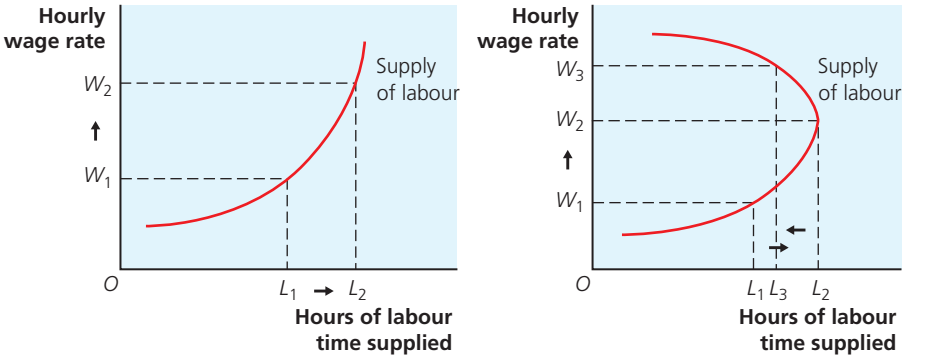


Figure 31 An individual worker's supply of labour

The elasticity of the supply of labour

The supply of labour to different labour markets will also be affected by the **elasticity of supply of labour**. Factors that determine the wage elasticity of supply of labour include the following:

- The supply of unskilled labour is usually more elastic than the supply of a particular type of skilled labour. This is because the training period of unskilled labour is usually very short, and any innate abilities required are unlikely to be restricted to a small proportion of the total population.
- Factors that reduce the occupational and geographical mobility of labour tend to reduce the elasticity of labour supply.
- The supply of labour is also likely to be more elastic in the long run than in the short run.
- The availability of a pool of unemployed labour increases the elasticity of supply of labour, while full employment has the opposite effect.

The determination of relative wage rates and levels of employment in perfectly competitive labour markets

A perfectly competitive labour market, if it were to exist, would have to meet all of the following conditions at the same time. As would be the case in a perfectly competitive goods market, a perfectly competitive labour market would have to contain a large number of buyers and sellers, each unable to influence the ruling

Exam tip

This is another application of the concept of marginal utility, which is explained in the second topic in this guide, Individual economic decision making.

Elasticity of supply of labour Proportionate change in the supply of labour following a change in the wage rate

Knowledge check 27

Write out the formula for the elasticity of supply of labour with respect to a change in the wage rate.

market price (in this case, the ruling market wage), and operating in conditions of perfect market information. Employers and workers would be free to enter the labour market in the long run, but an individual employer or firm could not influence the ruling market wage through its independent action.

Just as in a perfectly competitive goods market, it is impossible for all these requirements to be met simultaneously. It follows that perfectly competitive labour markets do not exist in the real world. Some labour markets, such as the market for building labourers in a city or region in which there is a large number of small building companies, approximate to perfect competition, but nevertheless they are not perfectly competitive.

However, assuming perfect competition does exist, we are now in a position to show the determination of the equilibrium wage rate and level of employment, both for a single firm or employer within the market and for the whole of the labour market. These are shown respectively in panel (a) and panel (b) of Figure 32. The equilibrium wage rate W^* and level of employment L^* are determined in panel (b), where market demand equals market supply. Panel (a) then shows each firm as a price-taker at wage rate W*, which, as well as being the perfectly elastic supply curve of labour facing each firm, is the **average cost of labour** (AC_L) curve and the **marginal cost of labour** (MC_L) curve. Because a firm can hire as many workers as it wants at W^*, every time an extra worker is hired, the firm's total wage bill rises by the wage paid to the new worker. Thus, MC_L equals the ruling wage, which is also the AC_L (wage cost per worker).

Average cost of labour Total wage costs divided by the number of workers employed.

Marginal cost of labour Addition to a firm's total cost of production resulting from employing one more worker.

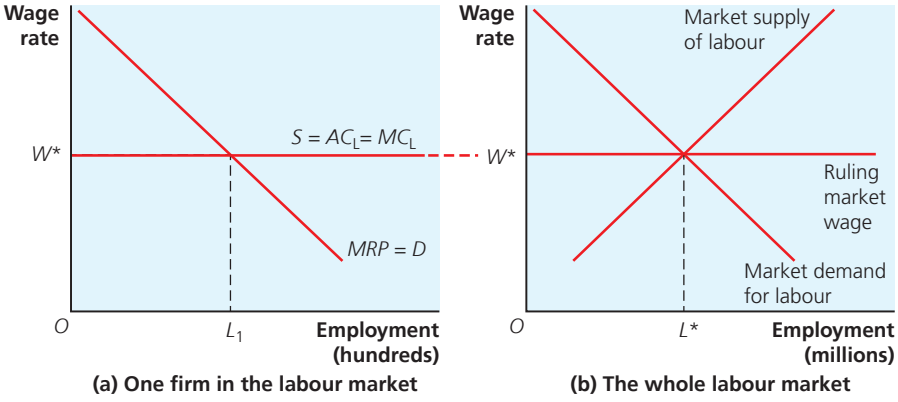


Figure 32 The equilibrium wage rate and level of employment in a perfectly competitive labour market

To maximise profit when eventually selling the output produced by labour, the firm must demand labour up to the point at which:

- the addition to sales revenue from employing an extra worker = the addition to production costs from employing an extra worker

or:

- $MRP = MC_L$

In a perfectly competitive labour market, MC_L always equals the ruling wage, so the firm hires labour up to the point at which the marginal revenue product of labour equals the wage rate ($MRP = W$). This *is* L_1 in Figure 32(a).

Exam tip

It is important to remember that perfect competition does not occur in the real world — either in goods markets for products or in labour markets. Perfect competition should be regarded as a theoretical abstraction against which real-world markets can be judged, particularly in relation to competitiveness, efficiency and economic welfare.

Knowledge check 28

Explain why firms and workers in a perfectly competitive labour market are passive price-takers.

The determination of relative wage rates and levels of employment in imperfectly competitive labour markets

Various factors such as **monopsony** power, trade unions and imperfect information contribute to imperfections in a labour market. A labour market in which there is a single employer is called a monopsony, and a market dominated by a single employer, but in which there are some other employers, is monopsonistic. A feature of such labour markets is imperfect market information among both employers and workers. An employer, for example, may enforce non-disclosure conditions of service on workers which prevent them from revealing their rates of pay to other employees.

Monopsony A single buyer in a market.

A monopsony is similar to a monopoly in many ways. As in monopoly, where consumers cannot choose between alternative suppliers of the good, in a monopsony labour market, workers cannot choose between alternative employers: only one firm or employer is available to hire their services. And in the same way that the market demand curve facing a monopoly supplier of a good is also the monopolist's average revenue curve, so the market supply curve of labour is the monopsonist's average cost of labour (AC_L) curve.

The AC_L curve shows the different wage rates that the monopsonist must pay to attract labour forces of different sizes. For example, Figure 33(a) shows a monopsony employer hiring ten workers at a daily wage or AC_L of £100 each. The figure shows that with ten workers initially employed, the wage (or AC_L) must rise from £100 to £110 a day to attract an eleventh worker.

Knowledge check 29

Explain why a monopsony employer is a price-maker rather than a passive price-taker.

But in a monopsony labour market, the AC_L curve is *not* the marginal cost of labour (MC_L) curve. To attract extra workers, the monopsonist must raise the daily wage rate, paying the higher rate (in the absence of wage discrimination, which we shall explain shortly) to *all* its workers. In this situation, the MC_L incurred by employing an extra worker includes the total amount by which the wage bill rises and not just the wage paid to the additional worker hired. The MC_L curve is thus *above* the AC_L or supply curve (whereas in a goods or product market, a monopolist's MR curve is *below* its AR curve). In Figure 33(a), the MC_L of employing the eleventh worker is £210 a day. This comprises the £110 paid to the eleventh worker (the green area in Figure 33a), plus the £10 extra now paid to each

of the original ten workers, which totals £100 (shown by the orange shaded area in Figure 33a).

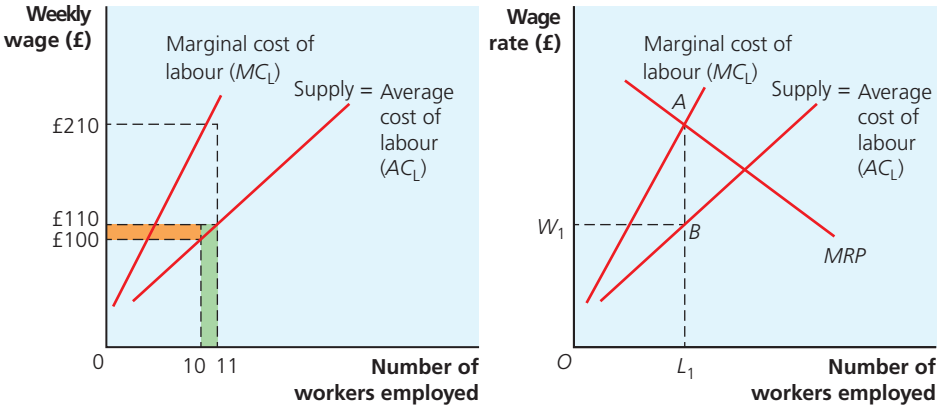


Figure 33 A monopsony labour market

Figure 33(b) shows the equilibrium wage and level of employment in a monopsony labour market. As in a perfectly competitive labour market, the firm's equilibrium level of employment is determined where $MRP = MC_L$. This is at point *A*. However, the equilibrium wage is *below* *A* and thus less than the *MRP* of labour, being determined at point *B* on the supply curve of labour. Although the monopsonist could pay a wage determined at *A* and equal to the *MRP* of labour, without incurring a loss on the last worker employed, it has no need to. The monopsonist can employ all the workers required by paying the wage W_1, determined at point *B*.

Exam tip

Exam questions on monopsony are usually quite difficult and the standard of students' answers is often quite low. Think carefully before you choose to answer a question on a monopsony labour market.

The influence of trade unions in determining wages and levels of employment

A **trade union** is an association of workers formed to protect and promote the interests of its members. A union's main function is to bargain with employers to improve wages and other conditions of work. Acting as a monopoly supplier of labour, a union may try to set the wage rate above the market-clearing wage rate, leaving employment to be determined by the amount of labour that employers hire at the wage set by the union.

Trade union An association of workers formed to protect and promote the interests of its members.

The effect of forming a trade union in a perfectly competitive labour market

Figure 34 shows the effect of a union setting the wage rate above the market-clearing rate in a (previously) perfectly competitive labour market. Employment falls from L_1 to L_2

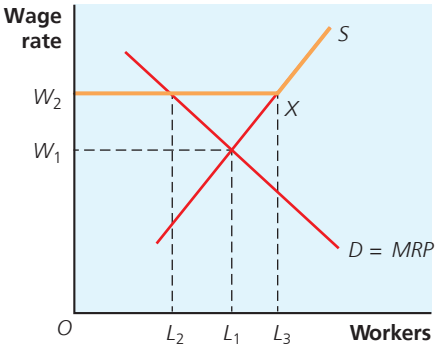


Figure 34 The effect of a union setting the wage rate above the market-clearing rate in a previously perfectly competitive labour market

The effect of forming a trade union in a monopsony labour market

In the monopsony labour market shown in Figure 35, a union may be able to raise both the wage rate and the employment level. In the absence of a union, the employment level is L_1, determined at point A where $MRP = MC_L$, and the wage rate is W_1, determined at point B. If the union sets the wage rate at W_2, the kinked line W_2XS becomes the labour supply curve (and also the AC_L curve) facing the monopsonist employer.

However, at wage rate W_2, the monopsonist's MC_L curve is the 'double-kinked' line W_2XZV. Employment rises to L_2, the level of employment at which the MRP curve intersects the vertical section between X and Z at point C on the double-kinked MC_L curve. Both the wage rate and employment have risen compared with the situation without a union.

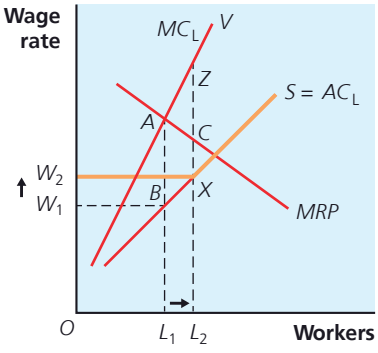


Figure 35 The effect of introducing a trade union into a monopsony labour market

The national minimum wage

Figure 35 can also be used to explain and analyse the possible effects of introducing a **national minimum wage (NMW)** rate. In a competitive labour market, a minimum wage rate set at W_2 increases wages for the workers who keep their jobs, but creates unemployment. By contrast, both wages and employment may rise if the labour market

Exam tip

Avoid confusing the two terms 'monopsony' and 'monopoly'.

National minimum wage (NMW) The lowest level of pay that is legally allowable.

is monopsonistic. In 2016, the UK government began the process of replacing the NMW with a **national living wage (NLW)**, a process that was completed in 2020.

The advantages and disadvantages of a national minimum wage

Advantages

- A national minimum wage reduces exploitation by employers of low-paid workers.
- Exploitation of low-paid workers also increases poverty. Supporters of the NMW believe that a national minimum wage, set at a 'fair' level, should be a vital part of government policy which aims to prevent 'rogue' employers driving wage rates down to 'poverty rates'.

Disadvantages

- A national minimum wage increases unemployment if set at too high a level.
- The incentive function of prices is distorted in the labour market, leading to less efficient functioning of the price mechanism.
- There is evidence that minimum wages lead firms to replace lower-skilled and less experienced younger workers with older workers.

Discrimination in the labour market

Just as price discrimination occurs when firms with monopoly power charge different prices based on customers' different willingness to pay, so **wage discrimination** takes place when employers with monopsony power pay different wage rates based on workers' different willingness to supply labour. In Figure 36, all workers in a competitive labour market receive a wage rate of W_1, determined by supply and demand. Employers' total wage costs are shown by the rectangle bounded by points OW_1AL_1. But if, instead of paying W_1 to all workers, employers pay each worker the minimum they are prepared to work for, the total wage bill falls to equal the shaded 'wedge' area bounded by the points $OBAL_1$. Employers thus gain at the expense of workers, which is why firms pay, and trade unions resist, discriminatory wages whenever possible.

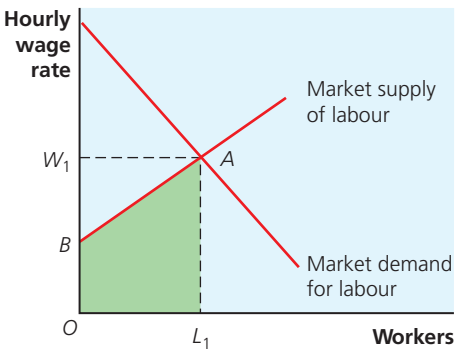


Figure 36 Wage discrimination

National living wage (NLW) The NLW is higher than the national minimum wage — workers get it if they are over 25. There are different rates for different age bands below the age of 25.

Exam tip

The specification was written before the national living wage was introduced, so for the time being exam questions will refer to the NMW rather than to the NLW. However, the advantages and disadvantages are the same.

Wage discrimination Paying different workers different wage rates for doing the same job.

Other explanations of different wages

Even in highly competitive labour markets, wage differences exist, largely because the labour demand and supply curves are in different positions in different labour markets, reflecting factors such as varying labour productivity, ability and required skill. Also, as noted above, different jobs have different non-monetary characteristics, often in the form of job satisfaction or dissatisfaction. Other things being equal, a worker must be paid a higher wage to compensate for any relative unpleasantness in the job. An equalising wage differential is the payment that must be made to compensate a worker for the different non-monetary characteristics of jobs so that, following the payment, the worker has no incentive to switch between jobs or labour markets.

Different wages paid to different groups of workers may also result from forms of labour market discrimination unrelated to the theory of wage discrimination. Some employers discriminate on the basis of race, religion, gender and age, even though such discrimination is illegal. Workers suffering labour market discrimination have poorer job opportunities and are generally less well paid than workers fortunate enough to avoid such discrimination.

Summary

- In labour markets, workers supply labour and firms or employers demand the services of labour.
- In a perfectly competitive labour market, the *MRP* curve is an employer's demand curve for labour.
- In perfect competition, the market supply of labour is the sum of the supply curves of labour of all the workers in the labour market.
- Likewise, in perfect competition, the market demand for labour is the sum of the demand curves of labour of all the employers in the labour market.
- In perfect competition, the equilibrium wage is set where market demand equals market supply, and all the firms and workers in the market are passive price-takers at the market-set wage.
- To maximise profit in all labour markets, firms employ workers up to the point at which $MRP = MC_L$.
- However, the wage rate and level of employment are lower in monopsony than for the whole of a perfectly competitive labour market.
- If a trade union is formed in a previously perfectly competitive labour market, any wage rate increase it achieves is likely to reduce the demand for labour.
- However, in monopsony, a trade union may, within certain limits, be able to increase both the wage rate and the level of employment.
- Similar outcomes result if a national minimum wage is imposed in perfectly competitive and monopsony labour markets.
- Wage discrimination takes place when employers with monopsony power pay different wage rates based on workers' different willingness to supply labour.

The distribution of income and wealth: poverty and inequality

The distribution of income and wealth

The difference between income and wealth

Income and **wealth** illustrate the key difference between **flow and stock concepts** in economics. Income is a flow, measured per period of time — for example, weekly, monthly or annually. The stock of wealth, by contrast, accumulates over time. The different factors of production receive different types of income. Employees are paid wages and salaries; owners of land and property receive rent; interest is paid to lenders of financial capital; and profit is the residual earned by the owners of businesses and entrepreneurs. Transfers, such as unemployment benefits, are another important type of income, especially for those who are without work or on very low incomes. As the name indicates, transfers shift income from taxpayers to benefit recipients, without production of a good or service by the person receiving the benefit.

Various factors which influence the distribution of income and wealth

In the UK, as in most countries, the distributions of income and wealth are both unequal, but the distribution of wealth is significantly more unequal than the distribution of income. The link between wealth and income partly explains this. For those who are better off, wealth generates investment income, part of which — being saved — then adds to wealth and generates even more income. Those at the other end of the scale, by contrast, who possess little or no wealth, have incomes (from low-paid jobs and/or welfare benefits) that are too low to allow saving and the accumulation of wealth.

The tax system also provides an explanation. In the UK, income is usually taxed, but wealth is generally untaxed, though when it is taxed, the rich elite can often avoid paying wealth taxes.

The distribution of income measures how personal or household income is distributed among different income groups in society. Among the factors that influence the distribution of income are employment and unemployment, sickness and disability, possession or lack of possession of qualifications and skills, age, and whether or not one lives in a country or region of high employment and pay.

Taking the working population as a whole, employed workers enjoy higher incomes than those who are unemployed, whose only source of income is welfare benefits set below average levels of wages and salaries. For a similar reason, people dependent on sickness and disability benefits generally suffer from low incomes. High incomes are also strongly correlated with qualifications and skills, and workers living in areas of high employment and job opportunity enjoy higher incomes than those living in areas of deprivation. While some elderly people living beyond retirement age enjoy comfortable private pensions, the same is not true for older people solely dependent on the state pension, which is regarded by many as a poverty income.

Knowledge check 30

Why is income a flow and wealth a stock?

Exam tip

Make sure you don't confuse income and wealth and flows and stocks.

Income Incoming payments received from the supply of factor services.

Wealth The current value of assets accumulated over time.

Flow and stock concepts A stock is measured at one specific time, while a flow is measured over a period of time, for example monthly, quarterly and annually.

People can hold wealth in physical assets such as land, houses, art and antiques, or in financial assets such as stocks and shares. Houses and shares are forms of **marketable wealth**, whose value can appreciate (go up in value) or depreciate (go down in value). Some forms of wealth are non-marketable. An example of **non-marketable wealth** is the stock of wealth accumulated when a person contributes to a pension scheme which cannot be sold to someone else.

Marketable wealth Assets, such as houses and stocks and shares, which can be sold.

Non-marketable wealth Assets, such as a claim on a future pension, which cannot be sold.

Wealth can be divided into inherited wealth, which is passed on from one generation to the next, and non-inherited wealth, which is often the result of entrepreneurial flair and skill. These are sometimes called 'old wealth' and 'new wealth', although the creators of 'new wealth' often decide, on or before their death, to pass it on to the next generations in their families. For the owners of both 'new' and 'old' wealth, the ability to save generated by the high incomes that wealth yields leads to the accumulation of even more wealth in future years.

The difference between equality and equity in relation to the distribution of income and wealth

Complete **equality of income** is achieved when each person receives exactly the same amount of income. The degree of inequality is indicated by the extent to which people's incomes differ. Equality (and inequality) are thus positive concepts, which can be measured. By contrast, **equity**, which means fairness or justness, is a normative concept, which cannot be measured. Different people form different value judgements on what is equitable or inequitable.

Equality of income When everybody, or all households, receive the same income.

Equity When people are treated fairly.

Horizontal equity involves treating people in the same circumstances equally, in contrast to **vertical equity**, which involves taking income from the richest people on the grounds that they do not need it and redistributing this income to the poorer people on the grounds that they do need it.

Horizontal equity Treating people in the same circumstances equally.

Vertical equity Taking income from rich people because they do not need it and redistributing this income to poor people because they do need it.

The Lorenz curve and Gini coefficient

The extent to which the distribution of income is equal or unequal can be illustrated on **Lorenz curve** diagrams, such as in Figure 37, with the degree of inequality measured by a statistic known as a **Gini coefficient**.

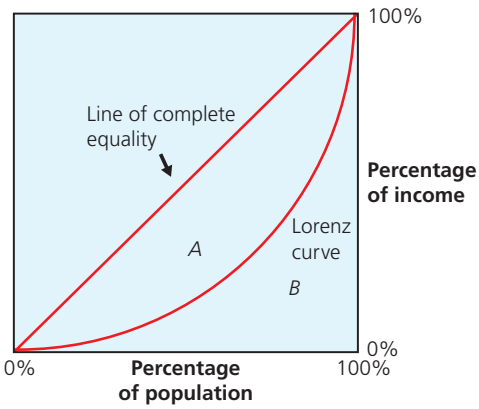


Figure 37 A Lorenz curve

Lorenz curve A line drawn on a graph that shows the extent to which income is equally or unequally distributed among the population.

Gini coefficient Measure of income inequality based on the Lorenz curve.

The Lorenz curve in Figure 37 shows population on the horizontal axis, measured in cumulative percentages from 0% to 100%. The vertical axis shows the cumulative percentage of income received by the population. If incomes were distributed equally, the Lorenz curve would lie along the diagonal line. The nearer the Lorenz curve is to the diagonal, the more equal is the distribution of income. The Gini coefficient measures the area between the Lorenz curve and the diagonal as a ratio of the total area under the diagonal.

In terms of the figure, the Gini coefficient is calculated using the following formula:

$$\text{Gini coefficient} = \frac{\text{area A}}{\text{area A} + \text{area B}}$$

The lower the value of the Gini coefficient, the more equally household income is distributed.

The likely benefits and costs of more equal and more unequal distributions

It is sometimes argued, particularly by economists who favour government intervention in markets, that a more equitable or fair distribution of income and wealth can lead to faster economic growth. People on low incomes and who possess little wealth generally spend all or most of their incomes on consumption and save very little or nothing at all. Greater spending on consumer goods and services increases aggregate demand in the economy, which promotes economic growth. By contrast, people on higher incomes and with a degree of wealth spend a smaller fraction of their incomes on consumption, which leads to slower growth.

Inequality can also mean that the talents of some people in society are wasted or, at least, not fully exploited. For example, children from families on low incomes are likely to do less well than children with affluent parents. Inequality in income usually means that there is also inequality of opportunity.

However, economists of a more pro-free-market persuasion generally reject this view of the world, partly on the grounds that it takes no account of incentives and disincentives. They argue that the progressive taxation of higher incomes and wealth, combined with the transfer of taxed income to those who are less well off, in the form of welfare benefits, significantly reduces the incentives to work hard, among both the high-income earners and those with low-paying jobs.

The problem of poverty

Poverty is the state of being extremely poor and not having enough money or income to meet basic needs, including food, clothing and shelter. The World Bank describes poverty as follows:

Poverty The state of being poor.

> **Poverty is hunger. Poverty is lack of shelter. Poverty is being sick and not being able to see a doctor. Poverty is not having access to school and not knowing how to read. Poverty is not having a job, is fear for the future, living one day at a time.**

The difference between relative and absolute poverty

Poverty is closely related to inequalities in the distribution of income and wealth. However, we must distinguish between **absolute poverty** and **relative poverty**. Because the UK is a high-income developed economy, in which welfare benefits provide a minimum income and safety net for those most in need, very few people are absolutely poor. For the most part, the problem of poverty in the UK is one of relative poverty.

A household is in relative poverty if its income is below a specified proportion of average income for all households — for example, less than a third of average income. Possible causes of relative poverty include: unemployment, especially long-term unemployment; old age and longevity; single parenthood; the decline of employment opportunities in traditional industries and skill fields; lack of education and training; the fall in the value of welfare and unemployment benefits relative to wages and salaries; and the higher incomes and tax cuts enjoyed by those who are already better off.

By contrast, absolute poverty occurs when income is below a particular specified level. When all incomes grow, absolute poverty falls, but relative poverty falls only if low incomes grow at a faster rate than average incomes.

Absolute poverty
Not having the income necessary to meet the minimum requirements for one or more basic living needs over a period of time.

Relative poverty When people lack the minimum amount of income needed in order to maintain the average standard of living in the society in which they live.

The causes and effects of poverty

Three of the main causes of relative poverty in the UK are old age, unemployment and the low wages of many of those in work. Old age causes relative poverty largely because many old people rely on the state pension and lack a private pension. Unemployment benefits are generally lower than the pay workers received before losing their jobs. An increase in unemployment therefore increases poverty. In terms of low wages, low-wage earners are almost always relatively poor rather than absolutely poor. In contrast, some unemployed people, including homeless people living on the street, fall into the category of absolute poverty.

In terms of its effects, poverty leads to educational deprivation, health deprivation and adverse effects on the communities in which poor people live. Children from poorer backgrounds lag at all stages of education. In terms of health deprivation, poverty is also associated with a higher risk of both illness and premature death. Poorer health over the course of a lifetime has an impact on life expectancy: professionals live, on average, 8 years longer than unskilled workers. And finally, with regard to the effect of poverty on communities, children living in poverty are likely to live in bad housing.

Old-age poverty, and the fact that people are living for longer, lead to other adverse effects on society. A recent example at the time of the coronavirus pandemic centred on elderly but poor hospital patients continuing to occupy hospital wards because they had nowhere else to go. The government reacted by moving many elderly patients into care homes, which had the unintended consequence of transmitting the virus into the wider population.

Exam tip

Exam questions may be set on absolute and relative poverty Make sure you don't confuse the causes of poverty with the economic *effects* of poverty, and with government policies to reduce poverty, particularly progressive taxation and transfers.

Knowledge check 31

How many people live in conditions of relative poverty at the time you are reading this topic?

Government policies to alleviate poverty and to influence the distribution of income and wealth

By reducing inequalities in the distribution of income, progressive taxation and transfers (welfare benefits) can reduce absolute and relative poverty — provided that labour market incentives, competitiveness and economic growth do not worsen significantly. Absolute poverty, though not necessarily relative poverty, can best be reduced by fast and sustained economic growth and by creating jobs. It follows that successful government policies that promote economic growth alleviate poverty and influence the distribution of income and wealth.

When describing and explaining the distribution of income, it is useful to understand the difference between various measures of income: original income, gross income, disposable income, post-tax income and final income. The relationship between these different measures of income is shown in Figure 38.

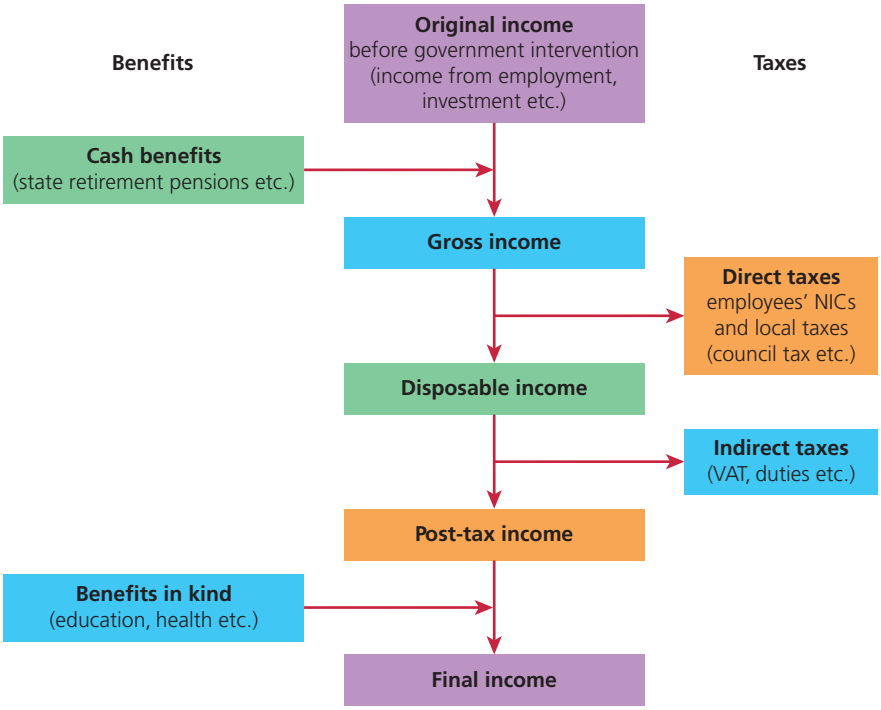


Figure 38 The different ways in which income is measured and their relation to taxes and benefits

The tax and welfare benefits systems affect income in the following ways. First, household members begin with income from employment, private pensions, investments such as the shares they own, and other non-government sources. This is referred to as '**original income**'. Second, some then receive income from cash benefits. The sum of cash benefits and original income is referred to as '**gross income**'. They then pay direct taxes. Income after direct taxes have been subtracted

Original income The income received by a household or person before allowing for any payment of income tax and other taxes, and the receipt of various transfer payments such as welfare benefits.

Gross income The sum of all wages, salaries, profits, interest payments, rents, and other forms of earnings, before any deductions or taxes.

from gross income and cash benefits are added is called '**disposable income**'. Indirect taxes are then paid on spending on goods and services. Disposable income minus indirect taxes is referred to as '**post-tax income**'. Finally, many households receive a benefit from services provided by the government (benefits in kind). Post-tax income plus benefits in kind is referred to as '**final income**'.

Disposable income Gross income minus direct taxes and social security payments.

Post-tax income Disposable income minus indirect taxes such as VAT.

Final income Post-tax income plus benefits in kind.

Summary

- Income is an example of an economic flow, whereas wealth is an example of an economic stock.
- In the UK, the distributions of income and wealth are both unequal, with the distribution of wealth being significantly more unequal.
- Different conditions in labour markets, differences between unwaged and low-waged people, and the nature of the welfare benefits system all provide explanations of income inequality.
- Horizontal equity involves treating people in the same circumstances equally, in contrast to vertical equity, which involves taking income from the richest in society on the grounds that they do not need it and redistributing this income to those at the bottom on the grounds that they do need it.
- Absolute poverty and relative poverty are the two main poverty concepts.
- Progressive taxation and transfers are two government (fiscal) policies used to reduce both absolute and relative poverty.

The market mechanism, market failure and government intervention in markets

To remind you, a market is a voluntary meeting of buyers and sellers. Both buyer and seller have to be willing partners to an exchange. Markets do not have to exist in a particular geographical location. Whenever a good or service is voluntarily bought and sold, a market transaction takes place. Over history, market transactions shifted away from open-air street markets to take place in shops. Shops have higher overhead costs, but they offer a permanent site of exchange and a continuing relationship between sellers and buyers. In recent years, the growth of the internet has allowed 24/7 e-commerce. As a result, many markets, especially those in commodities, raw materials and financial services, have become truly global.

A market is highly competitive when there is a large number of buyers and sellers all passively accepting the ruling market price that is set, not by individual decisions, but by the interaction of all those taking part in the market. The ruling market price (or equilibrium price) is set by supply and demand in the market as a whole. Highly competitive markets lack entry and exit barriers. This means that new buyers and sellers can easily enter the market without incurring costs. In the same way, buyers and sellers can leave the market if they wish to. Competitive markets also exhibit a high degree of transparency — buyers and sellers can quickly find out what everyone else in the market is doing.

Knowledge check 32

Name four markets which exist in specific geographical locations.

How markets and prices allocate resources

The four functions prices perform in markets

To understand how markets and prices allocate resources in an economy, we must explain the four functions that prices perform in markets. These are the:

- **signalling function**
- **incentive function**
- **rationing function**
- **allocative function**

In the first place, prices provide information that allows buyers and sellers in a market to plan and coordinate their economic activities. This is the signalling function of prices, which leads into the incentive function of prices. The information signalled by changing relative prices creates incentives for people to alter their economic behaviour. For example, a higher price creates an incentive for firms to supply more of a good or service. By contrast, the rationing function of prices relates to demand rather than to supply: a rising price rations demand for a product.

The rationing function of prices is related to, but not quite the same as, the allocative function of prices. The rationing function distributes scarce goods to those consumers who value them most highly. By contrast, the allocative function directs resources

Signalling function
Prices provide information to buyers and sellers.

Incentive function
Prices create incentives for people to alter their economic behaviour, e.g. a higher price creates an incentive for firms to supply more of a good or service.

Rationing function
Rising prices ration demand for a product.

Allocative function
Changing relative prices allocate scarce resources away from markets exhibiting excess supply and into markets in which there is excess demand.

between markets, away from the markets in which prices are too high and in which there is **excess supply**, towards the markets where there is **excess demand** and price is too low.

Excess supply When firms wish to sell more than consumers wish to buy, with the price above the equilibrium price.

Excess demand When consumers wish to buy more than firms wish to sell, with the price below the equilibrium price.

How prices coordinate the decisions of buyers and sellers in a market economy

Communication is essential in a well-functioning economy. In a market-based economic system, price signals help prevent shortages and surpluses. All of this happens without the need for government intervention and generally ensures that consumer wants are largely satisfied. Imagine that a natural disaster such as an extremely wet summer results in a shortage of vegetables. In a market-based economic system like the UK economy, the market, and not the government, determines the price of vegetables and the quantity produced and consumed. Price signals communicate in such a way that prevents massive shortages and surpluses and ensures that consumer wants are largely satisfied. The actual price of the good or service—in this case, vegetables —provides an incentive to buyers and sellers. In an extremely wet summer in which farmland is flooded for several months, the supply disruption would cause the price of vegetables to rise. Higher vegetable prices signal to buyers to reduce their consumption and to sellers to increase their production, and for imports to increase. Both buyers and sellers would have an incentive to do this. In this way, price signals allow markets to function efficiently and for buyers' and sellers' decisions to be coordinated.

The advantages and disadvantages of the price mechanism

Advantages

A market economy, in which the price mechanism operates, contains a very large number of individuals and firms, each pursuing their own self-interest. Yet, instead of anarchy and chaos resulting, the price mechanism can produce an economically more efficient outcome than can be achieved in command or centrally planned economies. Via the signalling, incentive, rationing and allocative functions of prices, and in the absence of various forms of market failure (which we shortly explain), the price mechanism decides what, how, how much and for whom to produce in a way that is best for society.

In competitive markets and assuming an absence of externalities, an allocatively efficient outcome is achieved with $P = MSC$ in all markets.

Consumer sovereignty results from the pursuit of individual self-interest on the part of individual consumers and households.

This outcome is achieved in an uncoordinated way with no bureaucracy costs.

Disadvantages

The advantages of the price mechanism are strongest when real-world markets approximate to perfect competition. However, when markets are monopolistic or oligopolistic, consumer sovereignty gives way to producer sovereignty. Prices end up being too high, leading to an allocatively inefficient outcome in which too little of a good or service is produced and consumed.

Prices can also be too low, being below the marginal social cost of production, leading to too much consumption of a good such as alcoholic beverages.

The price mechanism can contribute to highly inequitable distributions of income and wealth.

Knowledge check 33

Explain the difference between the rationing and the allocative functions of prices.

Extending the price mechanism into new areas of activity

Since about the 1980s, UK government, particularly Conservative governments, have tried to change the nature of the British economy through pursuing policies of economic liberalisation. One of these policies has been marketisation or commercialisation. This involves charging prices for goods and services previously provided free or heavily subsidised by the state. One of the oldest examples is charging for prescription medicines provided by the National Health Service.

The meaning of market failure

Market failure occurs whenever the market mechanism or price mechanism perform unsatisfactorily, with the result that resource misallocation occurs. There are two main ways in which markets fail. Markets can function inefficiently, or they can function inequitably.

It is also useful to distinguish between complete market failure, when the market simply does not exist, and partial market failure, when the market functions but produces the 'wrong' quantity of a good or service. In the former case, there is a **missing market**. In the latter case, the good or service may be provided too cheaply, in which case it is over-produced and over-consumed. Alternatively, as in monopoly, the good may be too expensive, in which case under-production and under-consumption result.

In the next sections, we explain how public goods, positive and negative externalities, merit and demerit goods, monopoly and other market imperfections, and inequalities in the distribution of income and wealth can lead to market failure.

Public goods, private goods and quasi-public goods

Public goods divide into **pure public goods** and **quasi-public goods**. Pure public goods, such as national defence, exhibit the characteristics of **non-excludability** and **non-rivalry**, together with a third characteristic of **non-rejectability**.

Market failure When the price mechanism leads to an inefficient or inequitable outcome.

Missing market A market functions, but it delivers the 'wrong' quantity of a good or service, which results in resource misallocation.

Public goods Goods that exhibit the characteristics of non-excludability and non-rivalry.

Pure public good A good that is always non-excludable and non-rival.

Quasi-public good A good that has characteristics of both a public and a private good.

Non-excludability It is impossible to provide the good to one person while preventing others from enjoying it.

Non-rivalry Such goods can be consumed by one consumer without preventing simultaneous consumption by others. For example, when a person benefits from national defence, it does not prevent other people from experiencing similar benefits. By contrast, an item of food, such as a strawberry, is a rival good, because when one person eats it, nobody else can consume it.

Non-rejectability A property of a public good which means that if the good is provided, it is impossible for a person to 'opt out' and not gain its benefits.

With public goods, individuals face the temptation to consume or benefit without paying, or to free-ride. Provided that most people believe nuclear defence to be necessary, a market is unlikely to provide a service which costs billions of pounds to provide. Few of the people who would benefit from nuclear defence would be willing to pay a high market price for the service. Instead, they would choose to benefit without paying, and the market would collapse. Public goods such as nuclear defence are thus associated with the **free-rider problem**. Free-riding occurs when people decide to gain the benefits of a good or service while refusing to pay for it.

Many, perhaps most, public goods are quasi-public goods (or non-pure public goods) to exclude free-riders. Non-pure public goods include roads, television and radio broadcasts, street lighting and lighthouses. In principle, roads can be converted into **private goods**, provided for profit through the market. This could be done by limiting points of access, by constructing toll gates or by introducing a scheme of electronic road pricing. However, the cost of making the good excludable may be very high, for example road-pricing schemes on minor roads.

Free-rider problem When non-excludability leads to a situation in which not enough customers choose to pay for a good, preferring instead to free-ride or benefit without paying.

Private goods Private goods possess the characteristics of excludability and rivalry.

Exam tip

Make sure you understand the difference between a public good and a quasi-public good and can give examples of both.

The tragedy of the commons

Fifty years ago, University of California professor Garrett Hardin published an influential essay in the journal *Science*. Hardin saw all humans as selfish herders: we worry that our neighbours' cattle will graze the best grass, so we send more of our cows out to consume that grass first. We take it first, before someone else steals our share. This creates a vicious cycle of environmental degradation that Hardin described as the 'tragedy of the commons'.

Knowledge check 34

What is the difference between a pure public good and a quasi-public good?

Public goods and allocative efficiency

The allocatively efficient or 'correct' quantity of any good produced and consumed is the quantity that people choose to consume when $P = MC$. However, in the case of a public good, assuming it is already being provided, the MC of providing the good to an extra consumer is zero. Allocative efficiency therefore occurs when $P = 0$ and the good is free for consumers. But private entrepreneurs willingly provide goods only if profit can be made, and for this to happen the price must be above zero ($P > 0$).

In the case of a quasi-public good such as a road, this means that markets can provide a road only if the price of road use is set above the marginal cost of supply ($P > MC$). This reduces road use to below the allocatively efficient level. Hence, to achieve an allocatively efficient level of road use, motorists should not be charged for driving on the road, at least until the road becomes congested.

Exam tip

Make sure you understand the concept of allocative efficiency.

Positive and negative externalities in consumption and production

An **externality** is a special type of public good or public bad, which is dumped by those who produce it on third parties who receive or consume it, whether or not they choose to. Because externalities are generated and received outside the market, they provide examples of missing markets.

Externality A special type of public good or public bad, which is dumped by those who produce it on third parties who receive or consume it, whether or not they choose to.

Externalities are classified in two main ways:

- as negative externalities and positive externalities, also known as external costs and external benefits
- as **production externalities** or as **consumption externalities**.

A negative externality is generated when an individual or firm making a decision to produce or consume a good or service does not have to pay the full cost of the decision. If the production of a good generates a negative externality, then the cost to society is greater than the cost incurred by the firm itself. Road congestion is a negative externality.

A positive externality is generated when an individual or firm making a decision does not receive the full benefit of the decision. The benefit to the individual or firm is less than the benefit to society. The pollination of fruit trees on neighbouring farms is a positive externality resulting from bee keeping.

As is the case with public goods, the production and consumption of externalities leads to the free-rider problem. The provider of an external benefit such as a beautiful view cannot charge a market price to any willing free-riders who enjoy it. Conversely, the unwilling free-riders who receive or consume external costs such as pollution and noise cannot charge a price to the polluter for the bad that they reluctantly consume.

A production externality is generated, usually by firms, in the course of producing a good or service. By contrast, a consumption externality is generated by households and individuals in the course of consuming a good or service.

Production externality When production of a good or a service imposes external costs or benefits on third parties outside of the market without these being reflected in market prices.

Consumption externality When consumption of a good or a service imposes external costs or benefits on third parties outside of the market without these being reflected in market prices.

Market imperfections

Market failures are often the result of market imperfections. In the case of merit goods and demerit goods, imperfect and asymmetric information can lead to market failure; the existence of monopoly and monopoly power can also lead to market failure. And the immobility of factors of production such as capital and labour can lead to regional inequalities, which are a form of market failure.

Many market imperfections and market failures can be analysed using the concepts of marginal private, external and social costs and benefits. Before we make use of these concepts, we shall first state a number of key definitions:

- Private benefit maximisation occurs when marginal private benefit (*MPB*) = marginal private cost (*MPC*).
- Social benefit maximisation, which maximises the public interest or the welfare of the whole community, occurs when marginal social benefit (*MSB*) = marginal social cost (*MSC*).
- Marginal social benefit (*MSB*) = marginal private benefit (*MPB*) + marginal external benefit (*MEB*).
- Marginal social cost (*MSC*) = marginal private cost (*MPC*) + marginal external cost (*MEC*).

Negative externalities and allocative inefficiency

Profit maximisation occurs in a perfect market at the price at which $P = MPC$. In the absence of externalities, this also means that the price equals the marginal social cost (MSC) of production: $P = MSC$. But when external costs are generated in the production of a good, $MSC > MPC$. So when $P = MPC$, $P < MSC$.

To achieve allocative efficiency, price must equal the true marginal cost of production: that is, the marginal social cost and not just the marginal private cost. But in a market situation, when externalities exist, the market mechanism fails to achieve an allocatively efficient outcome.

Negative production externalities and market failure

Consider a fossil-fuel burning power station which, while producing electricity, discharges negative externalities but not positive externalities. In this situation, the marginal private benefit accruing to the power station from the production of electricity, and the marginal social benefit received by the whole community, are the same and shown by the downward-sloping curve in Figure 39. But, because pollution is discharged in the course of production, the marginal social cost of electricity production exceeds the marginal private cost incurred by the power station. In Figure 39, the MSC curve is positioned above the MPC curve. The vertical distance between two curves shows the marginal external cost (MEC) at each level of electricity production.

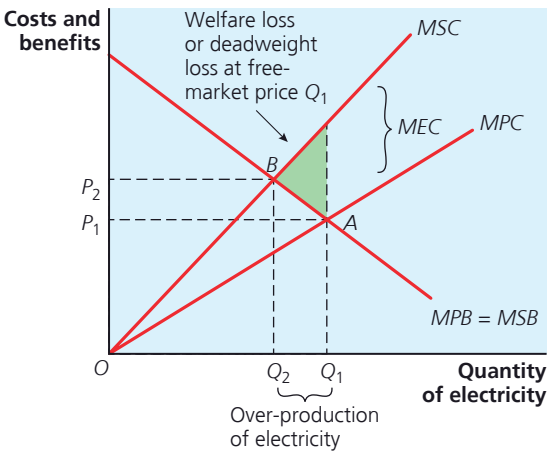


Figure 39 A coal-burning power station generating pollution (a negative production externality)

The power station maximises private benefit by producing output Q_1, where $MPC = MPB$. Q_1 is immediately below point A in Figure 39. However, the socially optimal level of output is Q_2, where $MSC = MSB$. Q_2 is immediately below point B. The shaded area illustrates the 'loss' of welfare or deadweight loss (DWL), which exists at the free-market output, Q_1 (where $MPC = MPB$), all the way back to the socially optimum output, Q_2.

Exam tip

Although coal-burning power stations are due to be phased out in the UK in the next few years, they can still be used as an example of negative production externalities.

Positive production externalities and market failure

When positive production externalities are generated, the marginal social costs of production lie below the marginal private costs incurred by the producers of the good or service. This is illustrated in Figure 40, which shows the costs incurred when a commercial forestry company plants trees.

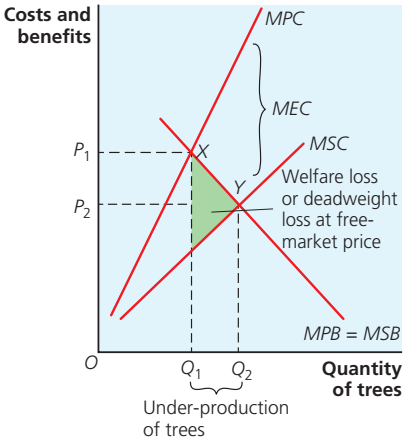


Figure 40 A commercial forestry company generating positive production externalities

Positive production externalities such as these mean that the *MSC* curve is positioned below the *MPC* curve. The vertical distance between the two curves shows a negative marginal external cost (*MEC*) at each level of tree planting. (A negative marginal external cost is exactly the same as a positive marginal external benefit enjoyed by society as a whole.)

Merit goods and demerit goods

A **merit good**, such as education or healthcare, has two important characteristics: positive externalities in consumption, and information problems which distort a consumer's choice on what is the privately optimal level of consumption. Also, whether a good is viewed as a merit good or a demerit good can depend on a value judgement.

When a person consumes a merit good such as healthcare, the resulting positive externalities benefit other people. An obvious example is that healthy people seldom spread diseases. The social benefit enjoyed by the wider community is greater than the private benefit enjoyed by the healthy person.

With regard to information problems, when deciding how much to consume, individuals take account of short-term costs and benefits, ignoring or undervaluing the long-term private costs and benefits. So whereas free-market provision leads to over-provision of a demerit good such as an alcoholic drink, it leads to under-provision of merit goods such as education and healthcare.

As is the case with a merit good, there are two main characteristics of a **demerit good**, such as tobacco and alcoholic beverages. In the first place, when people consume demerit goods, they discharge negative externalities which are dumped on other people or third parties. The marginal social costs suffered by the wider community are greater than the marginal private costs incurred by a smoker or drinker. The second characteristic of a demerit good centres on the distinction between the short-term and the long-term private costs incurred by the person consuming the demerit good. For

Merit good
Consumption of a merit good such as education generates positive externalities that are enjoyed by others. Also, the long-term private benefits of consumption exceed the short-term private benefits.

Demerit good
Consumption of a demerit good such as tobacco generates negative externalities that are suffered by others. Also, the long-term private costs of consumption exceed the short-term private costs.

example, when teenagers first get the 'habit' of smoking, drinking or drug taking, they may either ignore the long-term private costs they might suffer later in life, or downplay the significance of these costs. This behaviour illustrates an information problem.

These characteristics of demerit goods mean that free-market provision leads to market failure. Left to themselves, markets over-provide demerit goods.

Knowledge check 35

Explain the difference between a demerit good and an economic 'bad' such as garbage.

Positive consumption externalities and under-consumption of merit goods

Taking education as an example, when schooling is available only through the market, at prices unadjusted by subsidy, too little of the merit good ends up being consumed. In Figure 41, the privately optimal level of consumption is Q_1, determined at point A, where $MPC = MPB$. This is below the socially optimal level of consumption, Q_2.

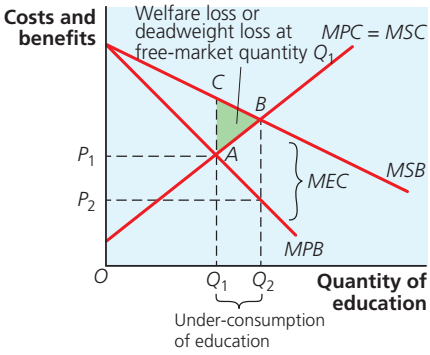


Figure 41 Under-consumption of a merit good in a free market

Negative consumption externalities and over-consumption of demerit goods

Figure 42 shows that too much tobacco is being consumed when bought at market prices unadjusted by taxes or by a minimum price law. At least in the short term, the privately optimal level of consumption is Q_1, where $MPC = MPB$. This is greater than the socially optimal level of consumption, Q_2, located where $MSC = MSB$. Free-market provision of demerit goods therefore leads to over-consumption and hence over-production.

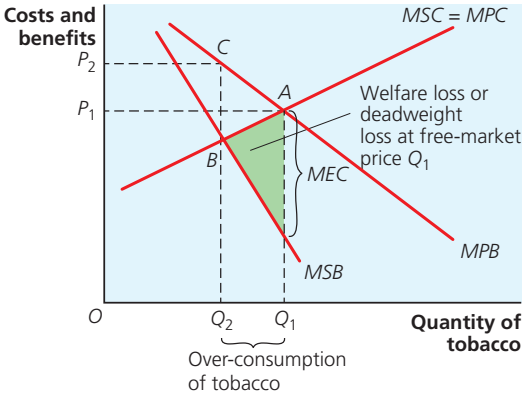


Figure 42 Over-consumption of a demerit good in a free market

The absence of property rights and market failure

A well-functioning market economy requires the establishment and enforcement of well-defined property rights. Throughout history many people have lived close to rivers. On the one hand, rivers have provided vital drinking water, but on the other hand they provide a tempting way of disposing of waste. However, while being an important natural resource, for much of history rivers have been owned by nobody. As a result of the absence of property rights with regard to the ownership, for either consumption or disposal purposes, river water has tended to be polluted and over-used.

Competition policy

As its name implies, **competition policy** is the part of government economic policy that tries to make the imperfectly competitive and monopolistic markets of the real world more competitive, and akin to the abstract model of perfect competition. The aims or objectives of competition policy include preventing the exploitation of monopoly power, efficiency, getting rid of excessive profit so that prices reflect costs of production, and the removal of entry and exit barriers that separate markets.

However, competition policy does recognise that there are two main circumstances in which monopoly may be preferable to small firms producing in a competitive market. First, when the size of the market is limited but economies of scale are possible, monopolies can produce at a lower average cost than smaller, more competitive firms. Second, under certain circumstances, firms with monopoly power may be more innovative than firms that are not protected by entry barriers. When this is the case, monopoly may be more dynamically efficient than a more competitive market.

Since 2014, UK competition policy has been implemented on behalf of the government by the **Competition and Markets Authority (CMA)**, which was formed through the merger of two older government agencies, the Office of Fair Trading (OFT) and the Competition Commission (CC).

Competition policy The part of the government's microeconomic policy and industrial policy which aims to make goods markets more competitive. It comprises policy towards monopoly, mergers and restrictive trading practices.

Competition and Markets Authority (CMA) The UK government agency which implements UK competition policy.

European Union competition policy

Before the United Kingdom left the European Union (EU) at the end of January 2020, UK competition policy was overarched by EU competition policy. EU merger policy, which is the main part of EU competition policy, is based on the principle of subsidiarity, which delegates policy as much as possible to national governments. Member countries continue to use national policy to deal with smaller mergers, but the European Commission adjudicates on larger mergers with a community dimension.

Exam tip

Although the A-level specification still states (in 2020) that students should be aware of EU competition policy and its costs and benefits, it is unlikely that exam questions will be set on these topics.

The costs and benefits of UK competition policy

Because economists recognise that monopoly can be either good or bad depending on circumstances, UK policy has always been based on the view that each case must be judged on its merits. If the likely costs resulting from the reduction of competition exceed the benefits, monopoly should be prevented, but if the likely benefits exceed the costs, monopoly should be permitted. Ongoing regulation is needed to make sure that firms, particularly large firms, continue to act in the public interest. However, very few mergers are actually investigated each year. Despite heavy fines being available for punishing collusion among firms, covert collusion is difficult to prove,

and tacit collusion is almost impossible to prove. 'Single' markets are inadequately defined. One particular criticism of regulation in general is that it is subject to regulatory capture — where the regulator is captured by the regulated (see below).

Public ownership, privatisation, regulation and deregulation of markets

Privatisation involves the transfer of firms and businesses from **public ownership** or state ownership to the private sector.

Economic **regulation** involves the imposition of rules, controls and constraints, which restrict freedom of economic action in the market place. **Deregulation** is the opposite — the removal of previously imposed regulations. Governments use regulation to try to correct market failures and to achieve a socially optimal level of production and consumption. In the case of monopoly, regulation is used to limit and deter monopoly exploitation of consumers.

Pro-free-market governments, including the current Conservative administration, generally prefer deregulation to ever more regulation. Deregulation is an important part of the Conservative government's policy of economic liberalisation.

Deregulation and the theory of contestable markets

Much of the justification for the policies of deregulation and economic liberalisation that have been pursued in recent years has been provided by the theory of contestable markets. Contestable market theory argues that the most effective way to promote competitive behaviour within markets is not to impose ever more regulation on firms and industries, but to carry out the opposite process of deregulation. According to this view, the main function of deregulation is to remove barriers to entry, thereby creating incentives both for new firms to enter and contest the market and for established firms to behave in a more competitive way.

Regulatory capture

Another theory that has had some influence on the trend towards deregulation is the theory of **regulatory capture**. This theory argues that regulatory agencies created by government can be 'captured' by the industries or firms they are intended to oversee and regulate. Following capture, the regulatory agencies begin to operate in the industry's interest rather than on behalf of the consumers they are supposed to protect.

Even if regulatory capture does not take place, the supporters of deregulation argue that much regulatory activity is unnecessary and ultimately burdensome on industry and consumers. Once established, the regulators have an incentive to extend their role by introducing ever more rules and regulations, since it is in this way that they justify their pay and their jobs. Regulation acts both as an informal 'tax' on the regulated, raising production costs and consumer prices, and as an extra barrier to market entry, restricting competition.

Privatisation The transfer of assets from the public sector to the private sector.

Public ownership The ownership of whole industries or individual firms by the state.

Regulation Imposing rules or limits on free-market activity.

Deregulation Removing previously imposed rules or limits on free-market activity.

Exam tip

While there is a strong case for removing many regulations, for example those that unnecessarily raise costs of production and consumer prices, many regulations can be justified on the grounds that they protect people from, for example, the abuse of monopoly power and harmful externalities.

Regulatory capture When regulatory agencies act in the interest of regulated firms rather than on behalf of the consumers they are supposed to protect.

Knowledge check 36

With an example of each, distinguish between privatisation and regulation.

Government intervention in markets

Governments can intervene in markets in two main ways: through regulation and through imposing taxes, for example to discourage consumption of a demerit good or encourage consumption of a merit good. Other forms of intervention include maximum and minimum price laws (**price ceilings** and **price floors**) and the creation of **markets in permits to pollute**.

In a permits-to-pollute scheme, maximum limits are imposed on the amount of pollution that industries are allowed to emit, followed by a steady reduction in each subsequent year (say by 5%) of the maximum amount. Once this regulatory framework has been established, a market in traded pollution permits takes over, creating market-orientated incentives for firms to reduce pollution because they can make money out of it.

The government can also enforce legal entitlement to private property rights. If the law provides people with the property right to breathe unpolluted air, breach of this right enables victims to sue polluting companies for financial compensation. To avoid having to pay financial compensation, polluters would have to take action to eliminate the discharge of negative externalities. Alternatively, if it is impossible to eliminate a negative externality without simultaneously eliminating production of the good that produces the externality, polluters could offer people money to sign away the right to breathe unpolluted air. Either way, if the externality persists, the people who suffer pollution receive financial compensation.

Government failure

Whereas market failure occurs when markets function inefficiently or inequitably, **government failure** can occur when government intervention to try to reduce or eliminate a market failure has the unintended consequence of creating a new problem or problems. If the government failure is greater than the market failure that the government intervention was intended to correct, it is better to leave the market failure uncorrected and to live with it.

Summary

- Market failure occurs whenever the market mechanism or price mechanism performs unsatisfactorily, either inequitably or inefficiently.
- The principal market causes of failures are public goods, externalities, merit and demerit goods and monopolies.
- Public goods divide into pure and quasi-public goods.
- The main 'add-on' topics for understanding market failure are the application of marginal analysis and the concept of allocative efficiency to market failure, and the case for extending private property rights.
- Competition policy, public ownership and regulation are three examples of government intervention in markets.
- The regulation of market activity can lead to the problem of regulatory capture.
- Deregulation reflects the impact of the theory of contestable markets.
- Government failure occurs when government intervention in markets to try to correct market failure is ineffective or damaging.

Price ceiling A limit on the price of a good or service imposed by the government to protect consumers by ensuring that prices do not become prohibitively expensive.

Price floor A mechanism to prevent prices from falling.

Market in permits to pollute A market in which pollution permits are traded. They give companies a legal right to pollute a certain amount per fixed time span. Firms that pollute less can then sell their leftover pollution permits to firms that pollute more.

Government failure When government intervention reduces economic welfare, leading to an allocation of resources that is worse than the free-market outcome.

Knowledge check 37

Distinguish between market failure and government failure.

Exam tip

You must avoid confusing government failure with market failure.

Questions & Answers

The A-level examination

A-level Paper 1

The A-level Paper 1, 'Markets and market failure', is 2 hours long and has a maximum mark of 80. The exam paper contains two sections, A and B, both of which must be answered. Section A, which accounts for 40 marks (50% of the total), comprises two data-response questions (DRQs), labelled Context 1 and Context 2, of which you should answer one. Section B, which also accounts for 40 marks (50% of the total), contains three essay questions (EQs), of which you should answer one.

This Student Guide contains one Context for Section A and one essay question for Section B.

A-level Paper 3

Paper 3, 'Economic principles and issues', also carries a maximum of 80 marks. Section A (30 marks) has 30 multiple-choice questions (MCQs), of which roughly half are on microeconomics and half on macroeconomics.

The MCQs that follow are similar to the sorts of microeconomic themed questions that you might see in the A-level Paper 3 exam.

In Section B of Paper 3, worth 50 marks, there are three extended-response questions based on a case study or investigation which require a student to draw together different areas of the specification. An 'extended response' is evidence generated by a student that is of sufficient length to allow that student to demonstrate the ability to construct and develop a sustained line of reasoning which is coherent, relevant, substantiated and logically structured. The case study in Part B is not pre-released.

This Student Guide provides one such investigation with questions.

Assessment objectives

Assessment objectives (AOs) are set by the government agency Ofqual and are the same across the A-level economics papers. The exams measure how students have achieved the following assessment objectives:

- AO1: Demonstrate knowledge of terms/concepts and theories/models to show an understanding of the behaviour of economic agents (consumers, workers and firms) and how they are affected by and respond to economic issues.
- AO2: Apply knowledge and understanding to various economic contexts to show how economic agents are affected by and respond to economic issues.
- AO3: Analyse issues within economics, showing an understanding of their impact on economic agents.
- AO4: Evaluate economic arguments and use qualitative and quantitative evidence to support informed judgements relating to economic issues.

Assessment objectives 1 and 2 are testing 'lower-order' skills, whereas objectives 3 and 4 test 'higher-order' skills.

Weighting of assessment objectives for A-level Economics

Assessment objectives (AOs)	Component weightings (approx. %)			Overall weighting (approx. %)
	Paper 1	Paper 2	Paper 3	
AO1	5–8	5–8	7–10	20–23
AO2	7–10	7–10	9–12	26–29
AO3	9–11	9–11	6–9	26–29
AO4	7–10	7–10	5–8	22–25
Overall weighting of components	33.3	33.3	33.3	100

The exam questions in this guide

This guide includes 19 examination-style questions designed to be a key learning, revision and exam preparation resource.

In Paper 1, Section A, there are four data-response questions. Then in Section B there are two essay questions (one worth 15 marks; one worth 25 marks).

In Paper 3, Section A, there are ten MCQs. Then in Section B there are three data-response questions (worth 10 marks, 15 marks and 25 marks). Section B is called the INVESTIGATION and all questions refer to the context given in the INSERT.

All the questions in this guide can be used 'en bloc' as part of a short trial or mock exam near the end of your course. Alternatively, as you study a topic in the Content Guidance section of this guide, you could refer selectively to particular questions in this section that assess aspects of the topic.

Note: no past-paper questions, nor students' answers to past-paper questions, have been included in this guide.

This section of the guide also contains:

- correct answers for the MCQs and a brief explanation of the correct answer
- comments on the MCQs, explaining particular features of each question
- student answers of grade-A standard (or sometimes A* standard) and grade-C standard for each DRQ and essay question
- comments on each student's answer, explaining, where relevant, how the answer could be improved.

Paper 1

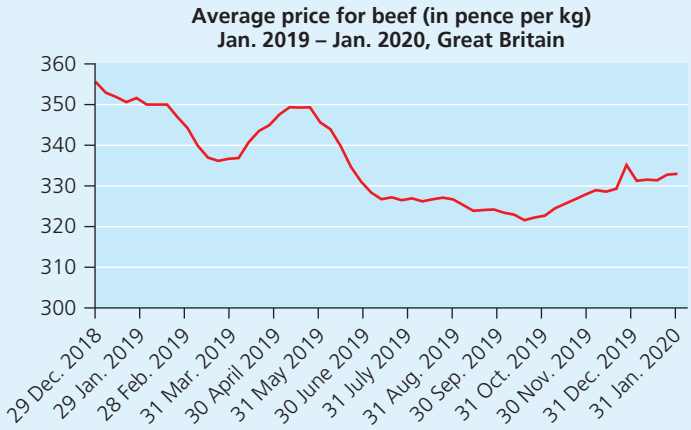
Section A

Total for this Context: 40 marks

Context 1: Veganuary

Study Extracts A, B and C and then answer all parts of Context 1 which follow.

Extract A The price of beef



Source: Agriculture and Horticulture Development Board (AHDB)

Extract B Burger King takes the lead in producing plant-based burgers

In January 2020, Burger King launched its first plant-based burger in the UK. The soy-based version of its Whopper burger is aimed at those who want to cut meat consumption.

It is a fact that more people are beginning to cut meat out of their diet in the UK. Various factors — such as the Netflix documentary ***Gamechangers***, the increasing concern about the meat industry contributing to climate change and meat-based health scares — have all contributed to this change in demand. As a result, many people are trying out Veganuary (which was launched by a non-profit firm which encourages people to go vegan for the month of January).

The fast-food market in the UK is highly competitive. As a result, other fast-food firms like McDonalds (veggie dippers) and Greggs (vegan sausage rolls) are following Burger King's lead and also producing non-meat alternatives.

Source: news reports (2020)

Extract C NFU calls for action due to the collapse of beef prices

The prime minister came under fire from a farmers' union, the National Farmers' Union (NFU) in early 2020 because of the collapse in beef prices. One union member stated that 'cattle farmers have lost nearly £170m in profit over 2019 because of the collapse in demand… current market prices barely cover our costs of production'. The NFU, which is registered as an association of employers under the 1974 Trade Union and Labour Relations Act, called for the government to introduce new subsidies to protect profits in the future. It is the largest farmers' organisation in the UK and boasts 55,000 farmer members.

The ability of trade unions to influence wages and levels of employment in the labour market is keenly debated in economics. In this instance, the prime minister was non-committal and replied that the real problem for cattle farmers was not the shortage of staff but the oversupply of imported meat from EU countries. 'Once we get Brexit done,' he said, 'we can limit the number of non-UK meat supplied into the UK which is causing the demand [and profit] for domestic producers to fall.'

Source: news reports (2019)

Question 1

Using the data in Extract A calculate the percentage change in UK beef prices from January 2019 to January 2020. Give your answer to one decimal place. 2 marks

For calculations questions it is important to show all the steps of your calculations, and to make sure that the final units are correct.

Student A answer

Price in January 2019: 355p/kg

Price in January 2020: 335p/kg

Percentage change is: ((335 – 355) / 355) x 100 = –5.6% to 1 decimal place.

2/2 mark awarded This is a correct calculation, which earns full marks. Showing the working like this is advisable because it shows steps that may have gone wrong. You may still get 1 mark for a particular stage of the calculation even if the full result is wrong.

Student B answer

Price in January 2019: 355p/kg

Price in January 2020: 335p/kg

The change is 20p.

1/2 marks awarded The student has not performed a percentage change calculation despite identifying the correct numbers.

Question 2

Explain how the information in Extract B show how the price mechanism allocates resources 4 marks

A good tip is to define the key word(s) in the question (in this case *price mechanism*). Then, make sure you explain the effect of the price mechanism mentioned in the extract in context.

Student A answer

The price mechanism is the process by which the market forces of supply and demand direct the allocation of resources in an economy. In this case, the demand for meat is falling and the demand for non-meat alternative is rising. While the price of meat falls, consumers of non-meat alternatives will bid up this price. This, in turn, creates an incentive for producers to reallocate their resources and produce more non-meat alternatives. Extract B shows that Burger King, McDonald's and Greggs have all recently introduced non-meat alternatives onto their menus, which shows that they are responding to those incentives.

4/4 marks awarded Although a definition is not essential, it can help to frame your answer. The economic theory offered in this answer is clear and correct. Mentioning an example from Extract B is neatly woven into the analysis, showing strong application skills.

Student B answer

The price mechanism is the process by which the market forces of supply and demand direct the allocation of resources in an economy. The price mechanism consists of the rationing, incentive and signalling functions. The rationing function will determine who gets to buy non-meat alternatives. The signalling and incentive functions will determine how many units of non-meat alternatives are produced.

2/4 marks awarded This answer demonstrates the skill of AO1 (knowledge) without applying it to the extract. It is not enough to use the words 'non-meat alternatives'; instead, you must use the extract to support your economic knowledge.

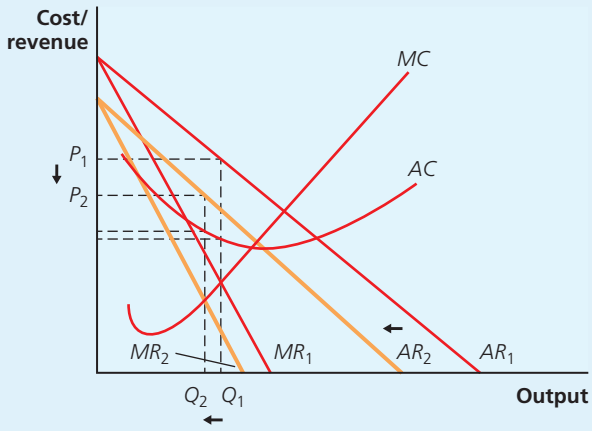
Question 3

Extract C states that '[UK] cattle farmers have lost nearly £170m in profit over 2019'. With the help of a cost and revenue diagram, explain why profits may have fallen for UK cattle farmers in 2019. 9 marks

Make sure that you include the correct type of diagram that the question asks for (in this case, demand should be falling). You won't be able to gain full marks for the question without including it.

Student A answer

Extract B says that more people are beginning to cut meat out of their diet in the UK. As a result, the price of UK beef is falling. This can be seen in Extract A where the price has fallen from 355p/kg in January 2019 to 335p/kg a year later. This demonstrates that the demand for meat-based foods, like burgers and sausage rolls, is falling. Since the demand for cattle is derived demand for burgers and sausage rolls, this is the reason for the 'collapse in demand' described in Extract C.



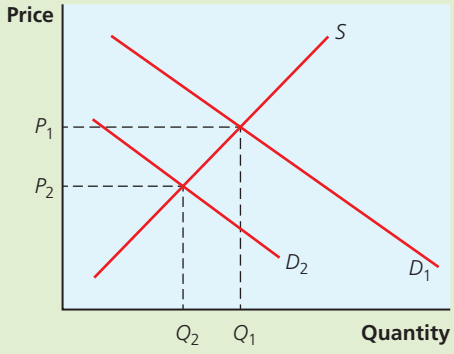
The diagram above shows the demand for cattle falling. The *AR* curve shifts to the left (AR_1 to AR_2). As a result, the *MR* curve also shifts left (MR_1 to MR_2) and the profit maximising level of output falls (Q_1 and Q_2). Although costs have not changed, the rectangle indicating supernormal profit has significantly decreased. This could be the lost '£170m in profit' described in Extract C. Also, it is clear that the new market price is much closer to the firms' average cost, which explains the quote: 'the current market prices barely cover our costs of production' from Extract C.

9/9 marks awarded The theory is well articulated and is correct. The diagram is well drawn and, crucially, is referred to throughout the answer; the simple use of brackets indicating the movements of lines when they are explained really helps. Quotes and figures from the extract are used to great effect throughout the answer, fully demonstrating the skill of application but also making sure the answer stays focused and actually answers the question.

Student B answer

Extract B says that more people are beginning to cut meat out of their diet in the UK. As a result, the price of UK beef is falling. This can be seen in Extract A where the price has fallen from 355p/kg in January 2019 to 335p/kg a year later. This demonstrates that the demand for meat-based foods, like burgers and sausage rolls, is falling.

When the demand curve decreases, this will lead to a decrease in the revenue gained by producers. Revenue is calculated by the formula *P* x *Q*. As the diagram shows, when demand falls the revenue box decreases in size (from P_1Q_1 to P_2Q_2), which indicates a fall in revenue.



4/9 marks awarded This answer starts well enough but it uses a simple demand and supply diagram rather than a cost and revenue diagram, which was specified in the question. In addition, the answer focuses on the issue of revenue rather than profit. Revenue and profit are different concepts and should not be used synonymously.

Question 4

Extract C states: 'The ability of trade unions to influence wages and levels of employment in the labour market is keenly debated in economics.' Use the extracts and your knowledge of economics, to assess the view that industry associations, like the National Farmers' Union, and other trade unions have a large influence on wages and the level of employment in the labour market.

25 marks

You should consider several factors (at least two) and focus on the impact on both wages and the level of employment in the labour market.

Student A answer

A trade union is an association of workers that acts as a bargaining unit in the determination of wages and working conditions. The success of a trade union, or its ability 'to influence wages and employment levels', depends on a multitude of factors.

First, a trade union's influence will depend on the percentage of workforce who are members of the trade union. Extract C states that the NFU boasts 55,000 farmer members. Although there are no figures which help determine the percentage of farmers who are trade union members, it can be assumed that as it is 'the largest farmers' organisation in the UK', it will wield a significant level of influence. In addition, farming is an industry which will undoubtedly have more economic leverage than, say, other professions such as checkout assistants in retail. Farmers are able to drive up the prices of food and limit the production of food in a country if they so wish. Recent farmers' strikes (over issues like milk prices) show that farmers are able and willing to use large machinery such as tractors to block road infrastructure, which can cause major disruption to an economy. Farmers can be extremely effective in driving up wages.

That said, the UK is currently in an economic climate where it is negotiating many new trade deals with other countries. Extract C implies that leaving the EU will improve UK farmers' wages but will not be the case if new trade deals (with non-EU countries) are formed with agricultural goods at the forefront. After Brexit, many countries will see the UK as a lucrative market for their agricultural produce. If the supply of imported agricultural goods increases then it follows that the demand for UK agricultural goods will fall and UK farmers' wages will fall, regardless of trade union membership. In this case, the NFU is asking for 'new subsidies to protect profits in the future'. The economic climate, as it is, does not lend itself to large public spending programmes on subsidies. The UK has a large national debt already and the government is committed to reducing the budget deficit. This is likely to reduce their influence.

Excellent application (AO2) to the extract, which is then developed to make appropriate points.

Excellent use of own knowledge to focus the answer on real-world events and context (AO2). This demonstrates the student's ability to fulfil the AO2 criteria to apply their economic knowledge in context. There is an expectation that, over 2 years of study, you will have built up a large amount of real-world subject matter knowledge and that you will be able to use this knowledge to support your analysis. From an examiner's perspective, use of own knowledge certainly helps to differentiate students.

Government legislation is also a significant factor affecting the influence of trade unions. There is no mention in the data of changing legislation but, since the 1970s, the strength of trade unions in the UK has diminished. The Conservative Party under Margaret Thatcher was synonymous with reducing trade union power and, in 2020, the UK now has a majority Conservative government again. If trade unions try to increase wage rates too much, the government might be concerned that firms might no longer be able to afford to employ workers or that the price of agricultural products becomes too high — this becoming uncompetitive. This would decrease the UK's ability to export and some might even argue that it causes real wage unemployment. Workers might prefer to earn low wages rather than be unemployed.

Trade unions are likely to have most success in a monopsony, where there is only one employer of specific skills. While farming skills are relatively narrow, there are many different firms in the market and so it is unlikely that the NFU will exert a large amount of power in this regard.

Overall, trade unions, like the National Farmers' Union, do not have a large influence on wages and the level of employment in the labour market. This will obviously depend on the percentage of workers who are in a trade union and whether they are a monopsony but, in most situations, they do not have a significant impact. In the case of the NFU, the government is likely to be more concerned with reducing the budget deficit, negotiating new trade deals with other countries and staying competitive internationally. The extract already states that the Prime Minister is avoiding their initial calls for subsidies. Trade unions wield little influence at the moment.

22/25 marks awarded

This answer demonstrates a good appreciation of the topic and links it to the extract and own knowledge at every stage (AO2). What is missing from the answer, however, is any analysis of market structure — is agriculture a monopoly or does it resemble perfect competition? The conclusion is not just a summary, but offers a judgement (AO4) which is supported by evidence from the essay.

Student B answer

A trade union is an association of workers that acts as a bargaining unit in the determination of wages and working conditions. The success of a trade union, or its ability 'to influence wages and employment levels' depends on a multitude of factors.

A trade union's influence will depend on the percentage of the workforce who are members of the trade union. If a trade union has more members it has more bargaining power. But when it has fewer members, it has less bargaining power. Therefore, we would expect to see that influence depends on size. This is because trade unions can threaten to go on strike if they do not succeed in achieving higher wages.

The comments made in this paragraph are too generic. There is no contextualisation of this analysis to the extract. Instead of talking about generic trade unions, you should try to talk about the specific trade union in the context. To make the same point, Student A wrote: '… the NFU boasts 55,000 farmer members… as it is "the largest farmers' organisation in the UK" it will wield a significant level of influence.'

In evaluation, it also depends on the current economic climate. If the economy is in a recession then the government may be more willing to offer subsidies because it might help to stimulate *AD*. If the economy is in an economic boom, then the government might be less willing to give out subsidies because the demand for workers is probably high anyway. And, as a result, their wages are probably rising too.

Government legislation is also a significant factor affecting the influence of trade unions. There are no data in the extract that suggests whether the government is imposing more or less legislation on trade unions. In Saudi Arabia, trade unions are illegal but in France they are powerful and there are lots of strikes. Extract C shows that the prime minister is ignoring the NFU at the moment.

Trade unions are likely to have most success in a monopsony, where there is only one employer of specific skills. While farming skills are relatively narrow, there are many different firms in the market and so it is unlikely that the NFU will exert a large amount of power in this regard.

Overall, the ability of trade unions depends on a number of factors. It depends on the number of members, the current economic climate, government legislation and whether the industry is a monopsony.

A missed opportunity to demonstrate own knowledge. 'If the economy is in a recession' could be improved by linking to the current economic situation at the time of writing and whether there is a recession or not. To make the same point, Student A referred to the current Brexit crisis ('the UK is currently in an economic climate where it is negotiating many new trade deals [with non-EU countries]'). Equally, the recession and job insecurity caused by the coronavirus pandemic could be alluded to.

14/25 marks awarded The answer identifies many of the key factors influencing the ability of trade unions' ability to influence wages and unemployment levels. However, at times there is a lack of application to the context (AO2) and a missed opportunity to demonstrate own knowledge. As a result, the answer lacks depth and feels too generic. Some of the examples are irrelevant (Saudi Arabia, France). Finally, the conclusion is weak. It offers a brief precis of the analysis and nothing else. There is no judgement supported by evidence.

Section B

Total for this essay: 40 marks

Essay 1

In a production possibility diagram, technological change is often given as a cause for the curve to shift outwards. 'Artificial Intelligence could potentially deliver additional economic output of around $13trn by 2030, boosting global GDP by about 1.2% per year' (headline from a report by McKinsey (September 2018).

Question 5

Explain what is meant by an outward shift of the production possibility curve AND explain why all points on the curve are considered to be productively efficient but not all points on the curve are allocatively efficient. You should use diagrams in your answer. 15 marks

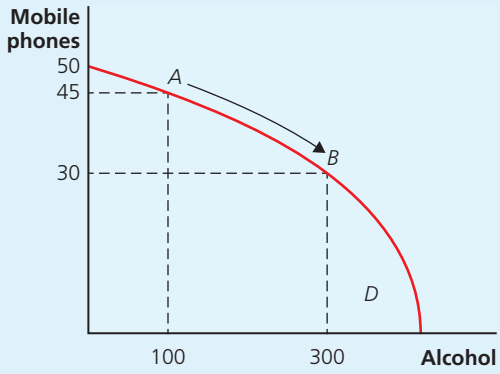
This is a long question with several instructions. Top tip: highlight all of the elements to ensure that you answer the full question and don't miss out parts.

Student A answer

A production possibility curve (*PPC*) is a line which shows the various combinations of two goods that an economy can produce when all of its resources are being fully utilised. The diagram below shows a *PPC* for a hypothetical economy that makes alcohol and mobile phones.

The *PPC* shifts outwards whenever there is any change to the resources of an economy that increases the maximum number of goods that can be produced. Technological change is often given as a cause for the curve to shift outwards because the productivity of capital goods is likely to increase. For example, machines may be able to make more mobile phones per hour than before. The outward shift is also shown on the diagram.

Productive efficiency is defined as the situation where it is not possible to produce any more of one good without sacrificing the production of another. Any point on the *PPC* is productively efficient because all resources are fully utilised at each of these points and so any reallocation of resources away from one industry (mobile phones) would lead to a loss in another industry (alcohol). This is shown in the diagram below in the movement from *A* to *B*.

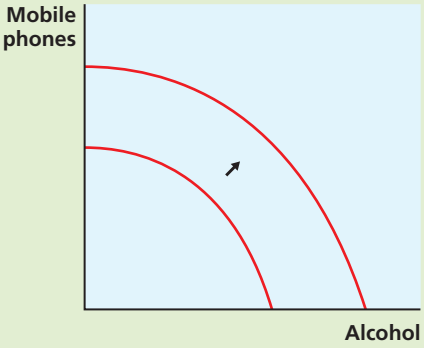


However, allocative efficiency occurs when community surplus is maximised. This can only be ascertained by a marginal analysis and the *PPC* does not provide any evidence of social benefits or social costs. For example, all resources could be allocated to alcohol — which imposes negative externalities onto the third party and clearly leads to welfare loss even though it is on the *PPC*.

13/15 marks awarded The student demonstrates a strong understanding of key concepts (A01) and develops each one to fully answer all criteria of the question (A03). The structure of the answer into three distinct paragraphs makes it easy for the examiner to identify each aspect of the question clearly. The diagram is drawn clearly and the answer uses its own examples to help support the analysis (A02). The student doesn't quite get full marks because the examples have not been fully explained.

Student B answer

A production possibility curve (*PPC*) is a line which shows the various combinations of two goods that an economy can produce when all of its resources are being fully utilised. The diagram below shows a *PPC* for a hypothetical economy that makes alcohol and mobile phones.



The *PPC* shifts outwards whenever there is any change to the resources of an economy that increases the maximum number of goods that can be produced. Technological change is often given as a cause for the curve to shift outwards because the productivity of capital goods is likely to increase. For example, machines may be able to make more mobile phones per hour than before. The outward shift is also shown on the diagram.

Allocative efficiency occurs when community surplus is maximised. Productive efficiency occurs at the point of production where the *AC* is at its lowest (the minimum point of the *AC* curve, where *AC* = *MC*).

The *PPC* is productively efficient because firms produce the most at the *PPC*. The *PPC* is not allocatively efficient because allocation is not determined by the *PPC* curve itself.

9/15 marks awarded The answer starts well: it offers a good understanding of the *PPC* and why technological change might shift the *PPC* outwards. Then there is a good definition of allocative efficiency and productive efficiency (AO1 only!) but the final paragraph demonstrates that the student cannot explain why all points on the curve are considered to be productively efficient but not all points on the curve are allocatively efficient. The depth of understanding simply isn't there and so the answer cannot receive top band analysis (AO3) marks.

Question 6

Evaluate the view that rapid technological change, such as the increased use of artificial intelligence (AI) in industry, will always lead to positive outcomes for firms. 25 marks

Do not be put off by open-ended questions like 'positive outcomes' or 'benefits' for economic agents. The question is intentionally open in order to give you greater freedom for your answer. As long as you understand the common aims/objectives of economic agents (be it producer, consumer or government) and structure your answer accordingly, your answer will be focused.

Student A answer

Firms may have a lot to gain from artificial intelligence (AI). For example, firms in manufacturing industries which employ workers to perform repetitive skills may be able to substitute AI for human labour. As a result, the costs of production are likely to fall. Artificial intelligence is a one-off cost rather than an ongoing variable cost and it does not attract employment taxes or pension contributions. In manufacturing industries where workers are often specialised to perform particular tasks on a production line, absences can delay production. Artificial Intelligence does not get ill, nor does it require absences. All of this means that a firm can lower its average cost curve and increase its profits.

However, it depends on the industry. In industries which require greater emotional intelligence and communicative skills such as patient care in the health sector, AI cannot (yet) replicate these skills. Artificial intelligence and human labour are not easily substitutable.

Similarly, AI may actually increase costs for a firm if it requires constant maintenance. Workers who repair, upgrade and fix AI are highly skilled and will demand high wages as a result. Ultimately, this may counter (or even completely outweigh) the cost savings made from losing unskilled staff.

Artificial intelligence may be beneficial for price discrimination. In markets with appropriate conditions (goods cannot be resold, different consumer groups have different elasticities), AI can help to identify different market segments more effectively and change prices more dynamically. Firms like Amazon and Facebook already use AI a great deal when analysing big data, which then helps other firms to advertise effectively. It is not uncommon for two consumers to be charged two different prices on Amazon according to their previous browsing history. This knowledge helps firms to maximise revenue.

Artificial intelligence also has the potential to improve the quality of a firm's products. New markets may emerge (such as apps that recommend certain cosmetic products based on digital only screening) or old products can be redesigned with AI in mind (for example, books which are now given augmented reality add-ons). This can increase the demand for a firm's products, increasing its revenue.

Overall, AI has the potential to lead to positive outcomes for firms. This might be improved quality, a wider range of products, more targeted advertising (and price discrimination) as well as the ability to substitute labour for AI-based solutions. In every scenario, costs are either decreasing or revenue is increasing: profits will rise. The only downside, for the moment, is cost — both the short-term purchasing of new AI and also the long-term maintenance. These costs may be expensive now but they are likely to come down in the future. Firms will undoubtedly benefit from AI in the future.

The comment 'it depends on the industry' is an excellent evaluative point (AO4) that helps to qualify the analysis in the preceding paragraph. However, at just three sentences it is slightly undeveloped, and ultimately doesn't hit top band level.

Excellent analysis (AO3) and application (AO2) combining here. The 'price discrimination' point is valid and well supported with examples.

21/25 marks awarded This is a strong answer which shows an excellent awareness of AI in the real world. There are two shortfalls: first, diagrammatic analysis would help to explain many of the points made about costs, revenue and profit. Second, the essay would benefit from more evaluation about the impact on different types of industry. It is mentioned in paragraph 2 but it is not referred to in the conclusion.

Student B answer

Firms may have a lot to gain from Artificial intelligence (AI). For example, firms in manufacturing industries which employ workers to perform repetitive skills may be able to substitute AI for human labour. As a result, the costs of production are likely to fall. Artificial intelligence is a one-off cost rather than an ongoing variable cost and it does not attract employment taxes or pension contributions. In manufacturing industries where workers are often specialised to perform particular tasks on a production line, absences can delay production. Artificial intelligence does not get ill, nor does it require absences. All of this means that a firm can lower its average cost curve and increase its profits.

Artificial intelligence may be beneficial for price discrimination. In markets with appropriate conditions (goods cannot be resold, different consumer groups have different elasticities), AI can help to identify different market segments more effectively and change prices more dynamically. Firms like Amazon and Facebook already use AI a great deal when analysing big data, which then helps other firms to advertise effectively. It is not uncommon for two consumers to be charged two different prices on Amazon according to their previous browsing history.

Artificial intelligence also has the potential to improve the quality of a firm's products. New markets may emerge (such as apps that recommend certain cosmetic products based on digital only screening) or old products can be redesigned with AI in mind (for example, books which are now given augmented reality add-ons). This can increase the demand for a firm's products, increasing its revenue.

Overall, AI has the potential to lead to positive outcomes for firms. This might be improved quality, a wider range of products, more targeted advertising (and price discrimination) as well as the ability to substitute labour for AI-based solutions. In every scenario, costs are either decreasing or revenue is increasing: profits will rise.

While it is important that you do not draw diagrams just for the sake of it, they can help to make your points more robust. There are two opportunities to draw diagrams here: first, a cost/revenue diagram and, second, a pair of price discrimination diagrams. It is a great shame to spend so long learning the diagrams in class and then not to use them in your actual exam.

Top tip: the opening sentence of your concluding paragraph should be a judgement that directly answers the question. The next couple of sentences should then review your strongest lines of analysis which support this judgement.

16/25 marks awarded The student considers three paths of analysis, each one of which is fully developed with deep analysis (AO3) and application (AO2). The student uses real-life examples to support their points throughout. Unfortunately, the answer is totally one-sided and offers no evaluation (AO4) at all. Evaluation does not have to be a complete counter-point (i.e. explaining why the opposite is true) but it must try to qualify those arguments. Evaluations such as time periods, different points of view, size of impact are all encouraged. This answer does not demonstrate AO4 at all, which also limits the amount of credit that the conclusion can be given.

Paper 3

Section A

Multiple-choice questions

The correct answers and brief explanatory notes follow at the end of the multiple-choice questions.

1 Good Z has a price elasticity of supply of 0. The demand curve shifts to the left. What will happen to the equilibrium price and quantity as a result?

	Price	Quantity
A	Increase	Increase
B	Unchanged	Decrease
C	Decrease	Unchanged
D	Decrease	Decrease

Try drawing the diagram on a scrap piece of paper before answering the question. The key is remembering how to draw a supply curve with *PES* = 0.

2 Hit and run competition occurs when a firm temporarily enters a market and then leaves because …

A supernormal profits are exhausted.
B it faces predatory pricing from incumbent firms.
C of creative destruction.
D of government regulation.

Read each of the possible answers carefully, as a full sentence. Eliminate the ones which are obviously false.

3 Starbucks coffee and Costa coffee are strong substitutes. Starbucks chooses to lower the price of its coffee.

Identify the correct impact on the Starbucks market and the Costa market, *all other things being equal*.

	Market for Starbucks coffee	Market for Costa coffee
A	Contraction along the demand curve.	Shift left of the demand curve.
B	Shift right of the demand curve.	Shift right of the demand curve.
C	Extension along the demand curve.	Shift left of the demand curve.
D	Shift right of the demand curve.	Contraction along the demand curve.

Consider drawing two demand curve diagrams (one for Starbucks and one for Costa) and then lower the price on the Starbucks one to find out what happens.

4 The total utility gained from eating several biscuits is shown in the table below. What would be the marginal utility of eating the fourth biscuit?

Biscuits eaten	Total utility
1	10
2	18
3	24
4	28
5	30

A 110
B 80
C 28
D 4

Even though this is a calculation question, it is testing your knowledge of the concepts of marginal utility and total utility rather than anything particularly numerical.

5 What is the four-firm concentration ratio for the market (UK publishers, 2014) below:

Firm	Market share
Penguin Random House	23.4
Hachette Livre	12.9
News Corporation	7.5
Holtzbrinck	3.7
Bloomsbury	2.3
Simon & Schuster	2.2
Others	48

A 95.5
B 91.8
C 47.5
D 3.7

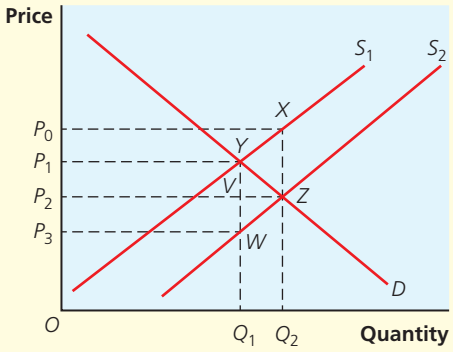
A common mistake would be to include 'Others' into the calculation. 'Others' is not the name of a firm.

6 Which one of the following is NOT a characteristic of perfect competition?

A large numbers of producers
B identical products
C perfect knowledge
D high barriers to entry and exit

There are mnemonics you can learn to remember groups of characteristics like this, such as FHIB (**F**irms, **H**eterogenous, **I**nfo, **B**arriers).

7 A firm receives a subsidy from the government. Which of the following areas represents the total government spending on the subsidy?



A P_1YWP_3
B P_0XZP_2
C P_0XQ_20
D P_1YVP_2

Many students confuse indirect tax diagrams with subsidy diagrams. This is a subsidy diagram.

8 A firm sells 1,500 units.
The marginal cost of the 1,500th is £5.50.
Its average revenue is £13.75.
Why is it not possible to calculate the firm's total profit?

A Data on average cost are required.
B Data on selling price are required.
C Data on the opportunity cost of production are required.
D All of the above.

Think about what values you would need to calculate profit and try to identify which one is missing from the list given.

9 The income elasticity of demand (YED) for coffee is 0.25. If George's monthly wage rises from £2,000 to £2,400 after a promotion at work, his consumption of coffee is likely to …

A increase by 2.5%.
B increase by 5%.
C increase by 10%.
D increase by 80%.

Write out the YED formula and rearrange.

10 A firm is dynamically efficient when …

A a firm reduces its cost curves by implementing new production processes.
B all of the firm's costs are directly related to the production process.
C it is impossible to reallocate resources so as to make one person better off without making someone else worse off.
D community surplus is maximised.

All of these definitions are different types of efficiency, but only one is the correct definition of dynamic efficiency.

Answers to multiple-choice questions

Question 1

A PES of zero implies that the supply curve is perfectly inelastic and is drawn as a straight vertical line. If the demand curve shifts inwards, the new equilibrium price has fallen but the quantity demanded remains unchanged. The answer is **C**.

Question 2

Hit and run competition occurs when a firm temporarily enters a market and then leaves because supernormal profits are exhausted. A good example of this is in the toy industry where firms quickly bring out the latest trend (slime, fidget spinners) and then leave as soon as the novelty — and their stock — runs out. The answer is **A**.

Question 3

If Starbucks lowers its prices, then there is a movement along the demand curve (known as an extension). The quantity demanded for Starbucks increases. As a result, the demand for Costa coffee decreases, which means the demand curve shifts to the left. The answer is **C**.

Question 4

Marginal utility is the utility gained from each additional unit. To find the marginal utility of the fourth biscuit you need to subtract the total utility of eating four biscuits by the total utility of eating three biscuits. This gives the marginal utility of that fourth biscuit. 28 – 24 = 4. The answer is **D**.

Question 5

A four-firm concentration ratio is calculated by adding up the market shares of the four largest firms. In this case, 23.4 + 12.9 + 7.5 + 3.7 = 47.5. The answer is **C**. Remember that 'Others' is not the name of a firm but a collective noun for all of the other firms in the market with a market share less than 2.2%.

Question 6

Perfect competition has low, or non-existent, barriers to entry and exit. The answer is **D**. The options A–C are all characteristics of perfect competition.

Question 7

The imposition of a subsidy causes the supply curve to shift left to S_2. The vertical straight line distance between S_1 and S_2 is the value of the subsidy (XZ). The new equilibrium is at Q_2 and all of those units receive a government subsidy. Therefore the government spending rectangle is P_0XZP_2. The answer is **B**.

Question 8

Total profit is calculated by the formula: total revenue – total costs. It is possible to calculate total revenue because you have *Q* and *AR*. However, you cannot calculate total costs because you do not have an *AC* figure. The answer is **A**.

Question 9

Income has risen by 20% because ((2,400 – 2,000) / 2,000) x 100 = 20%. If YED = 0.25 and the percentage change in income is 20%, then the percentage change in demand must be 5%. The answer is B. A common mistake here is get the YED equation the wrong way round — which, if you did it, would give you answer **D**.

Question 10

A firm is dynamically efficient when it reduces its cost curves by implementing new production processes. The answer is **A**. B is x-efficiency, C is productive efficiency and D is allocative efficiency.

Section B

Investigation

Total for investigation: 50 marks

Scenario

You are an economist working for an independent think-tank which has been employed by the Italian government to produce a report on tourism in Venice. As part of this investigation, you are required to answer three questions.

Study Extracts A, B, C and D, and then use these and your own economic knowledge to help you answer questions 1, 2 and 3.

Extract A Mayor of Venice blames climate change for flooding in Venice

In November 2019, severe flooding in Venice left much of the Italian city under water. The waters peaked at 1.87m (6ft), according to the tide monitoring centre. Only once since official records began in 1923 has the tide been higher, reaching 1.94m in 1966.

The mayor of Venice, Luigi Brugnaro, was quick to blame climate change: 'These are the effects of climate change... the costs will be high,' he said. A project to protect the city from flooding (known as the MOSE project) has been under way since 2003 but it has been hit by soaring costs, scandals and delays.
The mayor urged the government to intervene and provide the funds to finish the project. 'Now the government must listen… The situation is dramatic. We ask the government to help us,' he said on Twitter.
Recent news reports have estimated the final cost of the MOSE project at around €5.5 billion. This is equivalent to €90 per person in Italy (population 60 million).

Source: news reports (2019)

Extract B Love-hate relationship between tourists and residents in Venice

Famed for its classical architecture and lagoon waters, the city of Venice, in Italy, is one of the most visited cities in Europe. In 2019, it welcomed 5.5 million tourists. The city, however, has a love–hate relationship with its tourists.

Ultimately, the rapid growth of low-cost aviation, cruise ships and peer-to-peer home sharing platforms (like Airbnb) have all led to the increase in tourism. It is an industry that creates over 100,000 jobs but, on any given day, its 260,000 local residents are forced to negotiate crowds and put up with noisy wheelie suitcases, selfie sticks and disrespectful behaviour. They see the city being littered and its fragile infrastructure being damaged, while watching vast cruise ships chug up and down the canal system all day long. The rise of the day tripper is a huge problem too. Only half of the 5.5 million tourists actually sleep in the city; many just pour off a cruise ship on a whirlwind tour of Italy. They stay for just a few hours, see little, buy a few trinkets and leave. They bring no economic benefit to the city in this way.

Once upon a time, local residents would swim in the lagoon waters. But the onset of mass tourism and an increase in boat traffic in recent decades has made swimming a rarity. The water is now so polluted from diesel oil (from ships) and general litter that it is unsuitable for such a pastime anymore.

The love–hate relationship is no more evident than with the dual-pricing system that has been established there. Tourists and local residents are often quoted different prices for the same good or service. For example, a recent news report identified price disparities in what tourists are charged compared to what locals are charged for the same goods. This is shown in Extract C, Figure 1.

Source: news reports (2019)

Extract C Data on tourism in Venice

Good/Service	Tourist price	Local resident price
Entry to Doge's Palace (tourist attraction)	€18	Free
Waterbus ride	€7	€1.30
Public Wi-Fi	€5 a day	Free
Public bathroom	€1.50	€0.25

Figure 1 The price of selected goods and services in the city of Venice

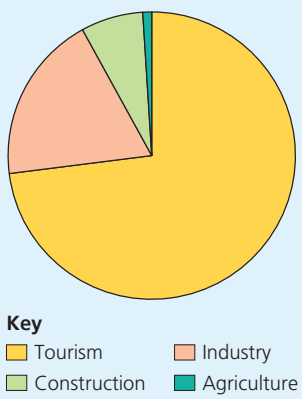


Figure 2 Venice's economy by sector

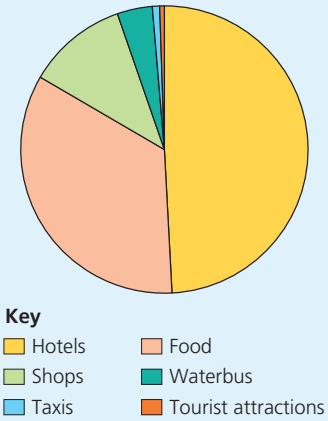


Figure 3 Percentage of tourist spending in Venice

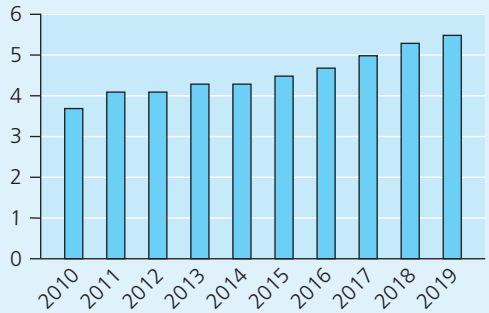


Figure 4 Number of tourists to Venice per year (2010–19), in millions

Extract D

In its latest bid to tackle over-tourism, Venice is considering a new tax on tourists. The tax would see all tourists pay a fee for every night that they stay in a hotel in Venice or, in the case of cruise ship passengers, an entry fee to the city as they disembark. The proposed tax could be quite popular: local residents have frequently protested against a tourism industry which they argue has eroded their quality of life, damaged the environment and driven residents away. The mayor, Luigi Brugnaro, has already stated that some of the extra cash from the fee will help fund the cleaning up of rubbish that tourists leave behind.

It is already the case that people cannot drive into Venice. People can arrive in the mainland area of the lagoon and park their vehicle for €29 a day. This was brought in to reduce traffic congestion and air pollution in the area. This has been a successful policy although one of the unintended consequences has been that returning family members have also been hit by the tax, angering local residents.

Source: news reports (2019)

Question 1

To what extent, if at all, do the data suggest that price discrimination in Venice is beneficial to consumers and producers. *You must use the data in Extract C to support your assessment*.

10 marks

You must look at both consumers and producers. In addition, the 'to what extent' command words imply that a final judgement is required from your essay.

Student A answer

The price discrimination described in Figure 1 of Extract C is very obviously undesirable for tourist consumers. The extract shows that tourists are being charged much more than local residents for the same products such as waterbus rides and public bathrooms. This means that tourist consumers will have much less consumer surplus than local residents when consuming the same goods. However, Figure 3 tourists are only spending nearly 5% of their expenditure on the waterbus and 1–2% on tourist attractions like the Doge's Palace.

An argument can be made that such price discrimination is beneficial for the local residents of Venice: for example, the profits gained from charging tourists a high price could be used to cross-subsidise lower prices for local residents. Furthermore, most of the goods and services that it applies to are unlikely to be regular purchases by local residents (such as entry to tourist attractions, public bathrooms etc.).

From the firm's perspective, it is highly beneficial to be able to charge tourist consumers a higher price. Figure 2 shows that about 73% of Venice's economy is based on the tourist sector. Therefore, tourists are the main source of revenue for many firms.

In addition, tourists are usually much less sensitive to price increases because of the fact that they have a limited time in the area. As a result, firms can charge a very high price and make a lot of revenue in a short period of time. The only drawback is if they exploit the situation too much and it attracts the attention of regulators, who may stop the practice in the future. That doesn't seem to have bothered tourists to Venice very much so far though. Tourist numbers have increased almost every year from 2010 (3.7 million) to 2019 (5.5 million). Tourists keep coming back and do not seem to be overly bothered by this overt price discrimination.

Therefore, price discrimination in Venice is seen to be beneficial for producers but not for consumers.

9/10 marks awarded There is no definition of price discrimination but there is implied understanding throughout this answer. The answer considers both consumers and producers in turn — and even separates tourist consumers from local consumers, which shows strong AO4 evaluative skills. The answer would be improved by a diagram even though the question doesn't specifically ask for one. For example, the argument 'charging a very high price and making a lot of revenue' would be greatly improved by accompanying diagrammatic analysis. The final line answers the question directly (good) but it feels a bit throwaway (bad) and should have been linked into the preceding paragraph for more impact.

Student B answer

Extract C states that tourists are being charged much more than local residents for the same products such as waterbus rides and public bathrooms. For example, tourists are paying €18 to go to the Doge's Palace but local residents do not pay anything. Tourists pay €7 for a waterbus ride, local residents pay just €1.30. Public Wi-Fi costs tourists €5 a day, it is free for local residents. Tourists pay €1.5 but local residents pay just €0.25.

Price discrimination occurs when firms identify different market segments; in this case, tourists and local residents. It must also be true that the elasticities of demand are different for both set of consumers. One would imagine that PED is inelastic for tourists, who are likely to visit Venice just for a few days (maybe even hours if they are on a cruise ship) and therefore they are less sensitive to price changes. Figure 4 shows that tourist numbers are increasing so they must not be too bothered. It must also be true that there is no resale market: that local residents cannot resell the products to tourists. This is clearly true for things like the entrance to tourist sites such as the Doge's Palace. Once you have paid to enter, you cannot then resell the ticket at the turnstile.

Simply quoting figures without using them to make an analytical point might gain low application (AO2) marks but not top band. A diagram showing the impact on consumer surplus for different market segments would improve this answer.

In the second paragraph, the conditions for price discrimination are explained but this has little relevance to the question about 'whether or not it is beneficial to consumers and producers'.

2/10 marks awarded This response never really answers the question. There is a good understanding of price discrimination and a description of why it is so prevalent in Venice. However, the student overuses examples in the first paragraph and fails to make an analytical point with them. The second, third and fourth examples add no more support to the point than the first example — and the only point being made here is that price discrimination exists! Only AO1 marks can be awarded for this.

Question 2

Explain why public goods, such as flood defence systems, cannot be provided by the free market.

15 marks

A common error might be to confuse 'public goods' with goods that are provided by the government. *Public goods* have a specific meaning in economics and you should focus on the theoretical aspects such as non-rivalry and non-excludability.

Student A answer

Public goods are goods that have two main characteristics: they are non-excludable and non-rival. Non-excludability means that once a good is provided, anyone can then access that good for free. Non-rivalry means that one person's consumption of a good does not stop or reduce another person's consumption of the good.

Flood defence systems like the one described in Extract A are a good example of public goods. Venice is vulnerable to flooding and so 'a project to protect the city from flooding (known as the MOSE project) has been under way since 2003'. Once the flood defence system is built (by the free market or not), households inside the defence system would still benefit from it even if they had not contributed to it. Similarly, the fact that any one household is benefiting from the defence system at any given moment does not stop or reduce another person's consumption of the defence system. Indeed, all households within the defence system are continually benefiting from it all the time, simultaneously.

In economic theory, public goods are not provided by the free market. This is because of the free-rider problem. Because households will be able to gain utility from the defence system without paying, there is no incentive no pay. If there is no incentive to pay, it follows that there will be no profit motive for a private firm to produce the defence system. The defence system will not be produced and there is a missing market. This is a market failure.

The nature of flood defence systems, and public goods in general, means that they cannot easily be made excludable. Therefore, the obvious solution to the market failure is for the government to provide public goods using tax revenue. Extract A suggests that 'the final cost of the MOSE project will be at around €5.5 billion, which is equivalent to €90 per person in Italy (population 60 million)'. Private business will not be able to afford this without a profit motive and so the government must step in.

This is a good example of AO1 (knowledge). It is not a requirement to define key terms at the beginning of an answer but it can be a useful approach because it helps to frame the rest of the answer.

This is a good example of AO3 (analysis). The understanding that is demonstrated in the first paragraph is developed here, using technical economic vocabulary, to explain why flood defences cannot be provided by the free market.

15/15 marks awarded The response neatly defines all of the key terms early on and then applies the theory to the context of flood defence systems. The analysis picks up marks in the top band by then developing the chain of analysis to thoroughly answer the question about free-market provision. The last paragraph is a fitting end but it is not necessary for this question; the answer has already scored full marks by the end of paragraph 3.

Student B answer

Public goods are goods that are provided by the government. They are most often found in mixed economies where there is an element of government intervention and resources are not always left to the free market.

Flood defences are very expensive for any one individual or one firm to produce. For a city the size of Venice, the cost of the MOSE project (in Extract A) is already said to be in the billions of euros. A government might be rich enough to provide this but not a private institution.

The free market is made up of consumers and producers. Prices are determined by the market forces of supply and demand. Where supply = demand, the price of a product is set. The demand for a flood defence system in Venice will be very high and so this will push the price very high. Also the number of potential suppliers will be very small so this means that the supply curve shifts to the left. The price, then, is incredibly high and this means that very few producers can afford to buy it.

In a mixed economy, people are taxed by the government through a range of taxes. The tax revenue that is received will be used by the government to provide flood defences.

Flood defences have positive externalities because they confer a positive spillover effect on the third party. Without government intervention, flood defences will be under-allocated by the free market and this is a market failure.

7/15 marks awarded The first sentence is wrong. The student has forgotten what a public good is and, therefore, the answer reflects their attempt to score points by writing lots. To some extent, this is a useful strategy. The worst thing you can do in an exam is to leave the answer box blank. This automatically scores 0 marks and will heavily affect your grade. This student picks up marks for their understanding of free markets, compulsory tax schemes (AO1) and some decent application to the city of Venice (AO2). The AO3 marks are low because, ultimately, there is no link to the idea of *public goods* and the associated concepts. But, there is some logic in the argument given (prices are too high!) so AO3 isn't a complete wipeout. The answer still scores 7/15 (nearly 50%), which is not bad if you didn't really understand the key concept of the question.

Question 3

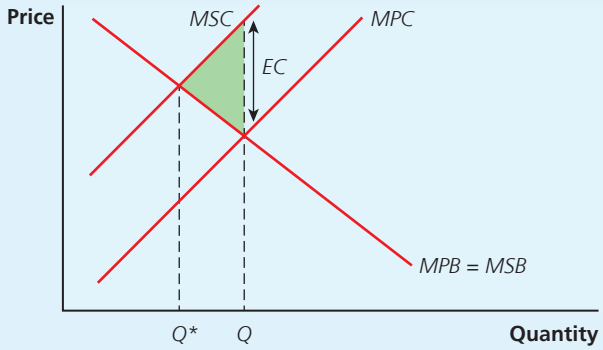
Taking into account all of the evidence from Extracts A, B, C and D, would you recommend that the city of Venice introduces tourist taxes? ***Justify your recommendation.***

25 marks

It is important to relate your answer to the city of Venice. It is also important to use data from all three extracts.

Student A answer

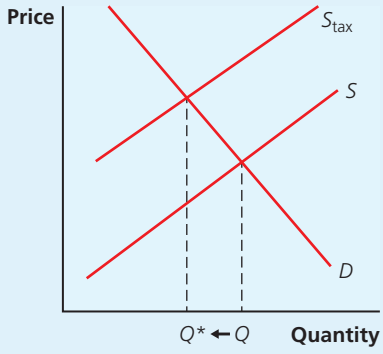
Extract B outlines the negative externalities associated with tourism in the city of Venice. A negative externality is a negative spillover effect of a private transaction on a third party (neither the consumer, nor the producer). For example, tourists are a cause of litter, noise pollution and traffic congestion. While they spend money in Venice, they are a great nuisance to local households. The diagram below shows that negative externalities can lead to market failure.



In the diagram above, the marginal social costs (*MSC*) are greater than the marginal private costs (*MPC*) due to the presence of the negative externality or external cost (*EC*). As a result, the amount of tourism generated is above the socially optimum level (Q^*) and a welfare loss is created (shown by the shaded triangle).

Best practice is to always refer to your diagrams. Diagrams without any reference will normally not gain any credit at all.

An indirect tax imposed on tourists would shift the *MPC* curve to the left, so that a new equilibrium would be created at the socially optimum level, thereby fixing the market failure. This is shown in the diagram below. The quantity demanded has fallen back to Q^* and the market failure is solved.



However, tourism is a key industry for Venice. Figure 4 in Extract C suggests that there were 5.5 million tourists in 2019, which would have generated a lot of economic growth for the city. By imposing a tourist tax, the number of tourists will fall (shown in the second diagram Q to Q^*) and, ultimately, this will lead to less profit for firms and less income for workers.

One of the major advantages to the tax scheme is that the government will gain a lot of revenue. Extract A suggests that the price elasticity of demand for tourism in Venice is inelastic (tourists continue to visit Venice despite the high prices). Thus, the amount of tax revenue will be very high indeed and the government is likely to have more money to spend on improving living standards within Venice. It may be able to fund a higher quality flood defence system (which Extract A states 'needs upgrading'), an underground rail network or other schemes like restoration projects which may actually reduce the negative externalities of tourism.

Excellent use of application (A02). A generic point about tax revenue is expertly analysed within the context of Venice and what it is likely to mean in this context. The paragraph could be improved further with a tax revenue diagram.

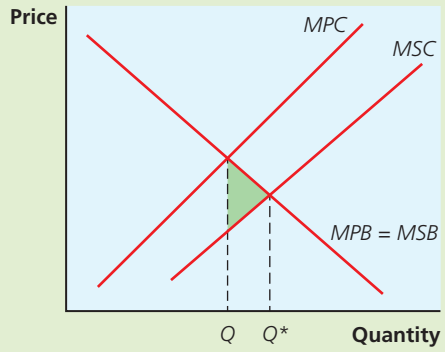
Sadly, tax schemes are not always effective. It is difficult to estimate the correct amount of tax to set. Too low and it won't solve the market failure; too high and it may lead to even greater welfare loss. It is also undesirable from the point of view of local households who have family elsewhere in Italy. These are not tourists but they will incur the tax when returning.

Overall, the imposition of a tourist tax is advantageous and it is highly recommended that Venice introduces a tourist tax scheme. The tax will help to reduce the welfare loss created by negative externalities of tourism such as noise pollution and damage to buildings, and it helps to generate tax revenue for the government which can be used for useful projects such as the MOSE flooding project. One potential downfall is that the government sets it so high that it deters tourism and economic growth in Venice. But, given that tourism is price inelastic, tourists will not be put off greatly. There are also alternative strategies such as tourist quotas and bureaucratic measures (for example, requiring tourists to get visas) but, again, these would impact the economy too severely. To that end, the imposition of a tourist seems like the most effective strategy and it is highly recommended.

20/25 marks awarded This is a very strong answer. AQA does not specify a particular structure for essay questions and you can obtain full marks in a number of different ways. However, you must be able to show A01, A02, A03 and A04 skills in your essay. The advantage of this essay structure (point, evaluation, point, evaluation, conclusion) is that the examiner can identify the separate chains of analysis and evaluation easily when marking. That said, the evaluation of this student's response does not achieve top band. The conclusion is more of a summary than a judgement and both of the evaluative paragraphs are relatively short, too generic and do not hit the same analytical standard of the analysis paragraphs. In short, to gain the highest marks the recommendation needs to be stronger, and the conclusion more substantial.

Student B answer

Figure 4 in Extract C shows that Venice had 5.5 million tourists in 2019. The extract then outlines the negative externalities associated with tourism in the city of Venice. A negative externality is a negative spillover effect of a private transaction on the third party (neither the consumer, nor the producer). For example, tourists are a cause of litter, noise pollution and traffic congestion. While they spend money in Venice, they are a great nuisance to local households. The diagram below shows that negative externalities can lead to market failure.



The diagram is incorrect as it shows *MSC* lower than *MPC*. The welfare loss, as a result, is shaded incorrectly. The socially optimum level (*Q**) is also in the incorrect position due to the *MPC* and *MSC* being drawn the wrong way round.

Social costs are higher than private costs, so the supply curve shifts to the left to illustrate this. This means that there are too many tourists (more than the socially optimum level) coming into Venice. As a result, there is deadweight welfare loss.

Taxes can be used to reduce the demand for tourism. Tourist tax is a direct tax, which means it cannot be passed on. It is a form of fiscal policy.

The main advantage of the tax is that it will reduce demand and so the welfare loss is reduced. On the other hand, tourism is a key industry for Venice. Extract C suggests that there were 5.5 million tourists in 2019, which would have generated a lot of economic growth for the city. By imposing a tourist tax, the number of tourists will fall and, ultimately, this will lead to less profit for firms and less income for local residents who work there.

14/25 awarded There are lots of positives with this essay but it doesn't quite hang together. The diagram and the attempt at application are weak because the same example is used repeatedly. The conclusion brings in new information that was not discussed earlier in the essay. It cannot gain top band (AO4) because it is now using supporting evidence that was not fully analysed earlier. Overall, the essay is focused and there is a good attempt at AO2, AO3 and AO4 but it fails to hit top band for any of them due to many inaccuracies, lack of application and underdeveloped evaluation.

On the other hand, another advantage to the tax scheme is that the government will gain a lot of revenue. Extract C says that 5.5 million people visited Venice last year — if each of them paid €29 a day, this would be €159.5m euros that the government could spend on flood defences or employing cleaners to help clear up the litter that the tourists are dropping.

Overall, I think the taxes will be useful because they reduce the amount of tourists in Venice. My diagram shows that too many (5.5 million) tourists visit Venice and so there is a welfare loss. Reducing the number of tourists is the most important objective at this stage. However, the government must ensure that the tax is set at a correct level (not too high and not too low) otherwise it may not affect tourists that much. For example, tourists are not very sensitive to price changes so even a large tax may not reduce demand too much. Assuming it was set at the same level as the externality, however, it would solve the market failure.

This is the fourth time the same example is used, which does not demonstrate top band application (AO2).

This brings in new information that was not discussed earlier in the essay.

Knowledge check answers

1 Scarcity means there is not enough of a good to meet demand. Economising means limiting the amount of a scarce good that you will buy and consume so as to be able to buy and consume other goods as well.

2 When economists refer to the 'opportunity cost' of a resource, they mean the value of the next-highest-valued alternative use of that resource. If, for example, you spend time and money going to watch a football match, you cannot spend that time at home reading a book, and you can't spend the money on another activity.

3 Total utility is the sum total of utility or welfare derived from the consumption of all the units of a commodity. Marginal utility refers to additional utility obtained from the consumption of an additional unit of a commodity.

4 Asymmetric information, also called 'information failure', is a situation in which one party in a transaction has more or superior information compared with another. Symmetric information is when the same information is available to all parties in the transaction.

5 Different answers possible.

6 The behaviour of economics students is likely to be affected by, for example, the way the teacher frames how the results of a weekly test are presented. Today the teacher says that 40% of the class will pass next week's test and be rewarded with a merit grade. A week later he says 60% of students will fail in the following week's test and not achieve a grade. The first response creates a greater incentive for the students to work hard than does the second response.

7 **An adjustment of demand:** if the income of the consumer, the prices of the related goods or services and the preferences of the consumer remain unchanged, then the change in quantity of a good or service demanded by the consumer will be negatively correlated to the change in the price of the good or service. The change in price will be reflected as a move along the demand curve.

A shift of demand: the demand curve will shift — move either inwards or outwards — as a result of non-price factors. A shift in demand can be related to the following factors: consumer preferences; consumer income; change in the price of related goods or services (compliments and substitutes); change in the number of buyers; consumer expectations.

8 Many people treat table salt as a necessity that must be used in cooking and for scattering on food. There are no close substitutes for table salt, so demand for table salt (as a generic product) is quite price inelastic. However, another brand of table salt, say the ASDA version, is an almost perfect substitute for Tesco's table salt. If Tesco increases the price of Tesco salt but ASDA leaves its price unchanged, some consumers may switch to the cheaper substitute. Demand for Tesco's salt is therefore more price elastic than demand for table salt as a generic product.

9 The plus sign (+) tells us that the good is a normal good, i.e. demand for the good increases as income rises. The absolute size of the elasticity statistic (2.3) tells us that demand is income elastic. With this statistic, a 10% increase in income induces a 23% increase in demand. The good is a superior good as well as a normal good. By contrast, an inferior good has a negative income elasticity of demand.

10 The minus sign (–) tells us that good A and good B are complementary goods or goods in joint demand. A decrease in the price of good B causes consumers to buy more good B and also increases their demand for the complementary good, good A. The absolute size of the elasticity statistic (0.8) tells us the strength of the joint demand relationship. A 10% decrease in the price of good B leads to an 8% increase in the demand for good A. This is quite a strong joint demand relationship.

11 Equilibrium is a state of rest or a state of balance between opposing forces. A market is in equilibrium when the amount consumers wish to buy exactly equals the amount producers wish to sell.

12 Water is in composite demand since among its many uses are drinking, cooling, heating, cooking and washing.

13 Diminishing marginal returns (diminishing marginal productivity) occur in the short run when at least one factor of production is fixed. Decreasing returns to scale occur in the long run when all factors of production are variable.

14 The rent paid for the lease of business premises is a fixed cost, as is the annual business rate paid to government. The hourly wage rate paid to casual labour is a variable cost, as is the electricity bill paid for the use of electricity for heating and lighting.

15 Both the *AVC* and the *ATC* fall when *MC* is below the curves. As soon as *MC* rises above the curves, the curves start to rise. Both the *AVC* and the *ATC* curves are U-shaped.

16 A rise in wage costs causes a firm's costs of production to increase, and, as labour becomes relatively more expensive than capital, firms are likely to employ less labour and more capital, adopting more capital-intensive methods of production.

17 Marginal revenue is the extra sales revenue a firm receives from selling one more unit of a good or service. Marginal returns refers to the extra output that an extra worker contributes to total output.

18 If $MR < AR$, the *AR* curve must fall.

19 Agricultural markets, stock markets and foreign exchange markets may approximate to perfect competition.

20 Because no real-world market displays simultaneously all the six conditions of perfect competition.

21 The concentration ratio for a pure monopoly is 100 or 100%.

22 Competitive oligopolists may be adversely affected by uncertainty as to how rival firms will react to this uncertainty. Such collusion may enable the participating firms to make larger monopoly profits.

23 No; differences can also be explained by differences in costs of production.

24 Monopoly may lead to market failure as a result of monopolies restricting output, raising the prices they charge, restricting consumer choice and exercising their producer sovereignty.

25 Composite demand is demand for a good which has more than one use: for example, barley can be used as an animal feed or to make beer. This example also illustrates derived demand, since the demand for barley is derived from the demand for meat products and the demand for beer. The demand for labour is also a derived demand, which increases when the demand for the goods workers help to produce goes up.

26 $$\text{wage elasticity of demand for labour} = \frac{\%\text{ change in demand for labour}}{\%\text{ change in wage rate}}$$

27 $$\text{wage elasticity of supply for labour} = \frac{\%\text{ change in supply of labour}}{\%\text{ change in wage rate}}$$

28 If workers attempt to sell their labour at a wage rate above the ruling market wage existent in a perfectly competitive labour market, they will price themselves out of jobs. To remain employed in such a market, workers must passively accept the ruling market wage determined by market forces in the labour market as a whole. Their employers are also passive price-takers in the sense that if they offer their workers a wage which is lower than the ruling market wage, their workers will leave their employment and work for employers who are willing to pay the ruling market wage.

29 Unlike employers in a perfectly competitive labour market who are passive price-takers, a monopsony employer has the power to force down the wage rate paid to workers. Possessing this power, the monopsonist is a price-maker, the wage rate being the price of labour.

30 Income is a continuous flow or stream of payments, whereas wealth is a historical accumulation that has built up over time. Wealth is measured as a 'snapshot' at a particular point in time, whereas income is measured over a particular time period, for example hourly, weekly, monthly, quarterly or annually.

31 In June 2019, the Institute for Fiscal Studies claimed that relative poverty, defined as a household income lower than 60% of median income, affects almost a fifth of people in working households in the UK. At the time, a relatively poor single-person household was one with an income of less than £152 a week, while for a couple with two young children, the relative poverty line was £367 a week.

32 The London Stock Exchange and the Lloyds Insurance market are both located in the City of London. The main fish market in Scotland is located in the town of Peterhead. Aalsmeer Flower Auction is a flower auction that takes place in Aalsmeer in the Netherlands. It is the largest flower auction in the world.

33 The rationing function of prices is related to, but not quite the same as, the allocative function of prices. The rationing function distributes scarce goods to those consumers who value them most highly. The allocative function directs resources between markets, away from the markets in which prices are too high and in which there is excess supply, towards the markets where there is excess demand and price is too low. The allocatively efficient or 'correct' quantity of any good produced and consumed is the quantity that people choose to consume when $P = MC$.

34 Pure public goods are completely non-excludable, non-rival and non-rejectable. Quasi-public goods have characteristics of both private and public goods, including partial excludability, partial rivalry, partial diminishability and partial rejectability. Examples include roads, tunnels and bridges. Markets for these goods are considered to be incomplete markets.

35 People gain utility, at least in the short run, from consuming a demerit good such as tobacco. By contrast, people suffer disutility or negative welfare when forced to be near an economic 'bad' such as garbage.

36 Privatisation is the transfer of assets, for example firms and industries, from public ownership to private ownership. Regulation is the imposition of rules and laws which limit and constrain people's freedom of economic action. The most extreme form of regulation is to make it illegal to emit an externality such as pollution. Lesser forms of regulation are restrictions on time of day or year when it is legal to emit the externality, maximum emission limits and forcing polluters to invest in clean technology.

37 Market failure occurs whenever markets perform badly or unsatisfactorily. Markets may fail either because they perform inequitably (unfairly or unjustly) or because they perform inefficiently. Government failure occurs when government intervention in the economy is ineffective, wasteful or damaging.

STUDENT GUIDE

AQA A-LEVEL

Economics

Individuals, firms, markets and market failure

Reinforce your understanding of Individuals, firms, markets and market failure with this A-level Economics Student Guide. This book covers Themes 1–8 of the 14 themes in A-level Economics which can be examined in A-level Paper 1 Markets and market failure and A-level Paper 3 Economic principles and issues:

- Economic methodology and the economic problem
- Individual economic decision making
- Price determination in a competitive market
- Production, costs and revenue
- Perfect competition, monopoly and imperfectly competitive markets
- The labour market
- The distribution of income and wealth: poverty and inequality
- The market mechanism, market failure and government intervention in markets

This Student Guide will help you to:

- Identify key content for the exams with our concise coverage of topics
- Avoid common pitfalls with clear definitions and exam tips throughout
- Reinforce your learning with bullet-list summaries at the end of each section
- Test your knowledge with rapid-fire knowledge check questions and answers
- Find out what examiners are looking for with our Questions & Answers section

Other Student Guides for AQA A-level Economics

9781510472013 The national and international economy
Ray Powell, James Powell, George Vlachonikolis

HODDER EDUCATION
t: 01235 827827
e: education@bookpoint.co.uk
w: hoddereducation.co.uk

ISBN 978-1-5104-7200-6