

A-level ECONOMICS

Paper 1 Markets and Market Failure

Time allowed: 2 hours

Materials

For this paper you must have:

- an AQA 12-page Answer Book
- a calculator.

Instructions

- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Write the information required on the front cover of your answer book.
The **Paper Reference** is 7136/1.
- In **Section A**, answer **EITHER** Context 1 **OR** Context 2.
- In **Section B**, answer **ONE** essay.

Information

- The marks for questions are shown in brackets.
- The maximum mark for this paper is 80.
- There are 40 marks for **Section A** and 40 marks for **Section B**.

Advice

- You are advised to spend 1 hour on **Section A** and 1 hour on **Section B**.

Section A

Answer **EITHER** Context 1 **OR** Context 2.

EITHER

Context 1

Total for this context: 40 marks

The pharmaceutical industry

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

Extract A

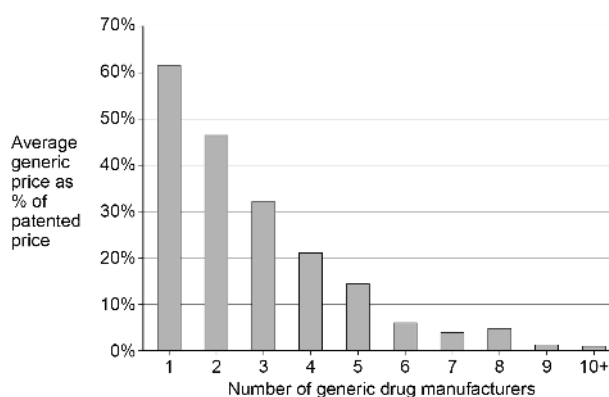
Figure 1: Profitability of new patented drugs

Year	Mean rate of return on new drugs (%)
2010	10.1
2011	7.6
2012	7.3
2013	4.8
2014	5.5
2015	4.2
2016	4.2
2017	3.7
2018	1.9
2019	1.8

Source: Deloitte report, 2019

Note: the mean rate of return is the average profit a pharmaceutical company can expect to make on a new drug over its lifetime, as a percentage of its development cost.

Figure 2: Price of generic drugs compared to patented drugs, and number of manufacturers



Source: Oxera report, 2019

Note: a generic drug is one which is no longer protected by a patent, and so can be made by any pharmaceutical firm, not only the original manufacturer.

Extract B: Rising prices in the global pharmaceutical market

Pharmaceutical drugs are very expensive to research and develop, costing an average of \$2.6bn. The most expensive, Zolgensma (a specialist gene therapy drug), costs \$2.1 million per patient. In order to encourage firms to invest in research and development, governments grant patents for new drugs. Patents preserve a company's intellectual property rights and give the company a legal monopoly that usually lasts for 20 years. Global drug prices have soared. The USA spent \$334bn on prescription drugs in 2017, up 41% from 10 years ago. Some price rises have been dramatic. The price of one antiparasitic drug, still protected by a patent, rose from \$13.50 to \$750 per pill.

After the monopoly's patent expires, other companies can manufacture and market the same drug under its generic name, usually at a lower price. For generic drugs, the UK government's policy is to rely on competition in the market between suppliers to control their price. Despite being a monopsony purchaser of these medicines in the UK, the NHS has relatively limited influence over their price in what is a global market. During 2017–18, the prices of certain generic drugs increased substantially, partly as a result of two large manufacturers having their production suspended because of quality issues. For example, the price of a generic mental health drug called Quetiapine peaked at £113.10 per 100mg, having previously cost £1.59.

The NHS budget of £138bn is not enough to meet all the demands for its services. Therefore, difficult allocation choices must be made. Rationing of care and treatment is inevitable. NHS spending on drugs in England has grown from £13bn in 2010–11 to £18.2bn in 2017–18, an average growth of around 5% per year. Policy on drugs in England aims to balance the competing goals of giving patients prompt access to effective treatments, incentivising the pharmaceutical sector to develop new products and ensuring that expenditure on drugs is affordable for the NHS. Today it is becoming harder to balance these objectives.

Source: News reports, 2019

Extract C: Is the market for pharmaceuticals working well?

The market for most generic drugs is oligopolistic, with just two or three manufacturers making each product. This enables significant economies of scale within firms, but also presents the risk of consumer interests being harmed. In 2019, US regulators accused 20 firms of colluding to raise generic drug prices by as much as 1000%. Combined with the monopoly for those drugs protected by patents, high prices and profits prevail in many parts of the industry. However, pharmaceutical firms argue that they take enormous financial risks in the search for new drugs. GlaxoSmithKline has spent just under £20bn on research and development over the past five years, while AstraZeneca spent nearly £30bn. These high development costs have led to new drugs becoming less profitable over the last decade. Pharmaceutical firms also provide thousands of highly skilled jobs and a reliable and safe product for consumers. Medicines and vaccines have played a crucial role in increasing UK life expectancy by 11 years since 1960. Over the last 40 years, cancer survival rates have doubled, and HIV/AIDS has been transformed into a manageable illness. UK firms are at the forefront of such developments and they make large profits in doing so. Many drugs have an extremely inelastic demand, as patients require them to maintain good health, which provides monopoly firms and collusive oligopolies with an opportunity to exploit their position. If this happens, the Competition and Markets Authority has the power to tackle the problem aggressively.

Source: News reports, 2019

- 0 1** Using the data in **Extract A** (Figure 1), calculate the mean amount of profit that pharmaceutical companies would have made on a new drug in 2019, given that the average development cost was \$2.6bn. Give your answer to the **nearest million dollars**. **[2 marks]**
- 0 2** Explain how the data in **Extract A** (Figure 2) show that greater competition leads to lower prices in the market for generic drugs. **[4 marks]**
- 0 3** **Extract B** (lines 9–10) states: ‘After the monopoly’s patent expires, other companies can manufacture and market the same drug under its generic name, usually at a lower price.’
- With the help of a diagram, explain why prices of pharmaceutical drugs supplied by a monopoly are likely to be higher than in markets which are more competitive. **[9 marks]**
- 0 4** **Extract C** (lines 1–3) states: ‘The market for most generic drugs is oligopolistic, with just two or three manufacturers making each product. This enables significant economies of scale within firms, but also presents the risk of consumer interests being harmed.’
- Using the extracts and your knowledge of economics, assess the view that the market structure of the pharmaceutical industry is damaging for consumers. **[25 marks]**

Turn over ►

Do **NOT** answer Context 2 if you have answered Context 1.

OR

Context 2

Total for this context: 40 marks

The ethnic pay gap

Study **Extracts D, E and F** and then answer **all** parts of Context 2 which follow.

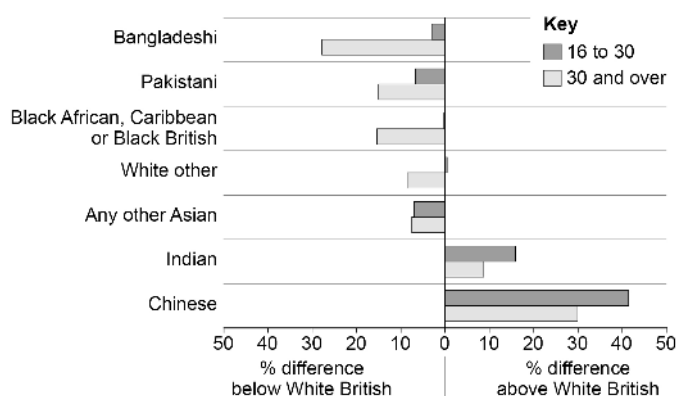
Extract D

Figure 3: Pay gap between ethnic groups and White British employees in the UK, 2018

Ethnic group	% pay gap compared with White British employees
Bangladeshi	−20.2
Pakistani	−16.9
Black African, Caribbean or Black British	−9.2
White other	−5.8
Any other Asian	−4.0
Indian	+12.0
Chinese	+30.9

Source: ONS, 2019

Figure 4: Pay gap between ethnic minority groups, by age, UK, 2018



Source: ONS, 2019

Extract E: The Ethnic Pay Gap

UK workers of Bangladeshi heritage have the lowest pay of any ethnic group, earning 20.2% less than White British workers. Analysis by the Office for National Statistics (ONS) shows significant pay gaps remain, even when education and occupation are taken into account, particularly for those born outside the UK. The median pay of White British workers last year was £12.03 an hour, compared with £9.60 for people of Bangladeshi ethnicity and £10.00 for those of Pakistani ethnicity. These groups also had the lowest employment rate. British-born employees of Chinese ethnicity earned the most of all groups, earning 30.9% more than White British workers.

1

5

Evidence suggests that the ethnic pay gap is reducing as the UK labour market evolves over time. People aged 16 to 30 from ethnic minority groups tended to have narrower pay gaps than older ethnic minority groups. For instance, the difference for the Bangladeshi ethnic group compared with White British workers was 3.1% among 16 to 30-year-olds but 27.9% for those over 30. Women in some ethnic groups were significantly less likely to be in the labour force than women in other ethnic groups. The ONS suggested that this could be a result of cultural differences as 38.1% of women from a Bangladeshi ethnic group and 32.1% of women from a Pakistani ethnic group were found to be economically inactive. This may be because they were looking after their family or home.

10

15

Source: News reports, 2019

Extract F: Tackling the ethnic pay gap

It is illegal to pay someone less for doing the same job based on their ethnicity. Andy Haldane, the Bank of England's Chief Economist, said that "the ethnicity pay gap problem in the UK is every bit as acute as the gender pay gap problem". A report by the management consultants, McKinsey, found that businesses with employees from the most diverse backgrounds were 15% more profitable than those firms with the least diverse workforce. A government report in 2017 suggested equal participation and progression across ethnicities could boost the UK's economy by an additional £24bn a year. BAME (Black, Asian or Minority Ethnic) workers are more likely to be in low-paid sectors such as catering, hairdressing and retail. These occupations have limited opportunities for workers to progress to more responsible, higher-paid roles. This may represent a misallocation of workers across the economy.

Only a handful of firms currently publish ethnic pay data. One of these, ITN, found that its BAME workforce were paid 21% less than White workers. The Conservative Government has indicated that it may introduce compulsory reporting of the ethnic pay gap for large firms. The Runnymede Trust said: "as well as getting companies to publish the ethnicity pay gap, you have to legally ask them to publish plans as to how they are going to close that gap. Otherwise, to be honest, nothing is going to happen". The government can ensure that public services such as the police and NHS provide such plans, but it is not intending to require all firms to do so. It could also take steps to recognise foreign qualifications, thus removing obstacles for well-qualified foreign-born workers. Others think the government should go further and consider ethnicity quotas on the management boards of major companies. In the absence of other reliable solutions, quotas could provide an urgent intervention to address the fact that currently only 4 of the biggest 100 listed companies are led by BAME executives, despite BAME groups accounting for 14% of the population. However, such quotas are rejected by some as discriminatory, patronising and inefficient, leading to the risk of government failure.

Source: News reports, 2019

- 0 5** Using the data in **Extract D** (Figure 3), calculate how much a worker of Bangladeshi ethnicity earned on average in 2018, given that a White British worker earned £28 677. Give your answer **to the nearest pound**. **[2 marks]**
- 0 6** Explain how the data in **Extract D** (Figure 4) show that the ethnic pay gap is more significant for older ethnic minority employees. **[4 marks]**
- 0 7** **Extract F** (lines 18–19) states: the Government 'could also take steps to recognise foreign qualifications, thus removing obstacles for well-qualified foreign-born workers.'
- With the help of a diagram, explain how UK firms not recognising foreign qualifications may lead to some ethnic groups receiving lower pay. **[9 marks]**
- 0 8** **Extract E** (lines 3–4) states: 'significant pay gaps remain, even when education and occupation are taken into account, particularly for those born outside the UK.'
- Using the extracts and your knowledge of economics, evaluate possible government policies to reduce the ethnic pay gap. **[25 marks]**

Turn over ►

Section B

Answer **one** essay from this section.

Each essay carries 40 marks.

EITHER**Essay 1**

Price comparison websites were used to purchase 53% of car insurance policies in 2017, improving competition and reducing prices for consumers. In other sectors, dominant firms such as Amazon have used innovation to capture market share. In the UK, 70% of online consumers say that Amazon is the first online retailer they visit when buying goods.

0	9
---	---

Explain how perfect competition should lead to outcomes which are **both** productively and allocatively efficient.

[15 marks]

1	0
---	---

Evaluate the view that technological change tends to bring industries closer to the market structure of perfect competition.

[25 marks]

OR**Essay 2**

Vaping, the use of e-cigarettes, is argued by many UK health experts to be a much less harmful substitute than smoking tobacco cigarettes. E-cigarettes contain fewer harmful chemicals such as tar, but they still contain addictive nicotine. The industry has been attacked for promoting sales amongst teenagers. This has led to San Francisco banning the sale of e-cigarettes.

1	1
---	---

Explain how the concept of cross elasticity of demand can be used to understand the relationship between markets.

[15 marks]

1	2
---	---

Discuss the view that individual economic agents will always act as rational decision makers so as to maximise their utility.

[25 marks]

OR

Essay 3

Aston Martin, the luxury car manufacturer, became a publicly listed company on the London Stock Exchange in October 2018, being valued at £3.89bn. The company had hoped to expand into new markets, enabling it to exploit economies of scale. However, following poor sales and weak profits, the company had lost 75% of its value by February 2020.

1	3
---	---

Explain how a firm may benefit from **both** internal and external economies of scale.

[15 marks]

1	4
---	---

Discuss how the divorce of ownership from control may affect **both** the conduct and performance of firms.

[25 marks]

END OF QUESTIONS

There are no questions printed on this page

Copyright information

For confidentiality purposes, all acknowledgements of third-party copyright material are published in a separate booklet. This booklet is published after each live examination series and is available for free download from www.aqa.org.uk.

Permission to reproduce all copyright material has been applied for. In some cases, efforts to contact copyright-holders may have been unsuccessful and AQA will be happy to rectify any omissions of acknowledgements. If you have any queries please contact the Copyright Team.

Copyright © 2021 AQA and its licensors. All rights reserved.



2 1 6 A 7 1 3 6 / 1

0 1

Mean profit = $1.8\% \times \$2.6\text{bn} = 0.018 \times \$2\,600\,000\,000 = \$46\,800\,000$

Answer = \$47 million (to the nearest million dollars)

0 2

The number of generic manufacturers is a measure of competition: the more firms producing the same off-patent drug, the more each must undercut its rivals to win sales. Figure 2 shows that as the number of manufacturers rises, the average generic price as a percentage of the patented price falls.

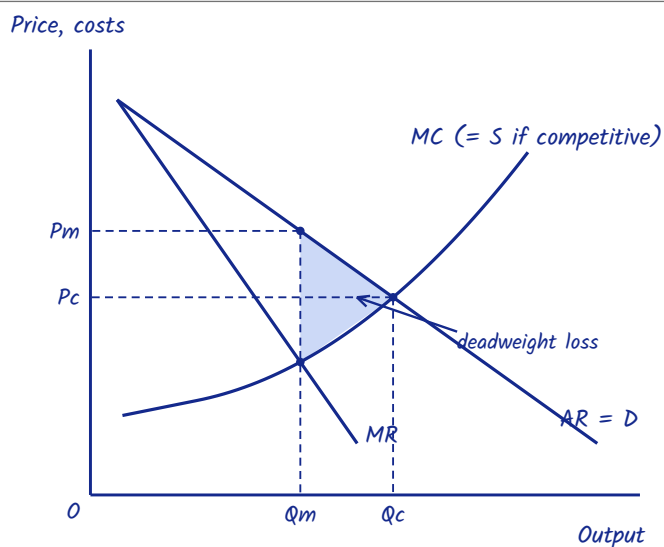
With one manufacturer the generic price is about 61% of the patented price. This falls to about 46% with two firms, 32% with three and 21% with four, and with 9 or 10+ firms the price is only about 1.2% and 1.0% of the patented price. The entry of the second firm has the biggest effect (a fall of about 15 percentage points), and each extra firm has a smaller marginal effect, with prices levelling off at around 4% to 6% beyond six firms (there is a slight rise at eight firms). This shows that greater competition removes the ability to set prices above cost, pushing price down towards the very low marginal cost of producing a generic drug.

0 3

A monopoly is a market with a single seller, or in UK law a firm with at least 25% market share, protected by barriers to entry. A patent gives a pharmaceutical firm a legal monopoly on a new drug for about 20 years (Extract B), so no other firm can produce it. In a more competitive market there are many firms selling a similar or identical product, so each firm has little or no power over price.

As shown in the diagram, the monopolist is a price maker facing the downward-sloping market demand curve (AR), so its marginal revenue (MR) lies below AR. To maximise profit it produces where $MC = MR$, at output Q_m , and charges the highest price consumers will pay for that output, P_m , read from the AR curve. In a competitive market firms are price takers and the industry supply curve is the sum of the firms' MC curves, so equilibrium is where $MC = AR$, at the higher output Q_c and the lower price P_c . The monopolist therefore restricts output to raise price, earning supernormal profit, and there is a deadweight loss of

welfare (shaded), as $P > MC$ so the outcome is allocatively inefficient.



Demand for many drugs is extremely price inelastic (Extract C) because patients need them to stay alive or healthy and there are no close substitutes while the patent lasts, so a monopolist can raise price with little loss of sales, as with the antiparasitic drug whose price rose from \$13.50 to \$750 per pill. Firms also argue that high prices are needed to recover research and development costs averaging \$2.6bn per drug, including testing and regulatory approval. A monopoly faces less pressure to keep costs down, so X-inefficiency may raise costs and prices further. When the patent expires, entry occurs. Because the marginal cost of producing a pill is very low, competition between generic manufacturers drives price down towards MC: Figure 2 shows generic prices falling to around 1% of the patented price with 10 or more firms. This is why prices under a patent monopoly are likely to be far higher than in more competitive markets.

0 4

The pharmaceutical industry has two main market structures. New drugs are supplied by monopolies protected by patents lasting about 20 years, while the market for most generic drugs is an oligopoly, with just two or three manufacturers making each product (Extract C). Consumer interests in this market include low and stable prices, safe and effective drugs, prompt access to treatment and a steady flow of new medicines in the future. Drugs are merit goods with large external benefits, so the outcomes of this market matter a great deal.

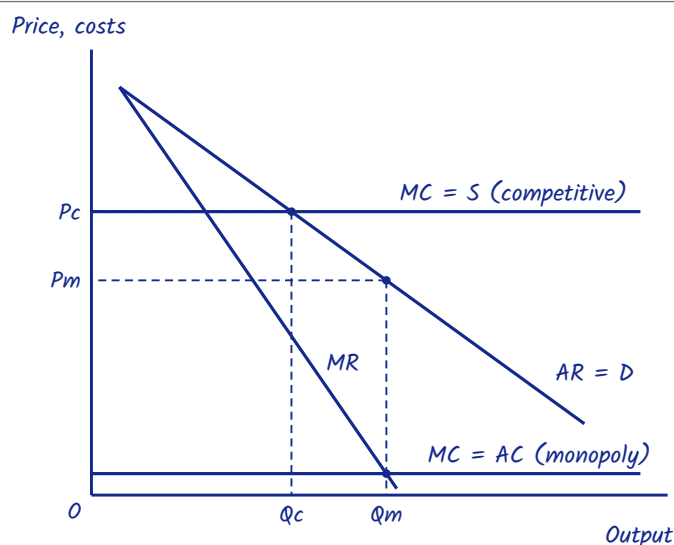
The main argument that the market structure is damaging is that monopoly and oligopoly power lead to high prices. As explained above, a patent holder maximises profit where $MC = MR$, restricting output and charging a price well above marginal cost, so there is allocative inefficiency and a loss of consumer surplus. Because demand for many drugs is extremely price inelastic, firms can exploit consumers: the price of one patented antiparasitic drug rose from \$13.50 to \$750 per pill, a rise of over 5400%, and Zolgensma costs \$2.1 million per patient. The USA spent \$334bn on prescription drugs in 2017, up 41% in ten years.

Oligopoly in the generics market also creates risks. With only two or three firms and products that are homogeneous with little branding, firms are highly interdependent and have a strong incentive to collude, tacitly or overtly, to act like a joint monopoly. In 2019 US regulators accused 20 firms of colluding to raise generic prices by as much as 1000%. With so few suppliers, the market is fragile: when two large manufacturers had production suspended in 2017 to 2018, the price of Quetiapine rose from £1.59 to £113.10, a rise of over 7000%. In a competitive market no single firm is critical, so supply shocks have far smaller effects on price.

High drug prices also harm consumers indirectly. The NHS budget of £138bn cannot meet all demands, and spending on drugs in England grew from £13bn in 2010 to 2011 to £18.2bn in 2017 to 2018. Every extra pound spent on overpriced drugs has an opportunity cost in nurses, operations or other treatments, so rationing becomes more severe. In poorer countries without a strong health service, high prices may mean patients go without treatment altogether, which is inequitable.

However, there are strong arguments that the market structure benefits consumers, especially in the long run. Drug development is extremely costly and risky: a new drug costs on average \$2.6bn, and GlaxoSmithKline and AstraZeneca spent nearly £20bn and £30bn on research and development over five years. Without patents, rivals could copy a new drug at low marginal cost as soon as it was launched, the innovator could not recover its sunk costs and there would be little incentive to innovate, which is a classic market failure. The supernormal

profit protected by patents finances research, giving dynamic efficiency. The benefits are huge: medicines have helped increase UK life expectancy by 11 years since 1960, cancer survival rates have doubled over 40 years and HIV/AIDS is now a manageable illness. Profits are also not as excessive as they seem: Figure 1 shows the mean rate of return on new drugs fell from 10.1% in 2010 to only 1.8% in 2019, suggesting that monopoly profits are only just covering development costs.



Large firms also benefit from economies of scale. Research, testing, regulatory approval and production of drugs involve very high fixed costs, so long-run average cost falls as output rises and the minimum efficient scale is large relative to the market. As shown in the diagram, if a large firm's economies of scale lower its marginal and average costs from $MC = S$ (competitive) to the lower $MC = AC$ (monopoly) curve, the profit-maximising monopoly price P_m can be below the price P_c that a fragmented competitive industry would charge, with higher output, Q_m compared with Q_c . Generic oligopolists can therefore be productively efficient and still pass on low prices, and the industry provides a reliable and safe product and thousands of highly skilled jobs.

The generics market also shows that the structure is not permanent. Once a patent expires the market becomes contestable, and Figure 2 shows that each extra manufacturer cuts prices sharply, so that with 10 or more firms the generic price is only 1% of the patented price. The harm from monopoly is therefore temporary, limited by the length of the patent.

Furthermore, consumers are protected by countervailing power and regulation. In the UK the NHS is a monopsony buyer of drugs, so it can use its buying power to negotiate lower prices for patented drugs and only fund those that are cost-effective, which limits exploitation, unlike in the USA where there are many buyers and prices are much higher. The Competition and Markets Authority can investigate abuse of dominance and collusion, impose fines of up to 10% of turnover and act "aggressively" (Extract C). However, the NHS has "relatively limited influence" over generic prices in a global market, and regulation may suffer from government failure, as regulators have imperfect information about firms' true costs and investigations are slow, so price spikes can happen before action is taken.

Whether the market structure is damaging therefore depends on several factors. First, the time period: in the short run monopoly pricing harms consumers, but in the long run the innovation it finances gives life-saving treatments that would not otherwise exist. Second, the part of the industry: patents are a deliberate trade-off that is justified, while collusion in the generics market brings no comparable benefit and is clearly damaging. Third, the strength of countervailing power: UK consumers are protected by the NHS monopsony far more than consumers in the USA. Fourth, whether firms actually use their profits for research or for marketing and dividends.

In conclusion, the market structure of the pharmaceutical industry is not damaging for consumers overall, but it does create a serious risk of harm. Patent monopolies lead to very high prices in the short run, yet they are the main source of the dynamic efficiency that has transformed life expectancy, and falling rates of return suggest profits are not excessive. The greater danger is in the oligopolistic generics market, where collusion and dependence on a few suppliers can lead to sudden price rises with no offsetting benefit. The best outcome is to keep the patent system, while the CMA and the NHS use their powers to prevent collusion and encourage more firms to enter generics markets.

0 5

Bangladeshi workers earned 20.2% less, so $100\% - 20.2\% = 79.8\%$

$$£28\,677 \times 0.798 = £22\,884.25$$

Answer = £22 884 (to the nearest pound)

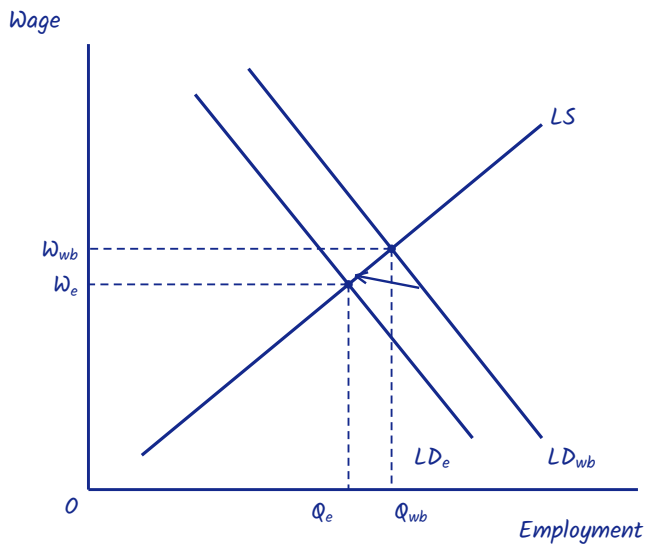
0 6

The ethnic pay gap is the percentage difference between the average pay of an ethnic minority group and that of White British employees. Figure 4 shows that for every group with a negative pay gap, the gap is larger for those aged 30 and over than for those aged 16 to 30.

The biggest difference is for Bangladeshi workers, whose pay gap is 27.9% for over-30s but only 3.1% for 16 to 30-year-olds. Pakistani workers aged 30+ earn about 15% less than White British workers compared with about 7% for younger workers, and Black workers aged 30+ earn about 15% less while younger Black workers have almost no gap. White other workers aged 16 to 30 earn slightly more than White British workers, but those aged 30+ earn about 8% less. The gap for Any other Asian is similar for both ages (about 7% to 8%), but still slightly larger for older workers. Even for the groups that earn more than White British workers, the advantage is smaller for older workers: Chinese workers earn about 42% more at 16 to 30 but only 30% more at 30+, and Indian workers about 16% compared with 9%. So older ethnic minority employees are relatively worse off in every group.

0 7

Employers demand labour because of the output it produces, so the demand for labour is a derived demand based on the marginal revenue product (MRP) of workers. Qualifications such as GCSEs, A-levels and degrees signal a worker's skills and productivity. If UK firms do not recognise or understand foreign qualifications, they treat well-qualified foreign-born workers as if they had lower skills, so their perceived MRP is lower, even if they are otherwise identically qualified and as productive as British-educated workers. Recruiting them also seems riskier, so firms prefer applicants with familiar UK qualifications for jobs and promotions.



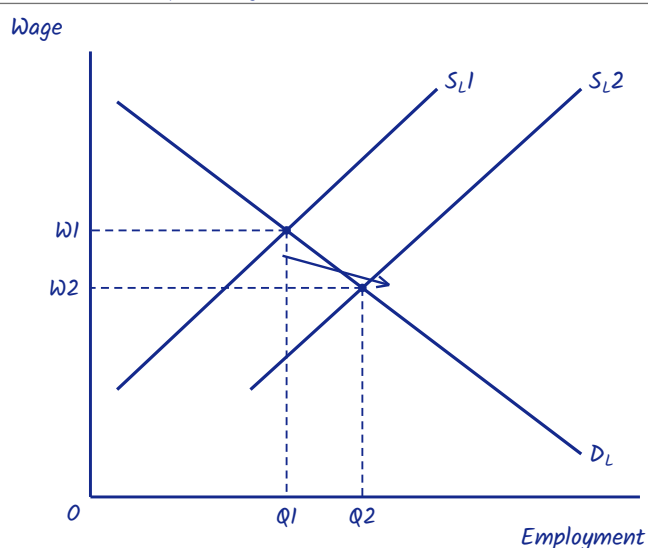
As shown in the diagram, the demand for labour of ethnic minority workers with foreign qualifications, LD_e , lies to the left of the demand for White British workers with UK qualifications, LD_{wb} . With a given labour supply curve LS , the equilibrium wage for these ethnic minority workers is w_e , below the wage w_{wb} of White British workers, and employment is lower at Q_e compared with Q_{wb} . Since many workers born outside the UK belong to ethnic minority groups, this leads to lower average pay for those groups.

In addition, workers whose qualifications are not recognised are often forced into occupations where few skills are needed, such as catering, hairdressing and retail (Extract F). This increases the supply of labour to low-paid jobs, shifting supply to the right and lowering wages further, while these jobs offer limited opportunities for progression. It also reduces the elasticity of supply of these workers to higher-paid jobs, since they cannot easily move into them.

This helps explain Extract E's finding that significant pay gaps remain even when education and occupation are taken into account, "particularly for those born outside the UK", and why British-born Chinese employees earn 30.9% more than White British workers. The result is a misallocation of labour, as skilled workers are not used where their MRP is highest.

The ethnic pay gap is the difference between the average pay of ethnic minority workers and White British workers. Figure 3 shows Bangladeshi workers earn 20.2% less and Pakistani workers 16.9% less, with median hourly pay of £9.60 and £10.00 compared with £12.03. Gaps arise partly from differences in education, occupation, age and labour force participation, but Extract E states that significant gaps remain even when education and occupation are taken into account, which suggests discrimination and the failure to recognise foreign qualifications also play a role.

There is a strong case for government intervention. The pay gap is a market failure: if workers are paid according to ethnicity rather than their MRP, labour is misallocated, with skilled BAME workers stuck in low-paid sectors such as catering and retail (Extract F). This is allocatively inefficient and reduces output: equal participation and progression could add £24bn a year to UK GDP, and firms with the most diverse workforces are 15% more profitable. The gap also causes inequity and poverty, and lowers tax revenue.



The first policy is anti-discrimination legislation, which already exists: it is illegal to pay someone less for doing the same job because of their ethnicity. In theory, this forces firms to pay equal wages for equal work. However, its effect is limited because most of the pay gap comes from occupational segregation and lack of promotion rather than unequal pay for the same job, and discrimination in hiring is hard to prove. The persistence of gaps of 20% shows legislation alone is not enough.

A second policy is compulsory publication of ethnic pay gap data for large firms, which the government has indicated it may introduce. This reduces information failure: firms, workers and investors can see where gaps exist, and firms with large gaps face reputational damage, which gives them an incentive to act. Gender pay gap reporting since 2017 has raised awareness at low cost. However, data alone may change little; as the Runnymede Trust says, "otherwise nothing is going to happen". Firms might also face compliance costs, and small firms would be excluded. Only a handful of firms currently publish the data: ITN found a 21% gap.

Requiring firms to publish plans to close the gap would make the policy stronger, and the public sector, such as the police and NHS, can lead by example as a very large employer. But plans may be vague and there is a risk of box-ticking without real change, and the government does not intend to require all firms to do so.

A third policy is to recognise foreign qualifications. As explained in the previous answer, this would increase the perceived MRP and so the demand for well-qualified foreign-born workers, raising their wages and employment. It would allow them to move out of low-paid occupations, reducing the supply of labour there; as shown in the diagram, a fall in supply from S_L2 back to S_L1 would raise wages in those occupations from $W2$ to $W1$. This tackles directly the gap "particularly for those born outside the UK" at low cost to the government. However, it would take time to establish equivalences between many different national systems, some qualifications may not match UK standards, and firms may still prefer familiar candidates due to unconscious bias.

A fourth policy is ethnicity quotas on the boards of major companies. Only 4 of the largest 100 listed companies are led by BAME executives, despite BAME groups being 14% of the population. Quotas would act quickly, create role models and could change the culture of firms, raising pay at the top. However, they are criticised as discriminatory, patronising and inefficient (Extract F): if appointments are made to meet a target rather than on merit, productivity may fall, and those appointed may face resentment. Quotas affect only a few senior jobs, so their

impact on the average pay gap would be small. This is a clear risk of government failure.

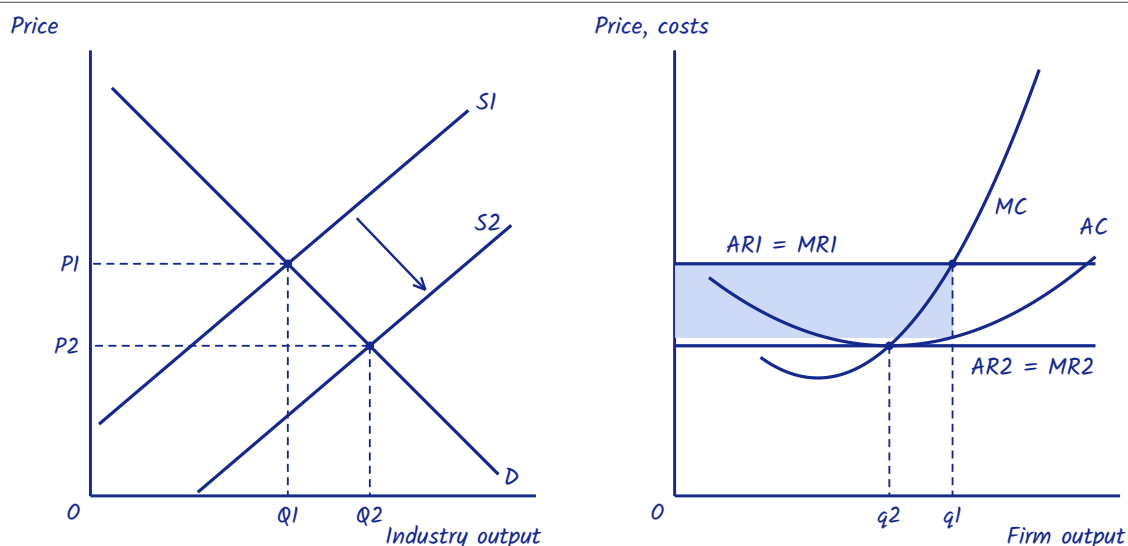
Other policies could address the underlying causes. Education and training, mentoring schemes and apprenticeships would raise the skills and occupational mobility of groups with low pay, while help with childcare could increase labour force participation, since 38.1% of Bangladeshi and 32.1% of Pakistani women are economically inactive. These are long-term supply-side policies with high costs but address root causes rather than symptoms.

In evaluating these policies, it is important to note that the pay gap is not uniform. Indian and Chinese workers earn more than White British workers, by 12.0% and 30.9%, suggesting that much of the gap reflects differences in education, occupation, age structure and culture rather than pure discrimination. Figure 4 and Extract E also show the gap is much narrower for younger workers (only 3.1% for young Bangladeshi workers compared with 27.9% for those over 30), so the gap may already be shrinking as the labour market evolves. This suggests that broad interventions such as quotas may be unnecessary and costly, while targeted policies may be more effective.

In conclusion, the most effective policies are likely to be compulsory reporting combined with action plans, and the recognition of foreign qualifications, because they tackle information failure and the misallocation of labour directly at relatively low cost and with little risk of government failure. Legislation is necessary but not sufficient, and quotas are a blunt tool that should only be a last resort if other policies fail. The success of any policy depends on enforcement, on whether firms change attitudes as well as following rules, and on how quickly the gap is already closing among younger workers. A combination of policies, including education and support for participation, is needed to close the gap in the long run.

Perfect competition is a market structure with a very large number of small firms selling an identical (homogeneous) product, perfect information for buyers and sellers, and no barriers to entry or exit. Because each firm is tiny relative to the market and the product is identical, no firm can influence price: each is a price taker, facing a perfectly elastic demand curve at the market price, so $P = AR = MR$.

Productive efficiency occurs when output is produced at the lowest possible average cost, the minimum point of the AC curve, where $MC = AC$. Allocative efficiency occurs when resources are allocated to produce the goods and services consumers value most, which happens when price equals marginal cost ($P = MC$), as the value consumers place on the last unit equals the cost of the resources used to produce it.



Firms are assumed to maximise profits, producing where $MC = MR$. Since $P = MR$ for a price taker, each firm produces where $P = MC$, so perfect competition is always allocatively efficient, in the short run as well as the long run. If a firm tried to charge above the market price it would sell nothing, as consumers with perfect information would buy the identical product elsewhere.

In the short run, the market price may be above average cost. In the diagram, the industry price is initially $P1$ and the firm produces $q1$, where $MC = MR1$. Since $AR1$ is above AC at $q1$, the firm makes supernormal profit (the shaded area).

With perfect information and no barriers to entry, new firms are attracted into the industry by these profits. Industry supply shifts from $S1$ to $S2$, so the price

falls from P_1 to P_2 , and output rises from Q_1 to Q_2 .

Entry continues until supernormal profits are competed away and firms earn only normal profit, where $AR_2 = AC$. This happens at the price P_2 , which is tangent to the lowest point of the AC curve, at output q_2 . At this point $MC = AC$, so each firm produces at minimum average cost and is productively efficient.

Conversely, if price were below AC, firms would make losses and some would leave, shifting supply left until price rose back to minimum AC.

Competition also forces firms to be productively efficient in another way. Any firm with higher costs than its rivals, for example through X-inefficiency, would make a loss at the market price and be forced out of the industry in the long run, so only firms producing at the lowest possible cost survive. Consumers benefit because price equals the minimum average cost, so they pay the lowest price consistent with firms covering their costs including normal profit.

Therefore, in long-run equilibrium, $P = MC = \text{minimum } AC = AR = MR$, so perfect competition leads to outcomes that are both allocatively and productively efficient. However, perfect competition is a theoretical benchmark. In the real world, markets such as price comparison websites for car insurance (where 53% of policies were bought in 2017) come closer to it by improving information, but few markets meet all the assumptions.

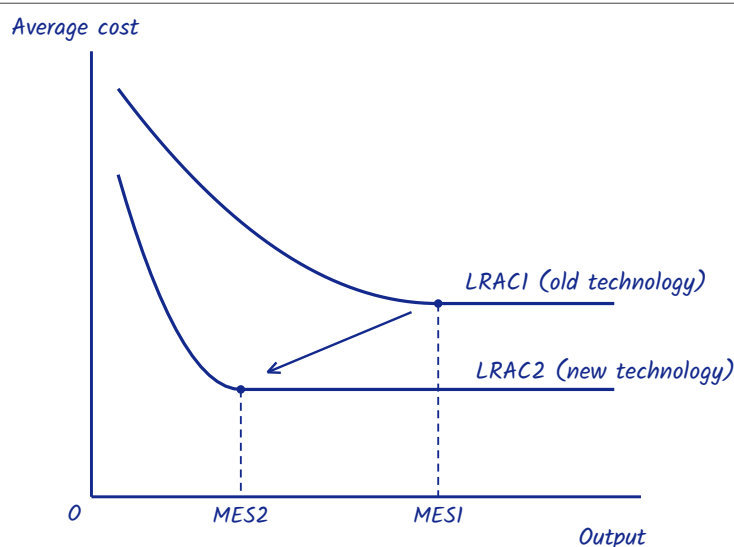
1 0

Technological change is the development and use of new methods of production, new products and new ways of communicating and selling, such as the internet, smartphones, automation and artificial intelligence. Perfect competition is a market structure with many small firms, identical products, perfect information and no barriers to entry or exit. In reality, perfect competition is an unattainable ideal in most industries, so the question is whether technological change moves industries towards this end of the spectrum of market structures, or towards oligopoly and monopoly.

There are strong arguments that technology increases competition. First, the internet has greatly improved price transparency and reduced information asymmetry. Price comparison websites were used to buy 53% of car insurance

policies in 2017, allowing consumers to compare many suppliers instantly. This makes demand facing each firm more price elastic, as consumers switch easily to the cheapest supplier, so firms behave more like price takers and prices fall towards marginal cost, closer to the perfect competition outcome.

Second, technology has lowered barriers to entry. A business can now be started with a laptop from a bedroom, selling worldwide through online marketplaces, social media and apps, without the need for expensive shops or advertising. Cloud computing turns fixed costs into variable costs, and 3D printing and small-scale digital production lower the minimum efficient scale.



As shown in the diagram, if new technology shifts the long-run average cost curve from LRAC1 to LRAC2 and reduces the minimum efficient scale from MES1 to MES2, many more firms can operate at the lowest cost within the market, increasing the number of sellers. Markets for taxis (ride-hailing apps), accommodation (home-sharing platforms) and online retail of crafts have all seen thousands of small sellers entering. Lower costs also allow lower prices, and contestability means even incumbent firms must keep prices close to costs.

Third, technology has increased the size of markets through globalisation, so local monopolies face competition from firms worldwide, and creative destruction, as Schumpeter argued, means that incumbent firms are constantly challenged by new entrants with better technologies, as streaming replaced CDs and video rental.

However, there are equally strong arguments that technology moves industries away from perfect competition. Many digital industries have very high fixed costs

and near-zero marginal costs, giving huge economies of scale, so the MES is very large and the market may be a natural monopoly. Network effects mean that the more users a platform has, the more valuable it becomes, so markets tend to "tip" towards one dominant firm: Amazon has used innovation, such as advanced logistics and inventory systems, to capture market share, and 70% of UK online consumers say Amazon is the first retailer they visit. Google, Facebook and Apple have similarly dominant positions.

Technology can also create new barriers to entry. Firms gather vast amounts of personal data and use artificial intelligence to target personalised adverts and prices, an advantage new entrants cannot copy. Patents and copyright on new technologies give legal monopolies. Technology also allows product differentiation through branding and features, making products less homogeneous, the opposite of perfect competition. Dominant firms often buy potential rivals, as with Facebook buying Instagram and WhatsApp, to prevent competition. Algorithms may even make tacit collusion easier, as firms can monitor and match each other's prices instantly.

Moreover, even where technology improves information, it can lead to new forms of market power. Price comparison websites themselves can become dominant and charge high commissions to insurers, and too much choice can overwhelm consumers, who may use rules of thumb rather than compare all options, so information is not truly perfect.

It can also be argued that moving away from perfect competition is not necessarily bad. Large firms with supernormal profits can fund research and development, giving dynamic efficiency, and economies of scale lower costs and prices, as Amazon's low prices and fast delivery show. Perfect competition, by contrast, may lack the funds for innovation.

Overall, the effect of technological change depends on the industry. Where technology mainly reduces information failures and entry costs, as in insurance, travel and small-scale retail, industries move closer to perfect competition. Where it creates economies of scale and network effects, as in online platforms, search and social media, it leads to greater concentration. It also depends on the time

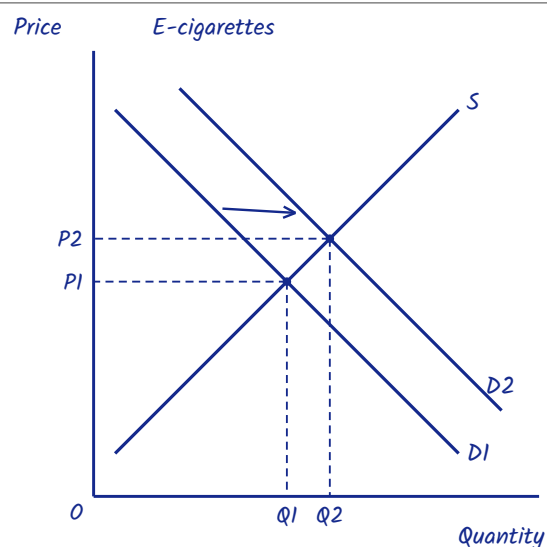
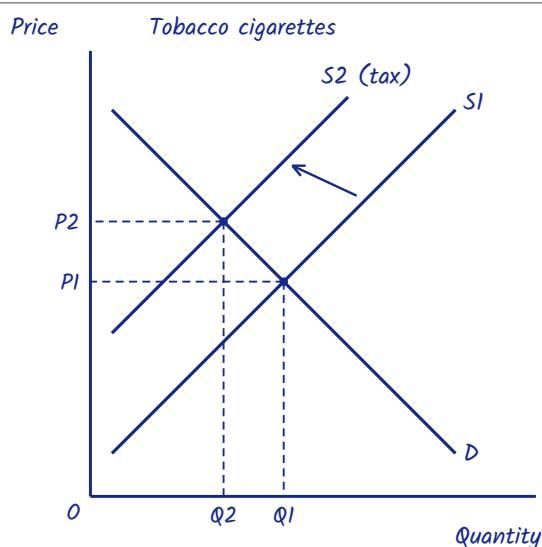
period: new technologies often first open up markets to many entrants, but over time a few winners emerge and markets consolidate, until a further wave of creative destruction challenges them.

In conclusion, technological change does not generally tend to bring industries closer to perfect competition. It increases competition in some respects, by improving information and contestability, but in many of the most important modern industries it has strengthened dominant firms, as the example of Amazon shows. Whether technology increases competition in the long run depends largely on competition policy, as regulators such as the CMA need to prevent dominant firms from abusing their power and buying up rivals, so that markets remain contestable.

Cross elasticity of demand (XED) measures the responsiveness of the quantity demanded of one good (good A) to a change in the price of another good (good B). It is calculated as $XED = \% \text{ change in quantity demanded of good A} \div \% \text{ change in price of good B}$. For example, if the price of tobacco cigarettes rises by 10% and the quantity demanded of e-cigarettes rises by 5%, $XED = +5 \div +10 = +0.5$.

The sign of XED shows the relationship between markets. If XED is positive, the goods are substitutes: a rise in the price of B causes consumers to switch to A, so demand for A rises. Examples include vaping and smoking, butter and margarine, or Coke and Pepsi. If XED is negative, the goods are complements, which are in joint demand: a rise in the price of B reduces the quantity demanded of B and therefore also of A, which is consumed with it, for example cigarettes and lighters, cars and petrol, or games consoles and games. If XED is zero or close to zero, the goods are unrelated, such as e-cigarettes and bread.

The size of XED shows the strength of the relationship. A high positive value (for example +2) shows close substitutes, so a small change in the price of one good has a large effect on demand for the other. A low positive value (for example +0.2) shows weak substitutes. Similarly, a large negative value shows strong complements.



The diagram shows how changes in one market affect a related market. Suppose the government increases the tax on tobacco cigarettes. In the market for tobacco cigarettes, supply shifts from S1 to S2, raising price from P1 to P2 and

reducing quantity from Q_1 to Q_2 . Because e-cigarettes are a substitute, with positive XED, demand for e-cigarettes shifts right from D_1 to D_2 , raising price from P_1 to P_2 and quantity from Q_1 to Q_2 . The size of the shift depends on the value of XED: the closer the substitutes, the bigger the increase. Similarly, a ban on e-cigarettes, as in San Francisco, would shift demand for tobacco cigarettes to the right. If XED were negative, for example between tobacco and cigarette papers, the tax would instead shift demand for the complement to the left.

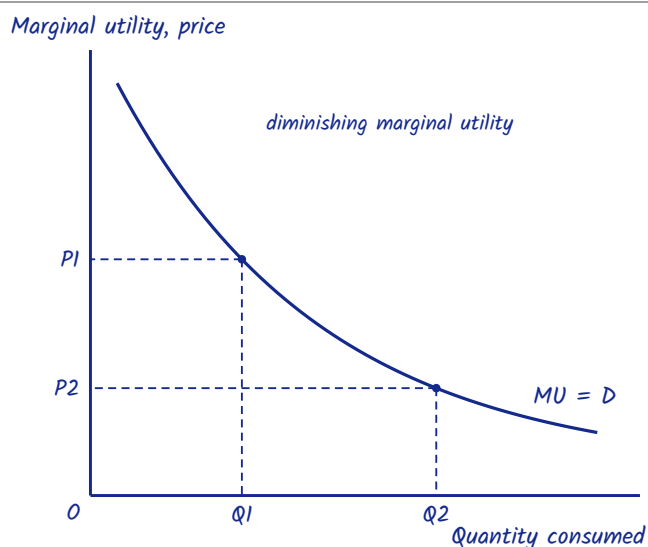
Understanding XED is useful for firms. A firm can estimate how its sales will be affected by a rival's price change: if XED with a rival's product is high, it may need to match a price cut. Firms selling complementary goods may use pricing strategies such as selling a console cheaply to increase demand for profitable games. Firms may also try to reduce XED with rivals' products through branding and advertising, so that their demand becomes less sensitive to competitors' prices. Merging firms are investigated by competition authorities, which use XED to define the market: a high XED between two firms' products suggests they compete in the same market.

XED is also important for government policy. When the government taxes a demerit good such as tobacco, it needs to know the XED with substitutes. If vaping is a much less harmful substitute, as many UK health experts argue, a tobacco tax that encourages switching would improve health outcomes, whereas a ban on e-cigarettes might increase smoking. Governments also consider effects on complements: a tax on petrol reduces demand for large cars.

However, XED estimates have limitations. They assume *ceteris paribus*, but in reality incomes, tastes and other prices change at the same time. Values also change over time as habits change: nicotine is addictive, so smokers may switch only slowly. Nevertheless, XED is a valuable concept for understanding how markets are linked.

Economic agents are individuals, firms and the government who make economic decisions. In traditional economic theory, individuals are assumed to be rational: they have stable and consistent preferences, perfect information and the ability to process it, and make choices that maximise their utility, which is the satisfaction or benefit gained from consuming goods and services. Firms are assumed to maximise profit.

Rational consumers make decisions at the margin. Total utility is the total satisfaction from consuming a given quantity, while marginal utility is the extra satisfaction from consuming one more unit. The law of diminishing marginal utility states that as more of a good is consumed, each additional unit gives less extra satisfaction than the one before.



A rational consumer maximises utility by buying a good until the marginal utility of the last unit equals its price, or more generally where the marginal utility per pound spent is equal across all goods ($MU_a/P_a = MU_b/P_b$). As shown in the diagram, because marginal utility diminishes, consumers will only buy more if the price falls: at P_1 they buy Q_1 , and at the lower price P_2 they buy Q_2 . This explains why the demand curve slopes downwards, and the assumption of rationality underlies much of microeconomic theory.

There is evidence that agents often do act rationally. Consumers compare prices, use price comparison websites and respond to price changes: demand curves slope downwards and taxes on cigarettes and sugary drinks do reduce consumption. Over time, consumers learn from mistakes, and in markets with repeated purchases,

such as groceries, they make reasonably rational choices. Even if not every individual is rational, the assumption may be a useful simplification that predicts the behaviour of markets overall.

However, behavioural economics argues that agents do not always act rationally.

First, information is often imperfect or asymmetric, so consumers cannot accurately judge the costs and benefits of a choice: for example, smokers may underestimate the long-term health risks. Measuring utility is also very difficult, so it is unclear how agents can maximise it.

Second, there is bounded rationality: people have limited time and limited ability to process information, so they make decisions that are "good enough" rather than optimal. They use rules of thumb (heuristics), such as always buying the same brand, which save time but may not maximise utility. They are also subject to biases. Anchoring means people rely too heavily on the first piece of information, such as a "was £100, now £50" price. The availability bias means people overestimate risks they can easily recall, such as plane crashes. Social norms mean people copy what others do, for example in fashion or in the teenage take-up of vaping, rather than making independent choices.

Third, there is bounded self-control. People often know what is best for them but fail to do it, such as people who intend to save for retirement or stop smoking but do not. Addictive goods such as nicotine make this especially strong. People are also present-biased, placing too much weight on immediate satisfaction compared with future costs, so they under-consume merit goods such as education and pensions and over-consume demerit goods.

Fourth, people do not always aim to maximise their own utility. Altruism, fairness and concern for others influence decisions: people give to charity, volunteer and reject unfair offers in experiments such as the ultimatum game even when this leaves them worse off. This could be included in a broad definition of utility, but then the theory risks explaining everything and predicting nothing.

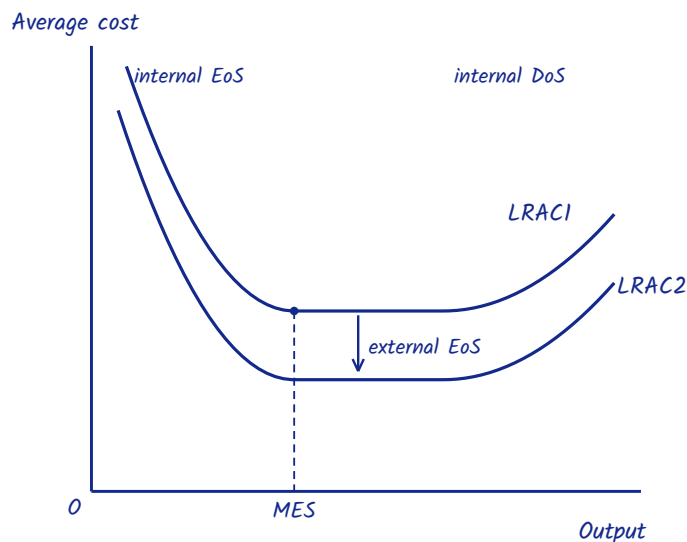
Firms and governments can exploit or correct irrational behaviour. Firms use marketing, default options and framing to encourage impulse purchases, for example promoting e-cigarettes to teenagers. Governments use nudges to help

people make better choices, such as auto-enrolment into workplace pensions, which greatly increased saving, and graphic health warnings on cigarettes. The success of nudges itself shows that people are not perfectly rational, because a rational agent would not be affected by how a choice is presented.

The extent of rationality depends on the type of decision. For simple, frequent, low-cost purchases with clear information, people tend to act fairly rationally. For complex, infrequent decisions with long-term consequences, such as pensions, mortgages and health, or for addictive goods, irrational behaviour is much more likely. It also depends on the individual: experience, education and the time available all matter.

In conclusion, the view that individual economic agents will always act rationally to maximise utility is not correct. The word "always" is too strong: behavioural economics shows that bounded rationality, bounded self-control, biases and social norms lead to systematic departures from utility maximisation, especially for complex or addictive choices. However, people often do act in a broadly rational way, and the assumption remains a useful starting point for explaining markets. The best approach is to treat rationality as a benchmark, while recognising that when agents act irrationally there may be market failure and a case for government intervention such as nudges, information provision and regulation.

Economies of scale are the reductions in long-run average cost that occur as a firm increases its scale of production. Internal economies of scale arise from the growth of the firm itself, while external economies of scale arise from the growth of the industry or the area in which the firm is located, and benefit all firms in that industry.



Internal economies of scale cause a movement down the firm's long-run average cost curve. As shown in the diagram, LRAC1 falls as output increases until the minimum efficient scale (MES), the lowest output at which average cost is minimised. LRAC may be L-shaped, remaining flat beyond MES, or U-shaped if diseconomies of scale occur at very high output.

Technical economies arise because large firms can use larger, more efficient capital equipment and production lines that would be too expensive for small firms. Fixed costs, such as the cost of designing a new car model, are spread over a larger output, so average fixed cost falls. The law of increased dimensions means that larger machinery and containers cost proportionately less to build and run. Large firms can also use the division of labour, with workers specialising in particular tasks, raising productivity. A car manufacturer like Aston Martin, producing a small number of cars, cannot benefit fully from these compared with mass-market producers.

Purchasing (bulk-buying) economies arise because large firms buy raw materials and components in large quantities and can negotiate discounts from suppliers using their buying power. Marketing economies arise because the cost of an

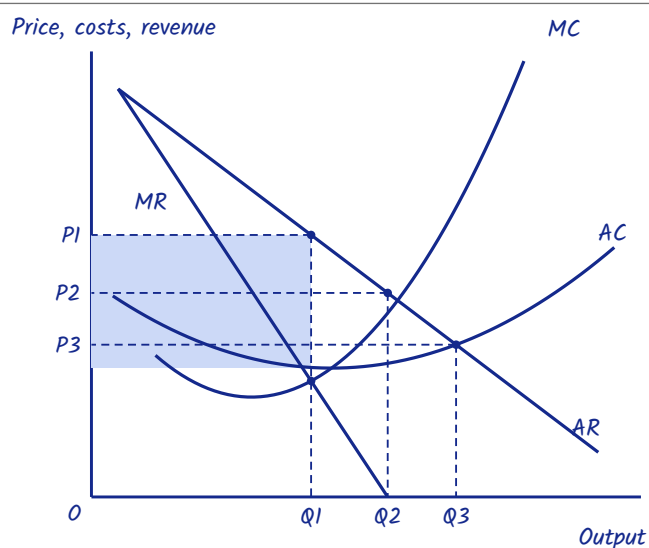
advertising campaign is spread over a larger output. Financial economies arise because large firms are seen as less risky by lenders, so they can borrow at lower interest rates and raise finance by issuing shares, as Aston Martin did on the London Stock Exchange. Managerial economies arise as large firms can employ specialist managers, raising efficiency. Risk-bearing economies arise because large firms can diversify into different products and markets, so a fall in demand in one market is offset by others.

Network economies arise in industries such as online platforms, where the value of the service increases as more users join, while the cost of serving each additional user is very low.

External economies of scale cause the whole LRAC curve to shift downwards, from LRAC1 to LRAC2 in the diagram, because they reduce costs at every level of output. They often arise when an industry is concentrated in one area. For example, a pool of skilled labour develops, as local colleges and universities train workers in the skills the industry needs, reducing training and recruitment costs. Specialist suppliers locate nearby, reducing transport costs and delivery times. Improved transport and communication links, such as roads, ports and broadband built to serve the industry, lower costs. Links with universities allow firms to share research and ideas, as in the motorsport cluster in the Midlands or the technology cluster in Cambridge and Silicon Valley.

A firm benefits from economies of scale in several ways. Lower average costs mean that, at a given price, it makes higher profits, which can fund investment and research. Alternatively, it can cut prices to increase market share and become more competitive, and lower costs may act as a barrier to entry against new firms that cannot achieve the MES. This is why Aston Martin hoped to expand into new markets to exploit economies of scale. However, the firm only benefits if demand is high enough for it to reach a large scale: Aston Martin's poor sales meant it could not achieve the expected economies, and beyond a certain size diseconomies such as communication problems may raise costs.

The divorce of ownership from control occurs when the owners of a firm, the shareholders, are different from the people who control it day to day, the directors and managers. It is typical of large public limited companies such as Aston Martin, whose shares are traded on the stock exchange and owned by thousands of shareholders, including pension funds, who play little part in running the business. The conduct of a firm refers to its behaviour, such as its objectives, pricing, investment and advertising decisions, while performance refers to the outcomes, such as profits, efficiency, growth, market share and share price. This creates a principal-agent problem. The shareholders (principals) want managers (agents) to act in their interests, which traditional theory assumes means profit maximisation, to give high dividends and a rising share price. However, managers may have different objectives, and because of asymmetric information, shareholders cannot easily monitor what managers are doing or whether profit could be higher.



This affects the conduct of firms because managers may pursue objectives other than profit maximisation. As shown in the diagram, a profit maximiser produces where $MC = MR$, at $Q1$ and price $P1$, earning the shaded supernormal profit. Managers whose salaries and status depend on the size of the firm may prefer revenue maximisation, producing where $MR = 0$, at the higher output $Q2$ and lower price $P2$, as Baumol suggested. Alternatively, they may aim for sales maximisation, producing the largest output possible while still covering costs,

where $AR = AC$, at $Q3$ and $P3$, earning only normal profit. Managers may also pursue growth through mergers and acquisitions, as larger firms bring higher pay and prestige, even if takeovers do not increase shareholder value. Some managers satisfice: they aim for a satisfactory level of profit to keep shareholders content, while pursuing their own goals, such as perks, large offices, or an easy life. The effect on performance is that profits are likely to be lower than they could be. A lack of pressure to minimise costs leads to X-inefficiency, with organisational slack, overstaffing and excessive spending on executive pay and perks, so the firm does not produce at the lowest possible cost and is productively inefficient. Rapid expansion to increase managers' status may lead to diseconomies of scale, such as poor communication and coordination. Aston Martin hoped to expand into new markets, but poor sales and weak profits meant the company had lost 75% of its value by February 2020, which shows how decisions by managers can damage shareholders' wealth. Managers may also focus on short-term results to meet targets and secure bonuses, cutting long-term investment and research.

However, there are reasons why the impact may not be significant. Owners can reduce the principal-agent problem by designing remuneration schemes that align the interests of managers with those of shareholders, such as paying managers in shares or share options and performance-related bonuses linked to profit or share price. Shareholders elect the board of directors and can vote at annual general meetings on key decisions, including executive pay, and large institutional shareholders actively monitor firms and can replace poorly performing managers. The threat of takeover also limits managers' freedom. If managers fail to maximise profits, the share price will fall, making the firm cheap for another firm to buy, after which the managers are likely to lose their jobs. Competition in the product market has a similar effect: in a competitive market, a firm that does not minimise costs will make losses and be forced out, so managers have little scope to pursue their own objectives. Divorce of ownership from control matters most for firms with monopoly power and many dispersed shareholders.

There may even be benefits. Professional managers may have greater expertise

than owners, which improves performance, and a long-term focus on growth and market share may raise future profits. A firm pursuing sales maximisation charges a lower price and produces a higher output, which benefits consumers and moves output closer to the allocatively efficient level. Pursuing growth can generate economies of scale, lowering costs. In contrast, family-owned firms and small businesses, where owners control the firm, avoid the principal-agent problem but may lack expertise and finance.

The extent to which divorce of ownership from control affects conduct and performance therefore depends on several factors: how dispersed the ownership is, how well managers' pay is linked to performance, the degree of competition in the market, how active shareholders are and the threat of takeover. It also depends on the time period: managers may be able to pursue their own objectives in the short run, but if profits remain poor shareholders will eventually act, as when a new investor consortium took a major stake in Aston Martin in 2020 and replaced its chief executive.

In conclusion, the divorce of ownership from control can significantly affect the conduct of firms, leading managers to pursue revenue, sales or growth maximisation and satisficing rather than profit maximisation, and this can reduce performance through X-inefficiency, diseconomies of scale and lower profits.

However, the effect is limited in practice by incentive schemes, shareholder voting, the threat of takeover and competition. The impact is likely to be greatest in large firms with dispersed shareholders and significant market power, and smallest where markets are competitive and managers' rewards are closely tied to the firm's results.