Cambridge Assessment International Education

Cambridge International AS & A Level

ECONOMICS **9708/41**

Paper 4 Data Response and Essays **May/June 2021**

2 hours 15 minutes

0010104515

You must answer on the enclosed answer booklet.

You will need: Answer booklet (enclosed)

INSTRUCTIONS

- Answer **three** questions in total:
  Section A: answer Question 1.
  Section B: answer **two** questions.
- Follow the instructions on the front cover of the answer booklet. If you need additional answer paper, ask the invigilator for a continuation booklet.
- You may use a calculator.
- You may answer with reference to any economy you have studied where relevant to the question.

INFORMATION

- The total mark for this paper is 70.
- The number of marks for each question or part question is shown in brackets [ ].

This document has **4** pages. Any blank pages are indicated.

DC (KS) 202768/1
© UCLES 2021

[Turn over

stemup.co.uk

PapaCambridge

Section A

Answer this question.

1 Public services – better or worse?

A government usually has macroeconomic policy aims that it hopes will enable some success in raising the quality of life, or well-being, of the population. In doing this the government also hopes to achieve efficiency in the use of resources.

Efficiency is measured by relating inputs to outputs. Inputs are relatively easy to count: they are financial costs of public services. Output can also be counted, but it is not necessarily a good measure of the outcome. The outcome is much harder to calculate. It is broader and more subjective – how do we assess whether a public service is 'better' or 'worse'?

This is where a knowledge of well-being makes an enormous difference. It could provide a much clearer view of the trade-offs that have to be made in allocating taxpayers' money to public services. Take the example of healthcare. To maximise the impact of expenditure on well-being, the budget may need to be adjusted to give more to mental health services and less to building general hospitals. For older people it could mean giving priority to programmes that would keep them out of hospital.

A focus on well-being should lead to better outcomes. This is where policymakers need a better understanding of behavioural economics. Governments have established Behavioural Insights teams, or Nudge Units. They have had some success. A small change in the wording of a letter to people who owed tax demonstrated how more behaviourally sensitive language sped up payments. The unit also found that jobseekers were nearly twice as likely to turn up for a job fair if the text message from the job centre used their names, and nearly three times as likely if the person sending the text message added 'good luck'.

Do tax reliefs persuade people to save? No. So enrol them instead automatically in a pensions programme as a default position, with the possibility of opting out. Allowing people to learn from mistakes is good: it reduces dependency on the public sector and helps people make better decisions for themselves. But some errors, such as failing to save anything until you are too old to earn, cannot be reversed. Then an early 'nudge' is justified. It has proved successful in spreading the habit of saving for retirement into groups not persuaded by tax reliefs alone.

In the long term this, and similar behavioural changes, may well have more influence on well-being than can be represented by concentrating on a monetary calculation of GDP. Other economic indicators could be used to assess this change in well-being.

Source: RSA Issue 1, 2017

(a) The article refers to macroeconomic policy aims. Identify and explain **two** such macroeconomic policy aims. [4]

(b) Is there evidence in the article that a knowledge of behavioural economics can help public policy? [4]

(c) The article says that 'efficiency is measured by relating inputs to outputs'. Is this how economic theory states that efficiency is determined? [5]

(d) The article deals with an improvement in well-being. Discuss whether there are any economic indicators that could be used to assess whether well-being has become better or worse. [7]

© UCLES 2021

stemup.co.uk

PapaCambridge

Section B

Answer **two** questions.

2 Assess the suggestion that a free market economy is neither possible nor desirable. [25]

3 **(a)** Explain what is meant by the concept of the 'equilibrium position of a consumer' and how the concept might be used to construct a demand curve for a good. [12]

(b) Distinguish between the income and substitution effects of a change in a good's price and analyse why the effect of a change in price is not always the same for different goods. [13]

4 **(a)** What costs ought a profit-maximising firm take into consideration when making decisions about price and output? [12]

(b) 'Price discrimination is always possible but never desirable.'

Do you agree with this opinion? [13]

5 'Wage differentials can be explained by economic theory. They are a sign of the power of a firm to exploit its workforce and are unjust.'

Do you agree with this view? [25]

6 **(a)** Distinguish between a country's national debt and its public sector budget deficit and consider which is the more important. [12]

(b) Discuss the effectiveness of alternative macroeconomic policies used to reduce a public sector budget deficit. [13]

7 **(a)** Explain and critically evaluate the quantity theory of money. [12]

(b) Monetary policy relies heavily on the theory of a monetary transmission mechanism.

Explain how a monetary transmission mechanism works and discuss its effectiveness. [13]

© UCLES 2021

stemup.co.uk

PapaCambridge

BLANK PAGE

Permission to reproduce items where third-party owned material protected by copyright is included has been sought and cleared where possible. Every reasonable effort has been made by the publisher (UCLES) to trace copyright holders, but if any items requiring clearance have unwittingly been included, the publisher will be pleased to make amends at the earliest possible opportunity.

To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced online in the Cambridge Assessment International Education Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download at www.cambridgeinternational.org after the live examination series.

Cambridge Assessment International Education is part of the Cambridge Assessment Group. Cambridge Assessment is the brand name of the University of Cambridge Local Examinations Syndicate (UCLES), which itself is a department of the University of Cambridge.

© UCLES 2021

stemup.co.uk

PapaCambridge

Surname	StemUp Model Solution
Answer booklet	Write the question number in the left-hand margin at the start of each answer.

Question number

Answers for paper 9708/41

1(a) Low and stable inflation: the government aims to keep the general price level rising slowly and predictably, for example a target of 2% a year. This protects the purchasing power of incomes and savings, reduces uncertainty for firms planning investment and keeps exports internationally competitive.

Low unemployment (full employment): the government aims to keep the percentage of the labour force who are willing and able to work but have no job as low as possible. This means scarce labour resources are not wasted, output and incomes are higher, and spending on unemployment benefits falls, which raises well-being.

1(b) Yes, there is some evidence. A small change in the wording of a letter to people who owed tax sped up payments, so tax revenue was collected sooner, which helps to reduce the government's budget deficit. Jobseekers were nearly twice as likely to attend a job fair when a text used their name, and nearly three times as likely when it added 'good luck', which could help reduce unemployment.

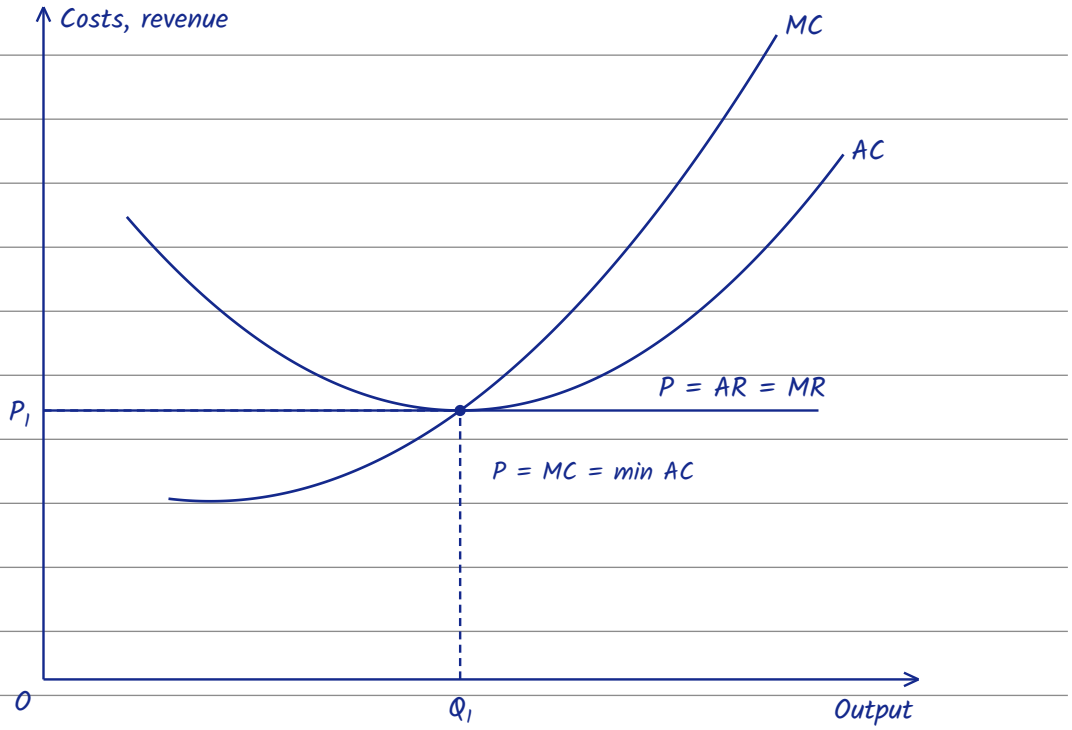
Automatic enrolment in a pension scheme, with the option to opt out, spread the habit of saving to groups not persuaded by tax reliefs. Higher saving reduces consumption, which could help to limit demand-pull inflation, and reduces dependency on the state in old age.

However, the evidence is limited: the article only says the units had 'some success', it gives no figures for the actual increase in tax paid, jobs found or savings, and going to a job fair is not the same as getting a job. So the evidence suggests behavioural economics can help, but its overall impact is not proven.

Question number

1(c)

Only partly. Relating inputs to outputs is how economic theory measures productive (technical) efficiency: a firm or public service is productively efficient when it produces its output at the lowest possible average cost, the minimum point of the average cost curve, so the fewest scarce resources are used for each unit of output. However, economic theory also defines allocative efficiency, which occurs where price equals marginal cost (P = MC). Here resources are allocated to producing the goods and services that consumers value most, so the value of the last unit to society equals the cost of the resources used to make it. This depends on the outcome for consumers, not just on the quantity of output, which is exactly the problem the article raises: output is 'not necessarily a good measure of the outcome'. A hospital could treat more patients cheaply and still be allocatively inefficient if money would raise well-being more in mental health services.



The diagram shows a firm in long-run perfect competition achieving both: P = MC (allocative) and production at minimum AC (productive). Theory also uses Pareto efficiency, where no one can be made better off without making someone else

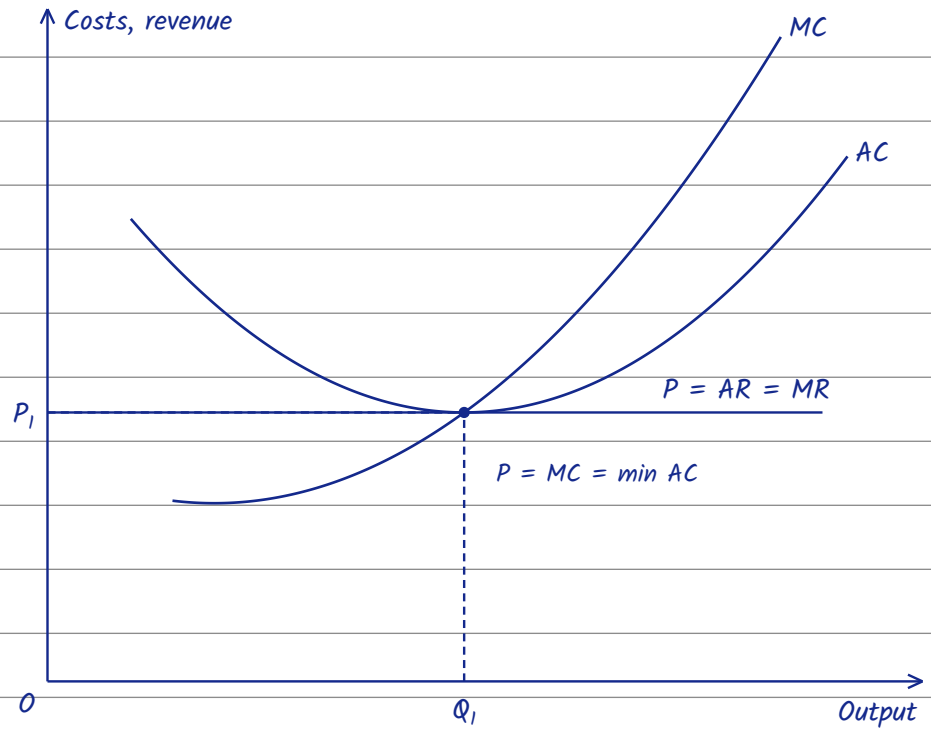
Question number	
	worse off, and dynamic efficiency over time. So the article's measure covers only one part of what economists mean by efficiency.
1(d)	Real GDP per head is the most widely used indicator. A rise means more goods and services per person, so material living standards usually improve. However, it ignores the distribution of income, leisure time, unpaid work, the informal economy and environmental damage, and the article itself warns that a 'monetary calculation of GDP' may not represent changes in well-being. The Human Development Index (HDI), published by the UNDP, is broader. It combines three dimensions: health (life expectancy at birth), education (mean years and expected years of schooling) and income (GNI per head at purchasing power parity), into an index from 0 to 1. A rising HDI shows that people are living longer and are better educated as well as richer, which is closer to well-being. But it uses only three measures, gives them equal weight and ignores inequality, mental health and political freedom. The Measure of Economic Welfare (MEW), developed by Nordhaus and Tobin, adjusts GDP by adding the value of leisure and of unpaid household work and subtracting 'regrettables' such as defence and commuting costs and the cost of environmental damage. This makes it more relevant to well-being, but the values of leisure and pollution are subjective estimates, so it is hard to calculate and rarely published. Other indicators include subjective measures of happiness and life satisfaction from surveys, and the Index of Sustainable Economic Welfare. These measure well-being directly, but answers depend on expectations and culture, so they are hard to compare between countries. In conclusion, there are useful indicators, but no single one fully measures well-being. A combination of real GDP per head with HDI and well-being surveys gives the best assessment of whether it is getting better or worse.

Question number

2

A free market economy is one in which resources are allocated entirely by the price mechanism, through the interaction of demand and supply, with private ownership of resources and no government intervention. Prices act as signals, incentives and rationing devices: if consumers want more of a good, its price rises, profits rise, and firms move resources into producing it.

Supporters argue it is desirable because it leads to an efficient allocation of resources. Productive efficiency is producing at the lowest point on the average cost curve, so no resources are wasted. Allocative efficiency is where P = MC, so resources produce exactly the goods consumers value most. In perfect competition, freedom of entry and exit means that in the long run supernormal profits are competed away and both are achieved.

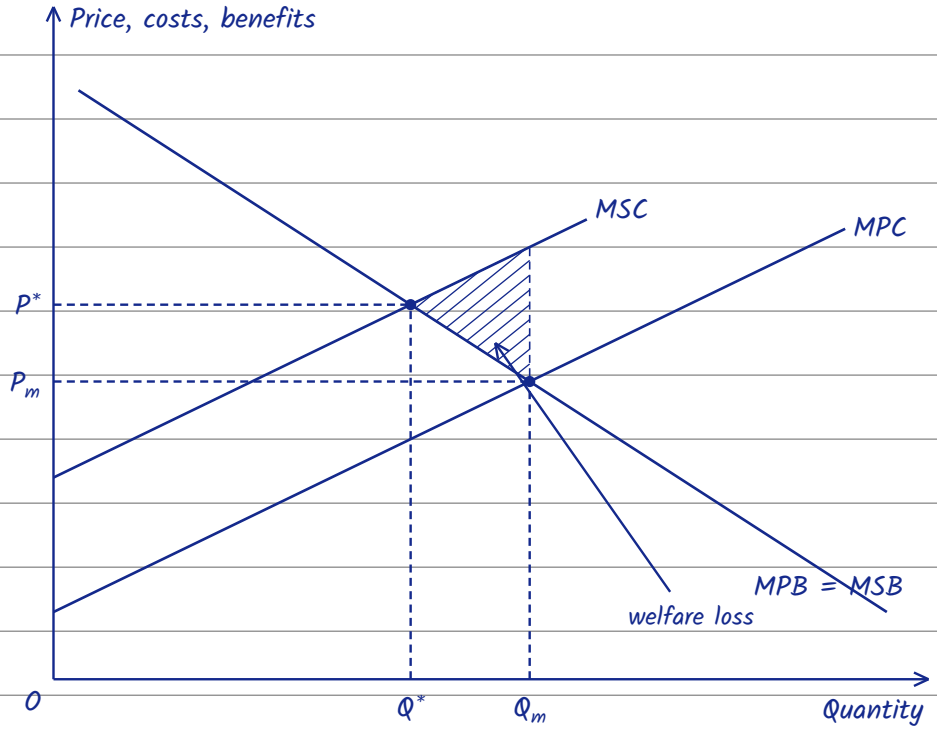


The diagram shows long-run equilibrium at Q1, where P = MC and P = minimum AC. Competition also gives consumer sovereignty and choice, and the profit motive creates incentives to innovate, which brings dynamic efficiency. Decisions are made by millions of individuals who have better information about their own wants than

Question number

a government planner, so a free market avoids the bureaucracy and shortages of command economies.

However, a free market is not desirable because individual actions are not always best for society as a whole. When production creates negative externalities such as pollution, firms consider only their private costs, so the market produces where MPC = MPB at Qm, more than the socially optimal output Q^* where MSC = MSB. The shaded triangle is the welfare loss, so resources are misallocated.



Merit goods such as education and healthcare have positive externalities and consumers lack information about their full benefits, so they are under-consumed, while demerit goods such as tobacco are over-consumed. A free market also tends to produce monopolies: a monopolist restricts output and sets price above MC, so it is allocatively inefficient. Finally, the market distributes income according to the ownership of factors of production, so it can create great inequality and leave people without access to basic needs, which many would see as unfair.

A pure free market is also not possible. Public goods such as defence, street lighting and flood defences are non-excludable and non-rival, so the free rider

Question number

problem means no private firm can charge for them and they would not be provided at all; this is a missing market. Markets themselves depend on government: they need a legal system to define and enforce property rights and contracts, a police force, and a stable currency, all provided by the state. This is why every real economy, even those close to a free market such as Hong Kong or the USA, is in practice a mixed economy.

On the other hand, government intervention can lead to government failure. Governments lack full information about the size of external costs and benefits, so taxes and subsidies may be set at the wrong level; intervention has administrative costs, may be influenced by political pressures and can create unintended consequences, such as black markets after maximum prices. Nationalised industries may lack the incentive to cut costs. So intervention will only improve efficiency if the welfare gain exceeds the cost of the government failure.

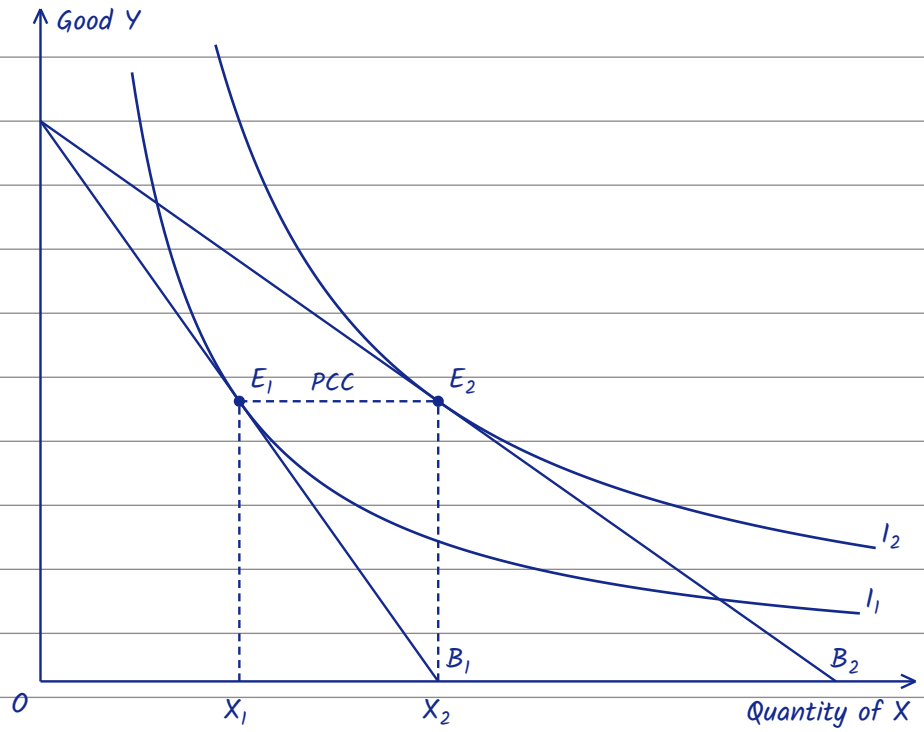
In conclusion, the suggestion is largely correct. A completely free market is not possible, because public goods and the legal framework that markets rely on must be provided by the state, and it is not desirable, because of externalities, merit goods, monopoly power and inequality. However, this does not mean the market mechanism itself is undesirable: it is the most efficient way to allocate most goods, so the best outcome is a mixed economy in which the market allocates most resources and the government intervenes only where there is clear market failure.

3(a) A consumer is in equilibrium when they have allocated their limited income between goods so that total utility (satisfaction) is maximised, so there is no incentive to change their spending. Using marginal utility theory, and the law of diminishing marginal utility, this happens when the marginal utility per $ spent is the same for every good (the equi-marginal principle): $MUx \div Px = MUy \div Py$, and all income is spent.

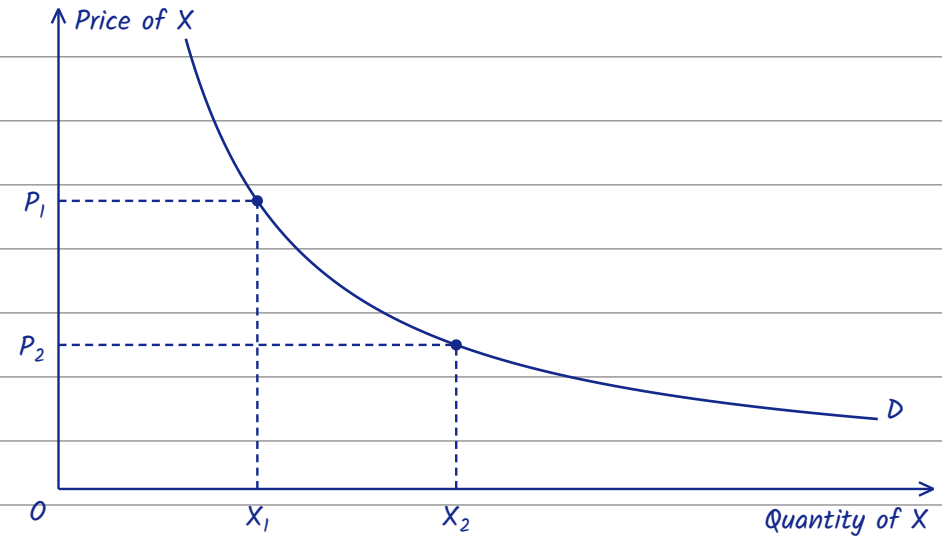
Question number

This can be used to build a demand curve. Suppose the price of X falls. Now MUx ÷ Px > MUy ÷ Py, so the last $ spent on X gives more satisfaction than the last $ spent on Y. The consumer buys more X and less Y. As they buy more X, its marginal utility falls, until MUx ÷ Px = MUy ÷ Py again at a larger quantity of X. So a lower price gives a larger quantity demanded, and plotting each price against the equilibrium quantity gives a downward sloping demand curve.

Using indifference curve analysis: an indifference curve shows combinations of two goods that give equal satisfaction; it is convex to the origin because of a diminishing marginal rate of substitution. The budget line shows the combinations the consumer can afford, with slope equal to the price ratio Px ÷ Py. The consumer is in equilibrium where the budget line is tangent to the highest attainable indifference curve, so MRS = Px ÷ Py.



Question number



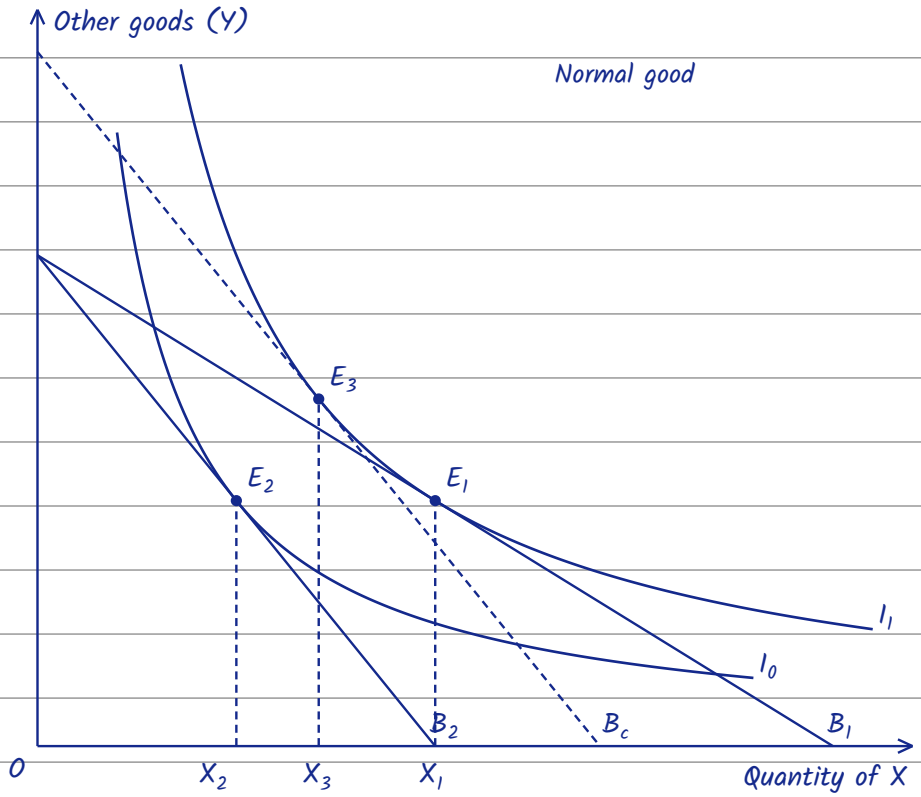
In the first diagram, with price P1 the budget line is B1 and the consumer is in equilibrium at E1 on I1, buying X1. When the price of X falls to P2, the budget line pivots outwards about the Y intercept to B2, because the same income now buys more X. The new equilibrium is E2 on the higher indifference curve I2, buying X2. Joining the equilibrium points gives the price consumption curve (PCC). Each price and the equilibrium quantity is then plotted in the second diagram: P1 with X1 and P2 with X2. Repeating this for every possible price traces out the individual's demand curve D, which slopes downwards for a normal good.

3(b) When the price of a good changes, the change in quantity demanded can be split into two effects. The substitution effect is the change caused only by the change in relative prices, holding real income constant: a good that becomes relatively more expensive is replaced by substitutes, so the substitution effect is always in the opposite direction to the price change. The income effect is the change caused by the change in real income (purchasing power): a price rise makes the consumer worse off in real terms. Its direction depends on the type of good.

In the diagrams, a rise in the price of X pivots the budget line inwards from B1 to B2. The compensated budget line Bc is drawn parallel to B2 (new relative prices) but tangent to the original curve I1. The movement E1 to E3 is the substitution

Question number

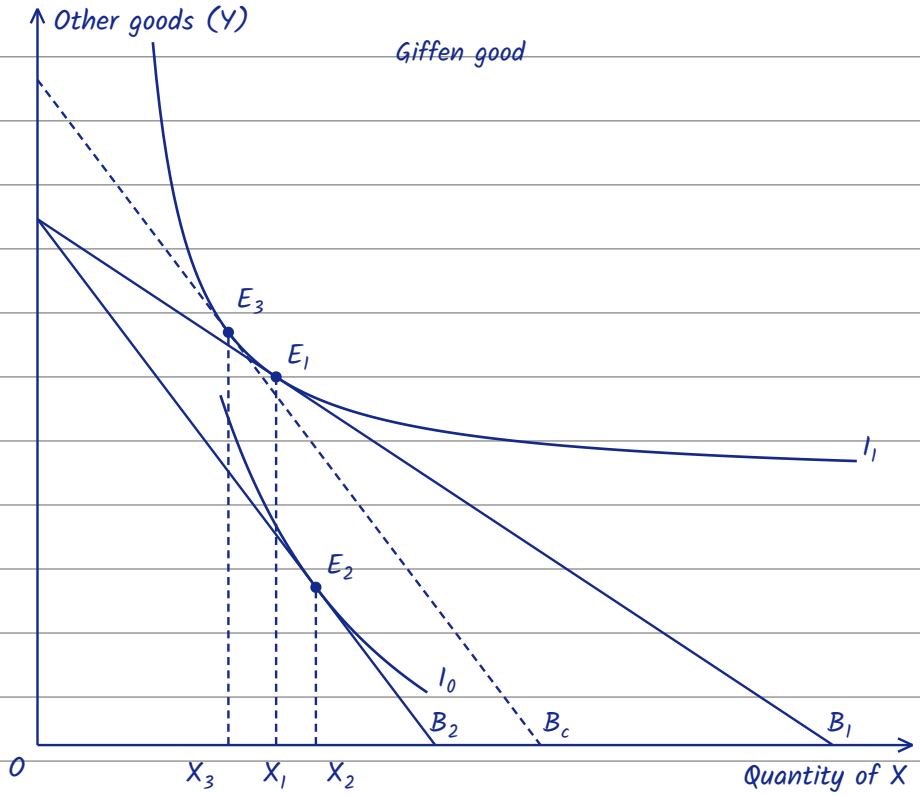
effect and the parallel shift from Bc to B2, E3 to E2, is the income effect.



For a normal good, demand falls when real income falls, so the income effect reinforces the substitution effect: demand falls from X1 to X3 and then further to X2. The demand curve slopes downwards. How much demand falls depends on the price elasticity of demand: if there are close substitutes, the substitution effect is large and demand is elastic; for a necessity with few substitutes, demand changes only a little.

For an inferior good, demand rises when real income falls, so the income effect works against the substitution effect. If, as usual, the substitution effect is larger, demand still falls when price rises, but by less than for a normal good, so demand is less price elastic.

Question number



A Giffen good is an inferior good where the positive income effect is larger than the substitution effect. In the diagram, the substitution effect reduces demand from X1 to X3, but the fall in real income increases demand from X3 to X2, so demand ends up higher than the original X1. A rise in price therefore increases the quantity demanded, and the demand curve slopes upwards. This happens with a staple food that takes a large share of a very poor household's income, because when its price rises the household can no longer afford better foods and buys more of the staple.

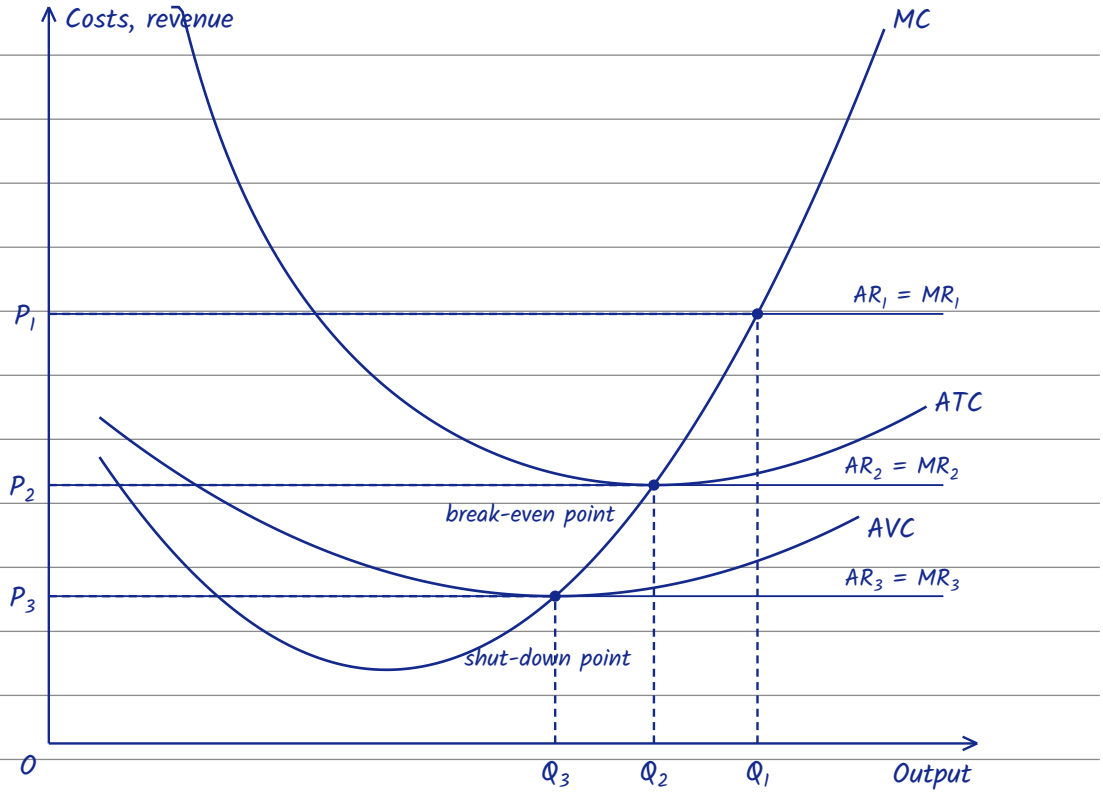
So the effect of a price change is not the same for all goods because the size and direction of the income effect differ between normal, inferior and Giffen goods, and the size of the substitution effect depends on the availability of substitutes. The difference is also greater when the good takes a large share of income, since the income effect is then larger. In practice Giffen goods are very rare, so for almost all goods a price rise reduces demand, differing only in how much.

Question number

4(a)

A firm's costs can be divided into fixed costs, which do not change with output in the short run (rent, interest on loans), and variable costs, which rise with output (raw materials, wages of hourly workers). Total cost = fixed cost + variable cost. Average total cost (ATC) is total cost ÷ output, average variable cost (AVC) is variable cost ÷ output, and marginal cost (MC) is the change in total cost from producing one more unit. Economists also include opportunity cost, so normal profit is counted as a cost.

Marginal cost is the key cost for the output decision. A profit-maximising firm produces where MC = MR, with MC rising and cutting MR from below. If MR > MC, producing one more unit adds more to revenue than to cost, so profit rises if output increases; if MC > MR, profit rises if output is cut. The price is then set from the average revenue (demand) curve at that output; in perfect competition the firm takes the market price.



Average total cost determines whether the firm makes a profit and is key in the long run. At price P1, the firm produces Q1 where MC = MR1, and because AR is

Question number

above ATC it makes supernormal profit. At P2, AR = ATC at the break-even point and only normal profit is made. In the long run all costs are variable, so the firm must cover ATC: if price stays below ATC it will leave the industry.

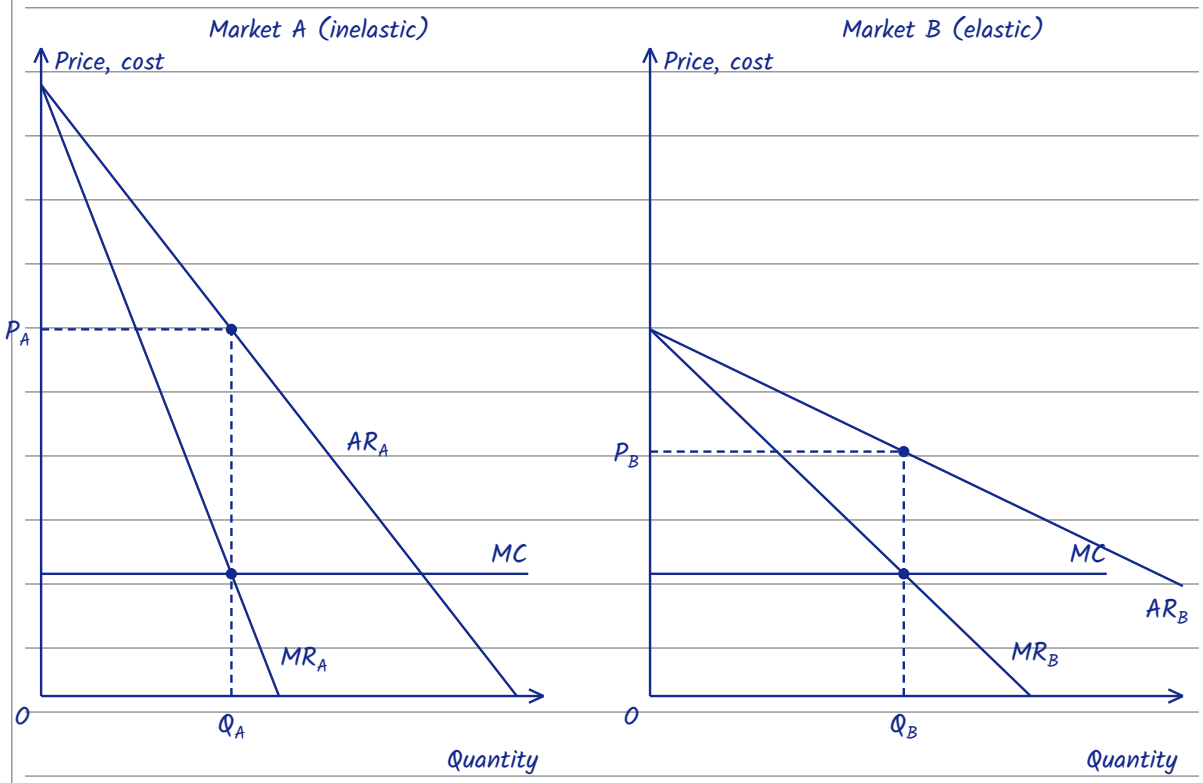
Average variable cost is key in the short run. Fixed costs must be paid whether or not the firm produces, so they do not affect the decision to produce in the short run. As long as price is at least equal to AVC, as at P3, the firm covers its variable costs and makes some contribution to fixed costs, so it should keep producing even though it makes a loss. If price falls below the minimum of AVC, the shut-down point, it loses less by closing down and paying only its fixed costs. So a profit-maximising firm should consider MC to choose output, AVC to decide whether to produce in the short run and ATC to decide whether to stay in the industry in the long run. In practice firms may not know their MC precisely, so many use average cost plus a mark-up when setting prices.

4(b)

Price discrimination is when a firm sells the same product to different consumers at different prices for reasons not related to differences in cost, for example peak and off-peak rail fares or student discounts.

It is not always possible. Three conditions are needed. First, the firm must have price-setting power, so it cannot happen in perfect competition, where firms are price takers. Second, the markets must be separated by time, age, location or income, and consumers must not be able to resell the product from the cheap market to the expensive one; this is easier for services such as haircuts and travel than for goods. Third, the price elasticity of demand must differ between the markets, and the cost of keeping them separate must be lower than the extra revenue gained.

Question number



The diagram shows third-degree price discrimination with constant MC. The firm maximises profit where MR = MC in each market. In market A demand is price inelastic, so it charges a high price PA; in market B demand is elastic, so it charges a lower price PB. Total profit is higher than with a single price.

It is often seen as undesirable. Consumers in the inelastic market pay more and lose consumer surplus, which becomes extra producer surplus, so income is redistributed from consumers to producers. Under first-degree discrimination all consumer surplus is taken. Higher profits may also be used to strengthen barriers to entry.

However, it is not 'never desirable'. Consumers in the elastic market, often those on lower incomes such as students and pensioners, pay a lower price than they would with a single price, and some who would otherwise be priced out can now buy the product, so output may increase. Price discrimination can also allow a firm to survive: a rural bus service or a rail company may find that at any single price AR is below ATC, but by charging different prices its total revenue covers total cost, so the service is provided rather than shut down. Extra profit can finance

Question number

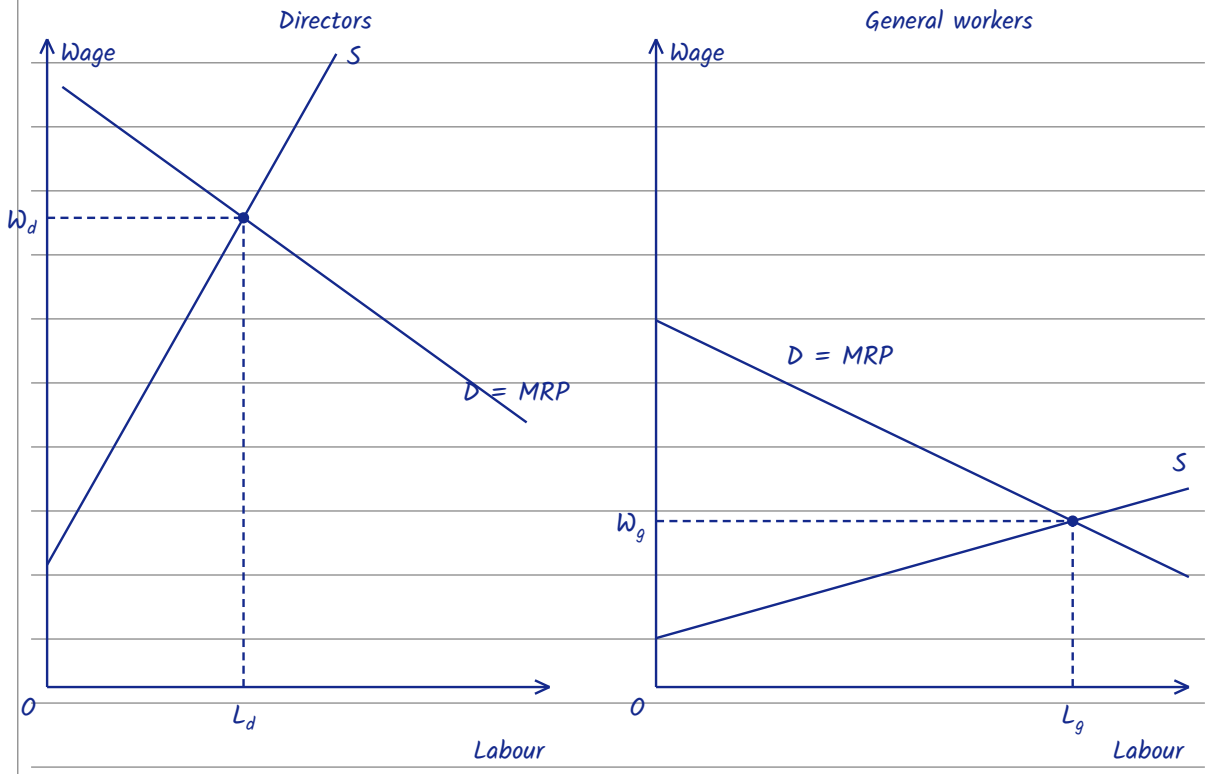
research and development, for example drug firms charging less in low-income countries. Under first-degree discrimination the firm produces up to the point where P = MC, the allocatively efficient output.

In conclusion, I disagree with the opinion. Price discrimination is not always possible, as it depends on market power, separate markets and different elasticities. It is not always undesirable either: whether it is desirable depends on whether output increases, who gains and loses, and whether it keeps a valuable service running. It is most harmful when a monopoly uses it simply to take consumer surplus.

5 A wage differential is the difference in wage rates between groups of workers. It can mean different wages for different occupations, such as doctors and cleaners, different wages for the same job in different firms, regions or sectors, or differences between workers with different skills, experience, gender or ethnicity. The view makes three claims: that theory explains differentials, that they reflect the power of firms to exploit workers, and that they are unjust. Each needs to be assessed.

Economic theory explains wages through demand and supply. The demand for labour is a derived demand, based on marginal revenue product (MRP = marginal physical product × marginal revenue). A profit-maximising firm employs workers up to the point where the wage equals MRP, so the MRP curve is the demand curve for labour. The supply of labour depends on the wage, the skills and training required, non-monetary benefits and the size of the working population. In a competitive labour market the wage is set where demand equals supply.

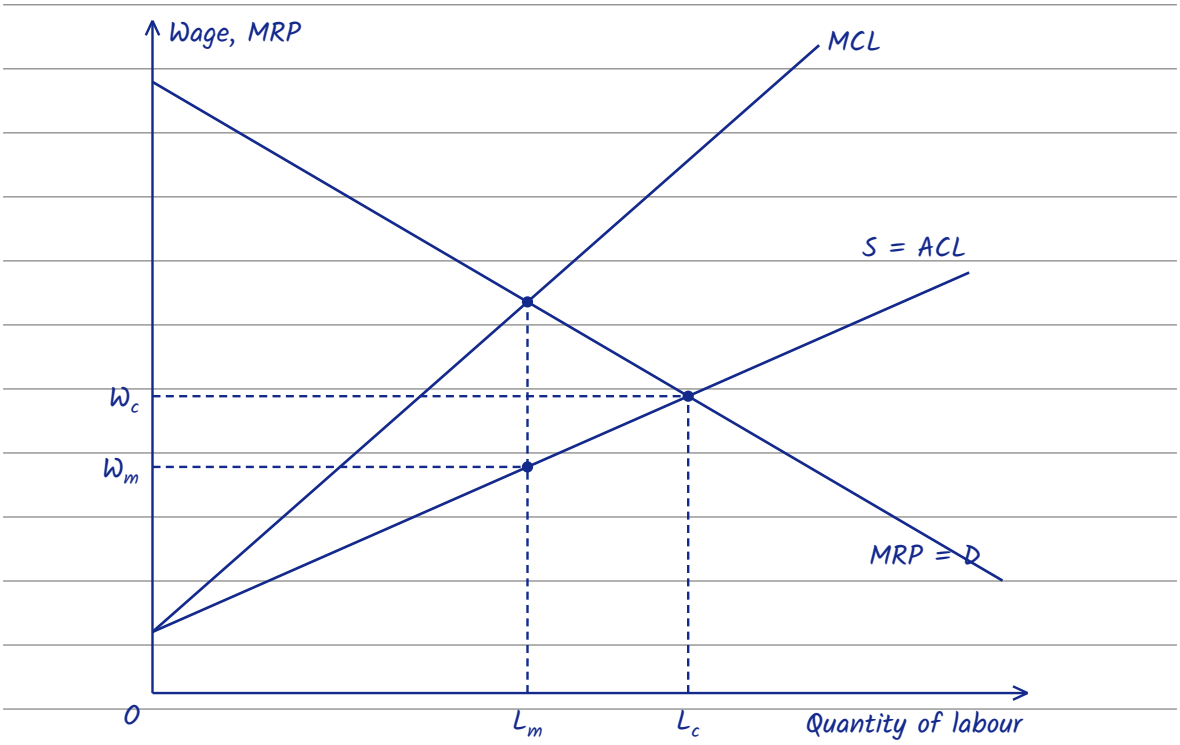
Question number



The diagram shows how this explains differentials between occupations. For skilled jobs such as company directors, MRP is high, and supply is low and inelastic because few people have the ability, qualifications and experience, so the wage Wd is high. For general workers MRP is lower and supply is large and elastic, so the wage Wg is low. The high wage includes a large element of economic rent, the payment above transfer earnings. Theory also explains compensating differentials, where dangerous or unpleasant jobs pay more, and differences between firms when productivity or the price of the product differ, since these change MRP.

The second claim is that differentials show the power of a firm to exploit its workforce. This is true where there is monopsony, a single or dominant buyer of labour, such as a large employer in a small town. To employ more workers the monopsonist must raise the wage for all of them, so the marginal cost of labour (MCL) is above the average cost of labour (the supply curve).

Question number



The monopsonist employs Lm where MCL = MRP, but pays only Wm, read from the supply curve. In a competitive market the wage would be Wc and employment Lc. So the monopsonist pays a lower wage and employs fewer workers, and workers are paid less than their MRP; this gap is what economists mean by exploitation of labour. So some differentials, for example between workers doing the same job for a monopsony employer and a competitive one, are due to the power of the firm. However, many differentials are not caused by firms' power. They come from differences in MRP and in labour supply, which reflect differences in skills, productivity and the demand for the product. Workers can also have power: strong trade unions can raise wages above the competitive level, and when a union faces a monopsonist (bilateral monopoly) the wage depends on bargaining strength. Wages in the public sector are often set by governments for political reasons, and minimum wage laws limit how low wages can go. Immobility of labour, imperfect information and discrimination also cause differentials that the simple model does not predict, so the theory explains differentials well in principle but needs these imperfections to explain real-world wages.

Question number

Whether differentials are unjust is a normative judgement. Differentials perform a useful economic function: they act as signals and incentives that move labour into jobs where it is most valued and reward people for the cost and time of training, so removing them would reduce efficiency. Differentials that reflect differences in productivity, skill or unpleasant conditions can therefore be seen as fair. But differentials caused by monopsony exploitation, discrimination by gender or ethnicity, or directors effectively setting their own pay are hard to justify, and very large differentials create inequality of income and wealth.

In conclusion, I only partly agree. Economic theory, using MRP and labour supply, does explain most wage differentials. Some, where firms have monopsony power, do show exploitation and may be unjust, but many result from differences in skills and productivity, or from the power of unions and governments, and these play an important role in allocating labour. So differentials are not always a sign of exploitation, nor always unjust.

6(a) The public sector budget deficit is a flow: it is the amount by which government spending exceeds government revenue (mainly taxation) in a particular time period, usually one year. The national debt is a stock: it is the total accumulated amount the government owes, the sum of all past borrowing not yet repaid, for example through selling government bonds to the private sector, banks and foreign investors.

The two are linked, because each year's deficit must be financed by borrowing, so it is added to the national debt. A budget surplus allows part of the debt to be repaid. Both are usually compared using a percentage of GDP, because a larger economy can afford to borrow more.

The budget deficit is important because it shows the current fiscal stance. A deficit injects demand into the economy, so it can help reduce unemployment in a recession, partly through automatic stabilisers. But a large deficit in a boom can

Question number

cause demand-pull inflation, financing it by borrowing may raise interest rates and crowd out private investment, and financing it by creating money can be inflationary. A deficit financed from abroad may also worsen the balance of payments.

The national debt is important because interest has to be paid on it every year. Debt interest has an opportunity cost, since the money could instead be spent on education or healthcare, and it shifts the burden to future generations, who face higher taxes. A high debt may lead to a lower credit rating, so lenders demand higher interest rates, and in extreme cases a debt crisis, as in Greece after 2010.

Which is more important depends on the situation. A large deficit for one or two years in a recession is not serious if the debt is low; the UK ran a large deficit in 2020 during the pandemic. In the long run, the debt matters more, since persistent deficits build up debt, and its sustainability depends on whether the interest rate is below the growth rate of GDP, so that the debt-to-GDP ratio can fall. It also matters whether the debt is owed in domestic currency to domestic lenders (as in Japan, with debt over 200% of GDP) or in foreign currency. Overall, the national debt relative to GDP is the more important indicator of long-term risk, while the deficit matters more for short-term macroeconomic management and shows the direction the debt is moving.

6(b)

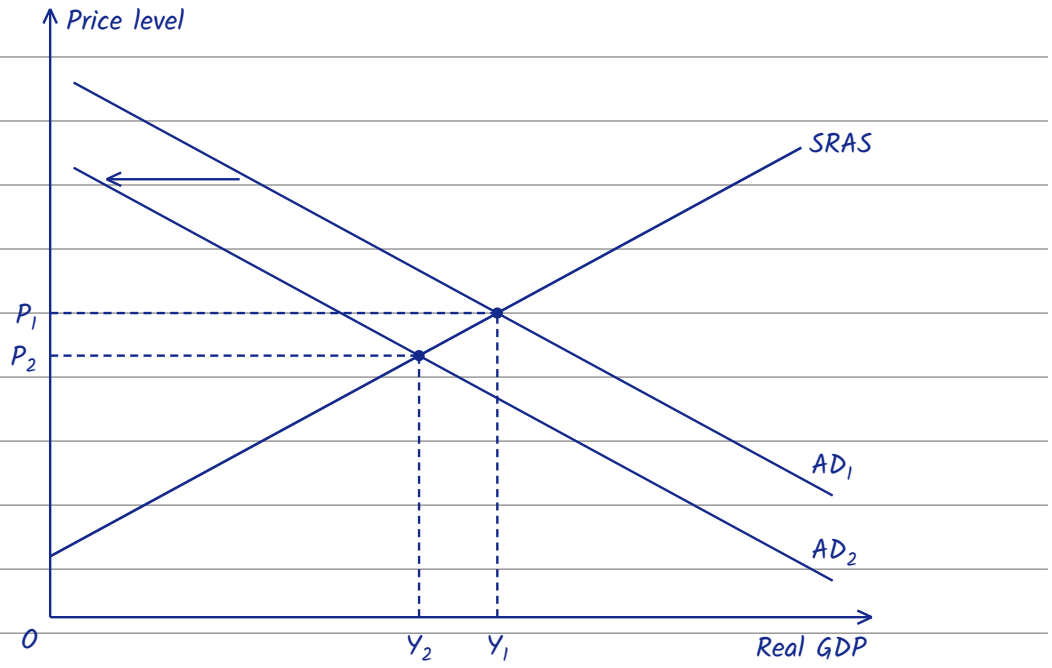
A government can reduce a budget deficit in the short run by raising taxes or cutting public spending, known as fiscal consolidation or austerity, and in the long run by policies that raise economic growth.

Raising taxes, such as income tax or VAT, increases revenue directly. Indirect taxes are quick to raise and hard to avoid. However, higher taxes reduce disposable income and consumption, so AD falls and, through the multiplier, national income falls by more than the initial change. Lower incomes reduce tax revenue and raise spending on benefits, so the deficit falls by less than expected. Very high income

Question number

tax rates may reduce the incentive to work and increase avoidance, and VAT is regressive, so it worsens inequality.

Cutting government spending reduces the deficit directly and may avoid the disincentive effects of higher taxes. Cuts to capital spending are politically easier but harm long-run growth, while cuts to current spending, such as public sector wages and benefits, are unpopular. Again, lower spending reduces AD.



The diagram shows that austerity shifts AD left from AD1 to AD2, so real GDP falls from Y1 to Y2 and unemployment rises. If the economy is already in recession, this is self-defeating, since automatic stabilisers increase the cyclical deficit. If the fiscal multiplier is small, or the economy is near full capacity, the cost in lost output is smaller and the policy is more effective.

Monetary policy can support fiscal consolidation: lower interest rates reduce the cost of servicing the national debt and encourage investment and consumption, offsetting the fall in AD. However, this is limited when interest rates are already close to zero.

In the long run, supply-side policies such as spending on education and training, infrastructure, and tax reforms that encourage enterprise can increase productive

Question number

capacity and shift LRAS to the right. Faster growth raises incomes and employment, so tax revenue rises and benefit spending falls without increasing tax rates. This is the least damaging way to reduce the deficit, but it is slow, its effects are uncertain and some policies initially increase spending.

In conclusion, the most effective approach depends on whether the deficit is cyclical or structural. A cyclical deficit falls automatically when growth returns, so cutting it in a recession is costly. A structural deficit needs tax rises or spending cuts, best introduced gradually once the economy is recovering, combined with supply-side policies for long-run growth and supportive monetary policy.

7(a) The quantity theory of money is based on the Fisher equation of exchange, $MV = PT$, where M is the money supply, V is the velocity of circulation (the number of times each unit of money is spent in a period), P is the general price level and T is the number of transactions (or real output, Y). MV is total spending and PT is the value of what is sold, so the equation is an identity: it is always true by definition.

The quantity theory turns this identity into a theory of inflation by making two assumptions. First, V is constant, or stable and predictable, because it depends on slowly changing institutional factors such as how often people are paid. Second, T is constant, because the economy tends towards full employment, so real output is determined by supply-side factors and does not depend on the money supply. It follows that any increase in M leads to a proportional increase in P: if the money supply rises by 10% faster than output, prices rise by 10%. Monetarists such as Milton Friedman used this to argue that inflation is always caused by excessive growth of the money supply, so it should be controlled by limiting money supply growth.

There is some evidence for the theory in the long run: in hyperinflations, such as Zimbabwe in 2008, very rapid money creation was followed by very rapid inflation.

Question number

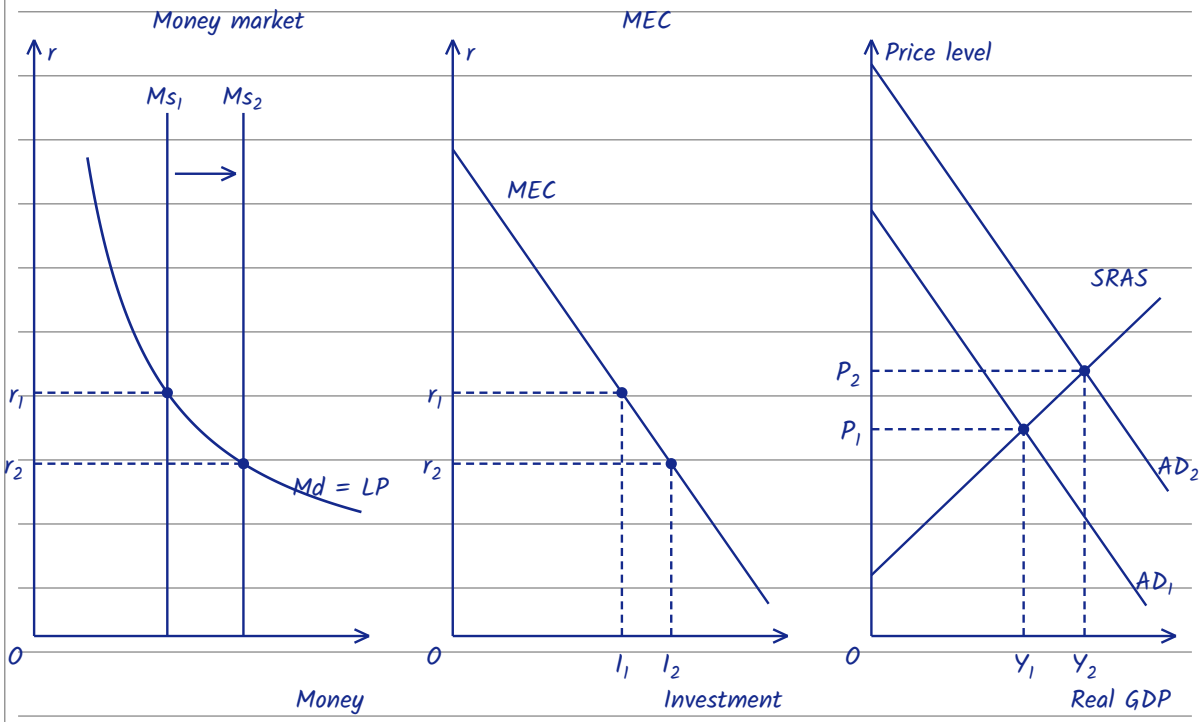
However, the assumptions can be criticised. Velocity is not constant: Keynesians argue that when the money supply rises, people may hold extra money as idle balances (the speculative demand for money), so V falls and offsets the rise in M. In many countries velocity fell sharply after quantitative easing from 2009, so huge increases in the money supply caused little inflation. Velocity also changes with financial innovation such as card and online payments.

T is not constant when there is spare capacity: if the economy is below full employment, extra spending increases real output rather than prices. Causation may also run the other way, from P to M: rising prices and wages lead firms and households to borrow more, so banks create more money in response.

Finally, the money supply is hard to measure, since there are several measures, narrow and broad, and they often move differently, and it is hard to control, because commercial banks create most money through credit creation. So the theory is a useful explanation of long-run inflation, but it is weak in the short run.

7(b) The monetary transmission mechanism is the process by which a change in the money supply or interest rates affects real variables such as output, employment and the price level. In the Keynesian version it links the money market to the goods market through the rate of interest.

Question number



Suppose the central bank increases the money supply from Ms1 to Ms2, for example by buying bonds. People now hold more money than they want, so they use the excess to buy bonds. This raises bond prices and, because bond prices and interest rates are inversely related, the interest rate falls from r1 to r2 along the liquidity preference (Md) curve. The lower interest rate reduces the cost of borrowing, so firms invest in projects whose expected return (marginal efficiency of capital) is now above the interest rate: investment rises from I1 to I2 along the MEC curve. Investment is an injection, so AD shifts from AD1 to AD2, and through the multiplier real GDP rises from Y1 to Y2 by more than the initial rise in investment, raising employment, with some rise in the price level from P1 to P2. Lower interest rates also raise consumption of durables bought on credit, raise asset prices (a wealth effect) and depreciate the exchange rate, increasing net exports.

However, each link can be weak. In a liquidity trap, when interest rates are already very low, the Md curve is almost horizontal, so extra money is held as idle balances and the interest rate does not fall; many economies faced this after 2008 with rates close to zero. If the MEC curve is interest inelastic, because

Question number

investment depends mostly on business expectations ('animal spirits'), a fall in interest rates causes only a small rise in investment; in a deep recession firms may not invest at all. The multiplier may be small if leakages, such as saving, taxes and imports, are high, as in a very open economy.

There is also a negative feedback effect: as income rises, the transactions demand for money increases, shifting Md to the right, which pushes interest rates back up and partly offsets the original fall. Commercial banks may not pass on lower rates or may be unwilling to lend, and the whole process has long and variable time lags, often up to two years.

In conclusion, the transmission mechanism explains how monetary policy can affect the real economy, and it works reasonably well in normal times when interest rates are well above zero and confidence is stable. It is much less effective in a deep recession, when a liquidity trap, low confidence and an inelastic MEC weaken the links, which is why governments have used quantitative easing and fiscal policy alongside it.

Cambridge Assessment International Education

Cambridge International AS & A Level

ECONOMICS **9708/42**

Paper 4 Data Response and Essays **May/June 2021**

2 hours 15 minutes

0420401845

You must answer on the enclosed answer booklet.

You will need: Answer booklet (enclosed)

INSTRUCTIONS

- Answer **three** questions in total:
  - Section A: answer Question 1.
  - Section B: answer **two** questions.
- Follow the instructions on the front cover of the answer booklet. If you need additional answer paper, ask the invigilator for a continuation booklet.
- You may use a calculator.
- You may answer with reference to any economy you have studied where relevant to the question.

INFORMATION

- The total mark for this paper is 70.
- The number of marks for each question or part question is shown in brackets [ ].

This document has **4** pages. Any blank pages are indicated.

DC (CJ) 202982/2
© UCLES 2021

[Turn over

stemup.co.uk

PapaCambridge

Section A

Answer this question.

1 Competition in passenger aircraft manufacturing

Two companies produce all of the world's passenger aircraft with a seating capacity of over 150 passengers. The supply of these is about 1600 aircraft a year and is divided almost equally between the two competing companies, who both benefit from extremely strong brand identity and customer loyalty.

The aircraft are made in two versions: single-aisle aircraft that carry 200 passengers and double-aisle aircraft that carry 300 to 350 passengers.

In 2018 both companies took over other firms which produced a range of aircraft seating between 40 and 150 passengers. The owners of the other firms had difficulties in financing and developing their new aircraft.

Recent competition in the aircraft manufacturing industry has aimed at improving fuel efficiency, developing longer non-stop flying ranges and developing very large aircraft with a capacity of 500 passengers. The companies have been successful in achieving the first two aims.

The third aim of developing very large aircraft was intended to move larger numbers of passengers between 'hub' airports, where they would transfer to smaller aircraft for onwards transport to regional airports. Whilst it has achieved increased passenger capacity, this development has not been a success for the airlines that bought these very large aircraft. The need for these was undermined by the success of the first two competitive developments.

The nature of the aircraft manufacturing industry is that it involves significant fixed costs. The research and development (R&D) costs of a new aircraft can rise to US$26 billion.

Currently the two companies are concerned that China, in its drive to increase the value of its exports, will enter the aircraft manufacturing industry in the next 10 to 15 years and add a new level of competition.

Source: adapted from The Economist, 5 February 2019 and 27 April 2019

(a) Using the information, explain the market structure that exists in aircraft manufacturing. [2]

(b) Explain why the two companies might be concerned about the potential entry of China into the aircraft manufacturing industry. [4]

(c) 'The nature of the aircraft manufacturing industry is that it involves significant fixed costs.'

(i) Using an example from the information, explain the meaning of fixed costs. [2]

(ii) With the aid of a diagram, explain the link between an increase in an aircraft manufacturer's level of output and its average fixed costs. [4]

(d) Discuss whether economic theory can explain how two companies will compete with each other when they dominate a market. [8]

© UCLES 2021

stemup.co.uk

PapaCambridge

Section B

Answer **two** questions.

2 **(a)** What is meant by 'efficiency' in relation to the use of resources? [12]

(b) 'Market failure can always be overcome to increase economic efficiency.'

How far do you agree with this statement? [13]

3 **(a)** Use indifference curve analysis to explain how an individual's demand curve for an inferior good is derived. [12]

(b) Discuss the relative importance of marginal cost and average variable cost in determining short-run production decisions. [13]

4 Evaluate whether the theory of wage determination can account for wage differentials (i) between a director and a general worker in the same company and (ii) between two workers doing the same job in different companies. [25]

5 Discuss the extent to which the Human Development Index (HDI) and the Measure of Economic Welfare (MEW) provide better measures of living standards than Gross National Income (GNI). [25]

6 **(a)** Explain the causes of unemployment and consider which is most likely to occur in a developing country. [12]

(b) Discuss whether solving the problem of unemployment should be the main policy objective for the government of a developing country. [13]

7 **(a)** Explain what is meant by quantitative easing (QE) and consider whether it is an effective policy to be used in a recession. [12]

(b) 'Policies to achieve economic growth will inevitably cause a government budget deficit.'

How far do you agree with this statement? [13]

© UCLES 2021

stemup.co.uk

PapaCambridge

BLANK PAGE

Permission to reproduce items where third-party owned material protected by copyright is included has been sought and cleared where possible. Every reasonable effort has been made by the publisher (UCLES) to trace copyright holders, but if any items requiring clearance have unwittingly been included, the publisher will be pleased to make amends at the earliest possible opportunity.

To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced online in the Cambridge Assessment International Education Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download at www.cambridgeinternational.org after the live examination series.

Cambridge Assessment International Education is part of the Cambridge Assessment Group. Cambridge Assessment is the brand name of the University of Cambridge Local Examinations Syndicate (UCLES), which itself is a department of the University of Cambridge.

© UCLES 2021

stemup.co.uk

PapaCambridge

Surname	StemUp Model Solution
Answer booklet	Write the question number in the left-hand margin at the start of each answer.

Question number

Answers for paper 9708/42

1(a) The market structure is an oligopoly, and more precisely a duopoly, because two companies dominate the whole market: they produce all of the world's passenger aircraft with over 150 seats, about 1600 a year, divided almost equally between them. There are high barriers to entry, such as R&D costs of up to US$26 billion and strong brand loyalty.

1(b) China's potential entry makes the market contestable. A new entrant, perhaps supported by its government in its drive to increase exports, could undercut the prices of the existing firms, which limits their ability to set high prices and so reduces their profits.

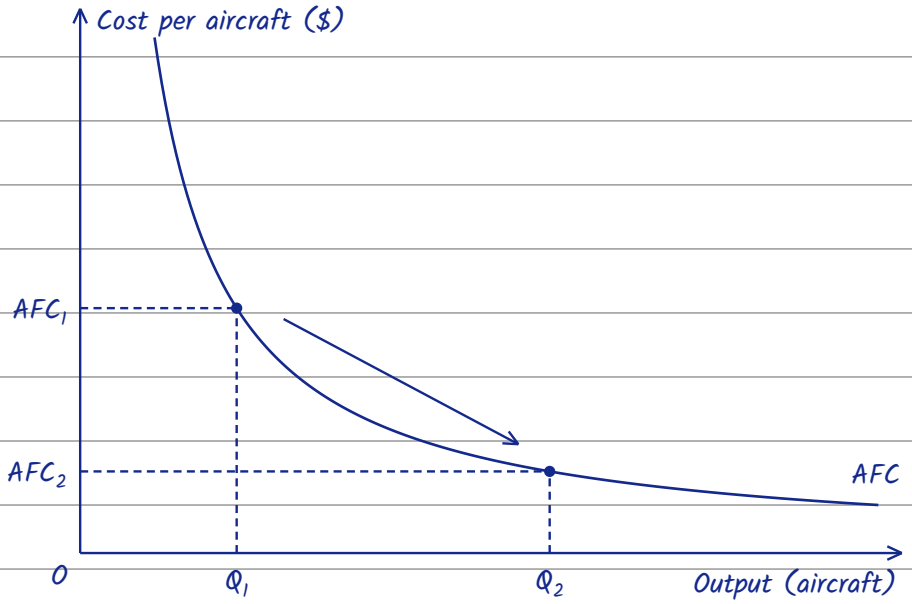
The presence of a third producer would also take customers away from the two existing firms, so demand for their aircraft and their market share would fall. To protect their position they may have to spend more on marketing and on research and development, which raises their costs and further reduces profit.

1(c)(i) Fixed costs are costs that do not change as the level of output changes in the short run; they must be paid even if output is zero. An example is the research and development cost of a new aircraft, which can rise to US$26 billion whether the firm builds 10 or 1000 of that aircraft.

1(c)(ii) Average fixed cost (AFC) = total fixed cost ÷ output. As the number of aircraft produced increases, AFC falls continuously, because the same total fixed cost is spread (divided) over a larger number of aircraft. For example, R&D of $26 billion

Question number

over 100 aircraft is $260 million per aircraft, but over 1000 aircraft it is only $26 million per aircraft.

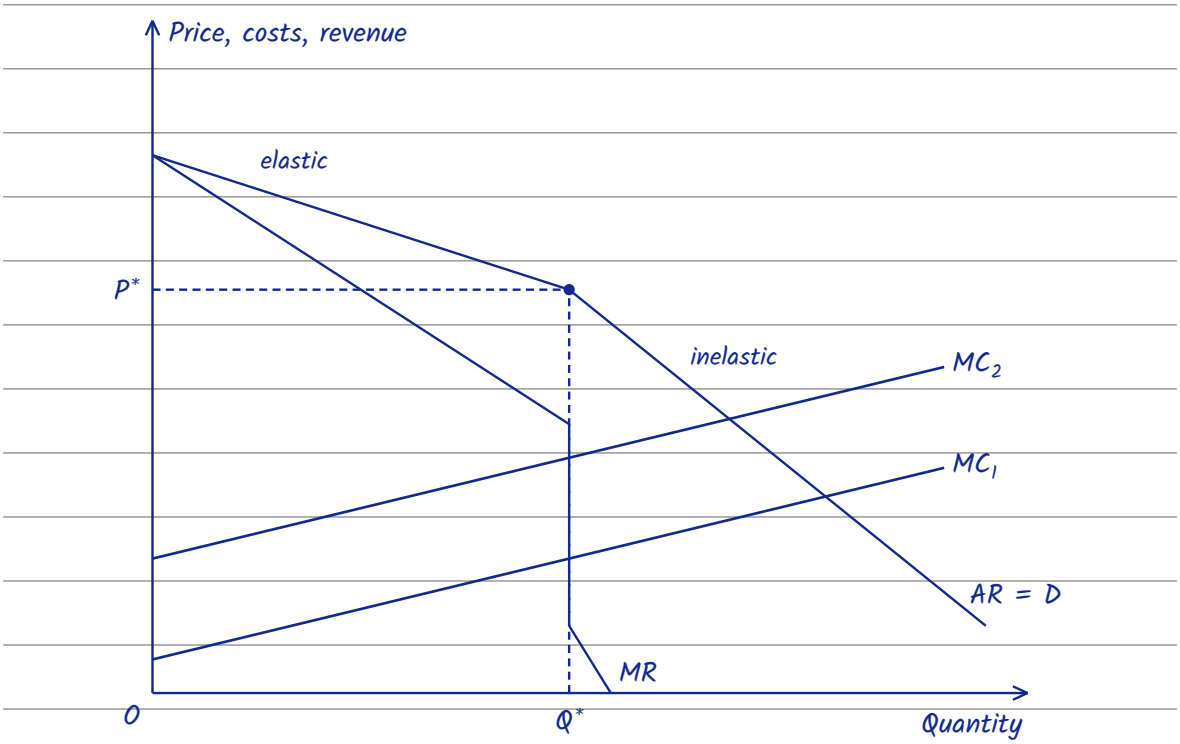


The AFC curve is a rectangular hyperbola: when output rises from Q1 to Q2, AFC falls from AFC1 to AFC2. It gets closer to the output axis but never reaches it, since fixed costs never fall to zero.

1(d) When two firms dominate a market (a duopoly), they are interdependent: each must consider how its rival will react before changing its price or output. Economic theory offers several explanations of how they compete.

The kinked demand curve model assumes that if one firm raises its price, the rival will not follow, so it loses many customers (demand is elastic above the current price); if it cuts its price, the rival will match the cut, so it gains few customers (demand is inelastic below the price).

Question number



This creates a kink at P* and a vertical gap in the MR curve, so MC can rise from MC1 to MC2 without changing the profit-maximising price and output. This explains why prices in oligopoly tend to be stable and why firms compete in other ways. However, it does not explain how P* was set in the first place.

Game theory can also be used. In this example payoff matrix (profits), each firm's dominant strategy is to set a low price whatever the other does, so both end up at $3bn each, although they would both earn more by keeping prices high. This prisoners' dilemma explains why there is pressure to collude, and why price wars can break out if collusion is illegal or breaks down.

	Firm B: high price	Firm B: low price
Firm A: high price	A $5bn, B $5bn	A $1bn, B $7bn
Firm A: low price	A $7bn, B $1bn	A $3bn, B $3bn

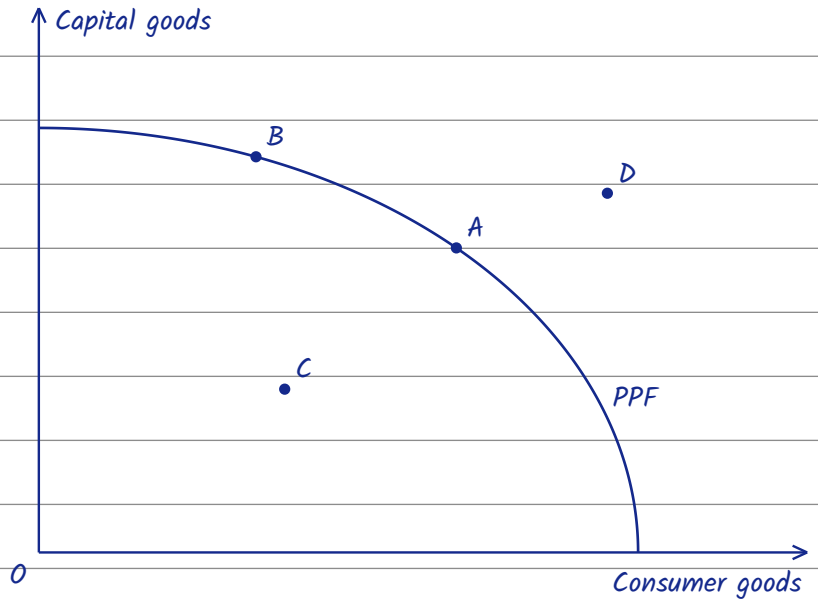
In practice, the extract shows that the two firms mainly use non-price competition: improving fuel efficiency, developing longer non-stop flying ranges and building very large aircraft, supported by strong brand identity. Theory predicts

Question number

this, because non-price competition is less easily copied than a price cut. The firms may also pursue other objectives, such as sales revenue maximisation or market share, rather than only profit. In conclusion, economic theory can explain the general pattern of competition well, especially price stability and non-price competition, but it cannot predict exactly how the firms will behave, since outcomes such as the failure of the very large aircraft depend on uncertain demand and on each firm's guesses about its rival.

2(a) Efficiency in the use of resources means making the best possible use of scarce resources (land, labour, capital and enterprise) to satisfy as many wants as possible. There are several types.

Productive efficiency is achieved when output is produced at the lowest possible average cost, the minimum point of the average cost curve. This means the maximum output is obtained from the resources used, so no resources are wasted. For the whole economy, it means producing on the production possibility frontier (PPF).

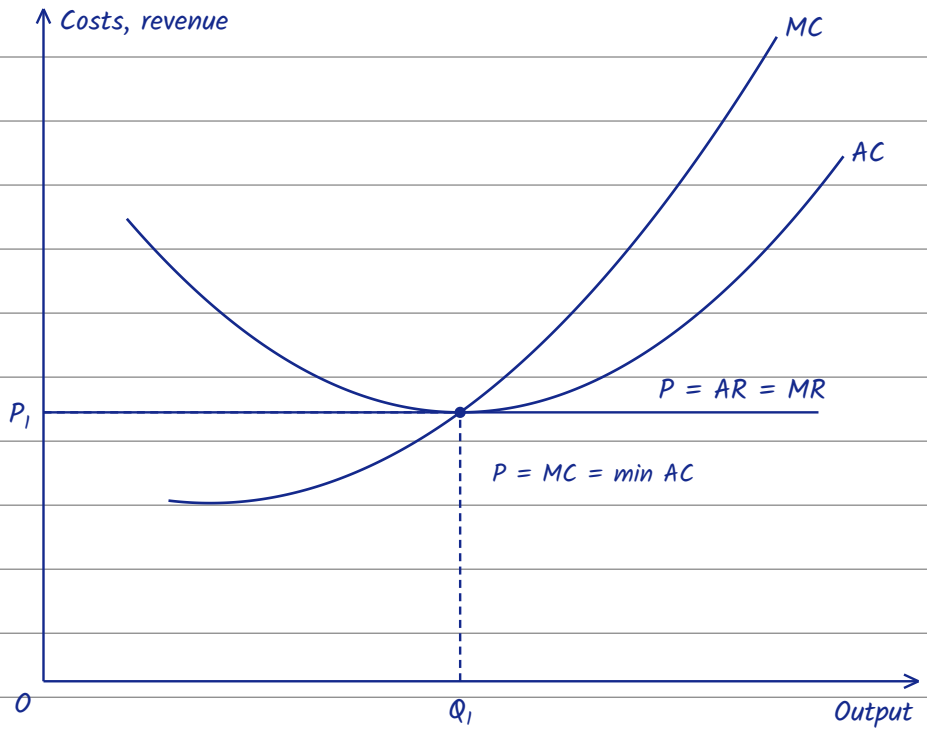


In the PPF diagram, points A and B are productively efficient because all resources are fully employed, while point C is inefficient, as resources are unemployed or

Question number

wasted. Point D is unattainable with current resources.

Allocative efficiency is achieved when resources are allocated to produce the combination of goods and services that consumers value most. This occurs where price equals marginal cost (P = MC), because the price shows the value consumers place on the last unit and MC shows the cost of the resources used to produce it. If P > MC, too few resources are allocated to the good; if P < MC, too many are. On the PPF only one point, say A, is allocatively efficient: the one that matches consumers' preferences.



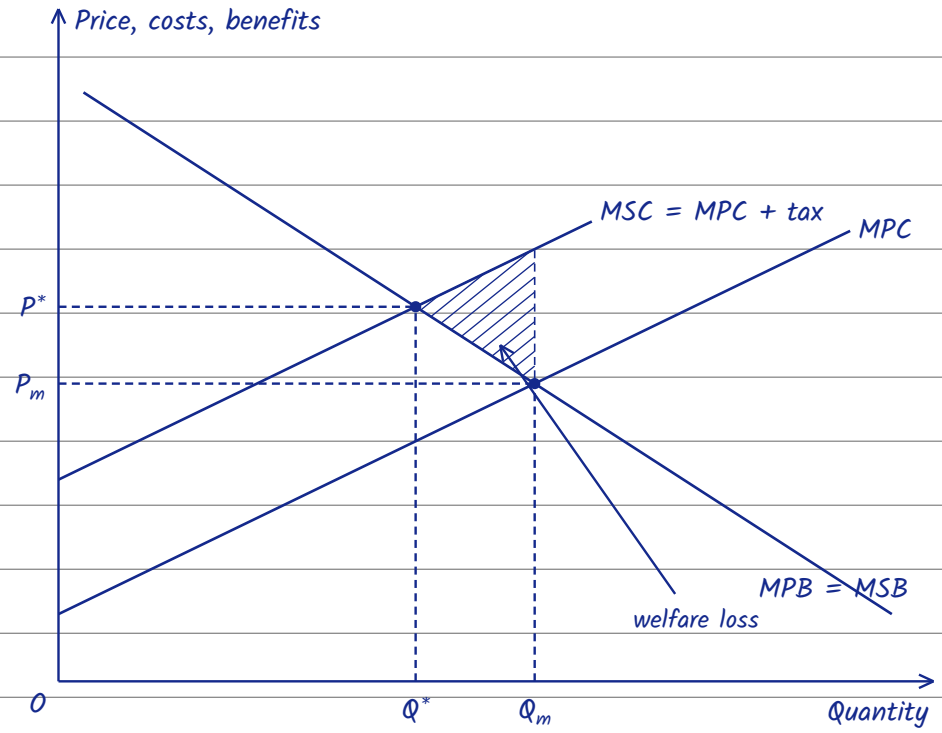
A perfectly competitive firm in long-run equilibrium achieves both, producing Q1 where P = MC and P = minimum AC. Pareto efficiency is when resources are allocated so that it is impossible to make one person better off without making someone else worse off; it requires both productive and allocative efficiency.

Dynamic efficiency is about improvements over time: using resources for research, innovation and new technology, which lower costs and improve products in the future and shift the PPF outwards. A monopoly may be dynamically efficient even though it is not statically efficient, because supernormal profits can fund R&D.

Question number

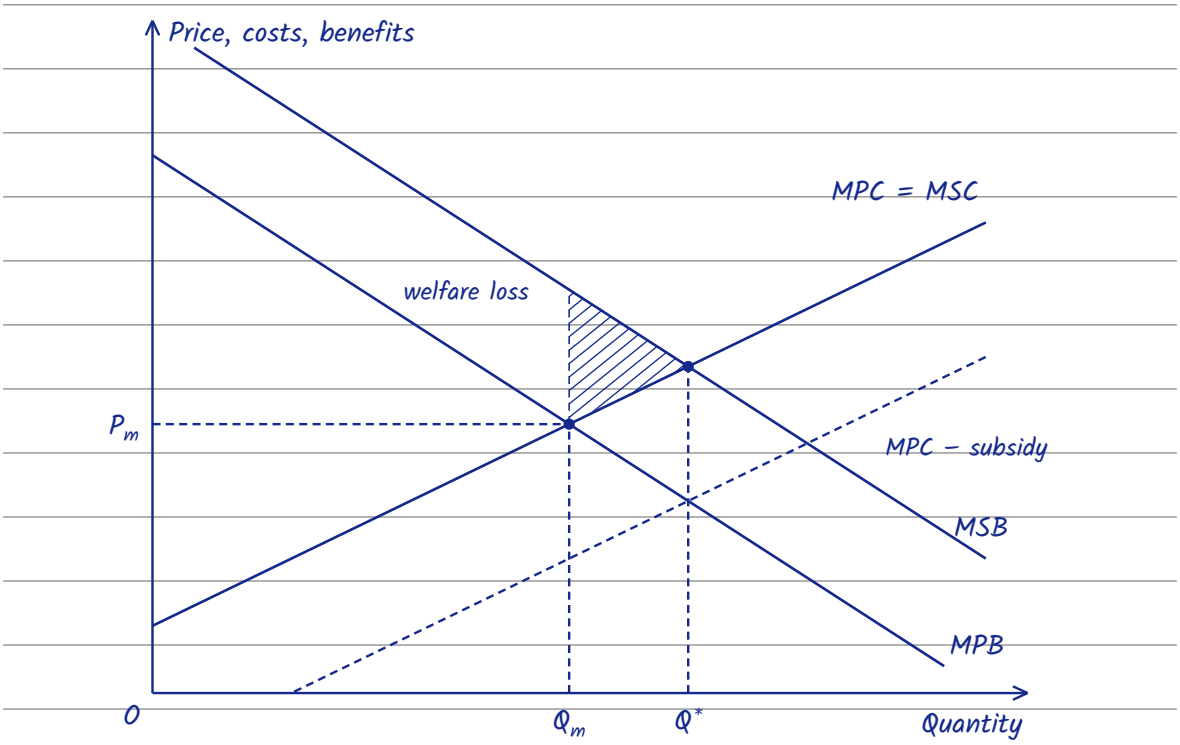
2(b) Market failure occurs when the free market fails to allocate resources efficiently, so the output of a good is not at the socially optimal level where MSB = MSC. Causes include externalities, merit and demerit goods, public goods, information failure and monopoly power.

A negative externality, such as pollution from a factory, means MSC is above MPC, so the market over-produces at Qm instead of Q^*. The government can impose an indirect tax equal to the marginal external cost, which shifts MPC up to MSC, so the market produces Q^* and the welfare loss is removed. The tax 'makes the polluter pay' and raises revenue.



Merit goods such as education create positive externalities, so MSB is above MPB and the market under-produces at Qm. A subsidy reduces the cost of production, shifting the supply curve down, so consumption rises to Q^*.

Question number



Public goods such as street lighting are non-excludable and non-rival, so the free rider problem means the market does not provide them at all. Here the government can provide them directly, financed by taxation. Legislation and regulation can ban or limit harmful activities and correct information failure, for example by requiring health warnings, and competition policy can control monopoly power.

However, market failure cannot always be overcome. The government rarely has perfect information: it is very hard to put a money value on external costs such as damage to health or the environment, so a tax or subsidy may be set at the wrong level and create a new misallocation. This is government failure. Taxes on demerit goods with inelastic demand, such as cigarettes, reduce consumption only a little and are regressive. Subsidies have an opportunity cost and may be hard to remove. Regulations need to be enforced and may be avoided, and intervention can have unintended consequences, such as black markets. Direct provision may be inefficient because there is no profit motive to keep costs down. Some externalities, such as climate change, cross national borders, so no single

Question number

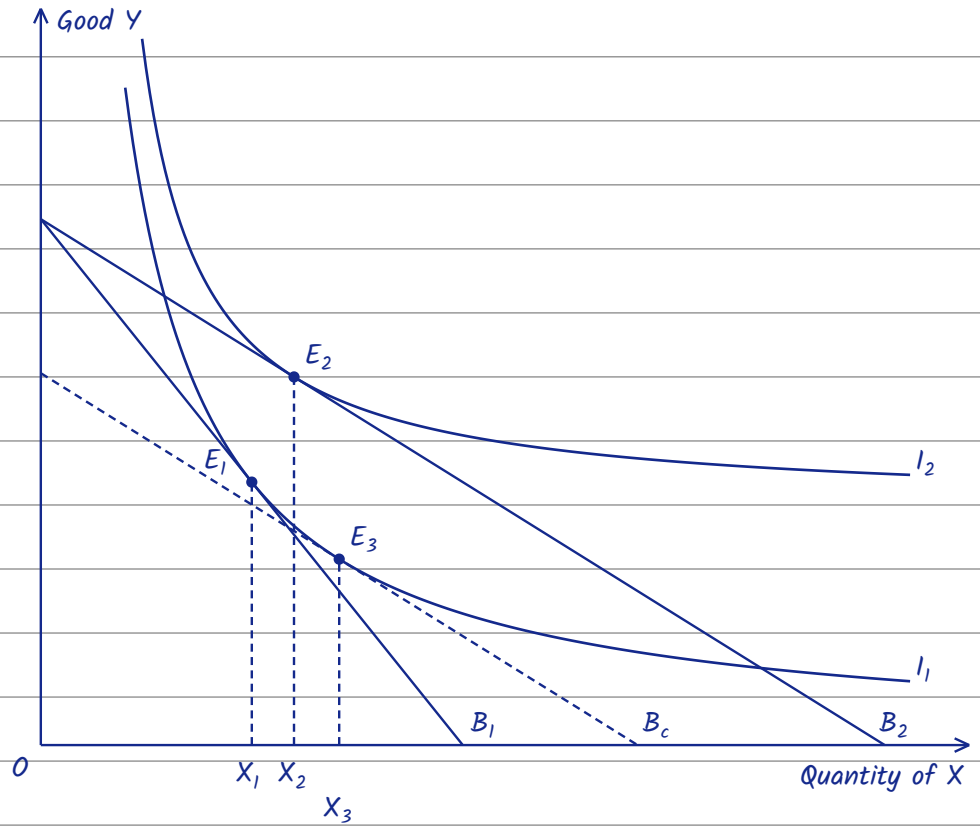
government can correct them.

In conclusion, I disagree that market failure can always be overcome. Government intervention can often reduce market failure and increase efficiency, especially for public goods and where external costs can be measured. But because of imperfect information, administrative costs and the risk of government failure, it will only increase efficiency when the benefits of intervention outweigh its costs, so market failure can be reduced but not always removed.

3(a) An indifference curve shows all the combinations of two goods that give a consumer the same level of satisfaction. It slopes downwards and is convex to the origin because of a diminishing marginal rate of substitution (MRS), and higher curves represent higher satisfaction. The budget line shows all the combinations of the two goods the consumer can afford with a given income and prices; its slope is Px ÷ Py. The consumer maximises satisfaction where the budget line is tangent to the highest indifference curve, where MRS = Px ÷ Py.

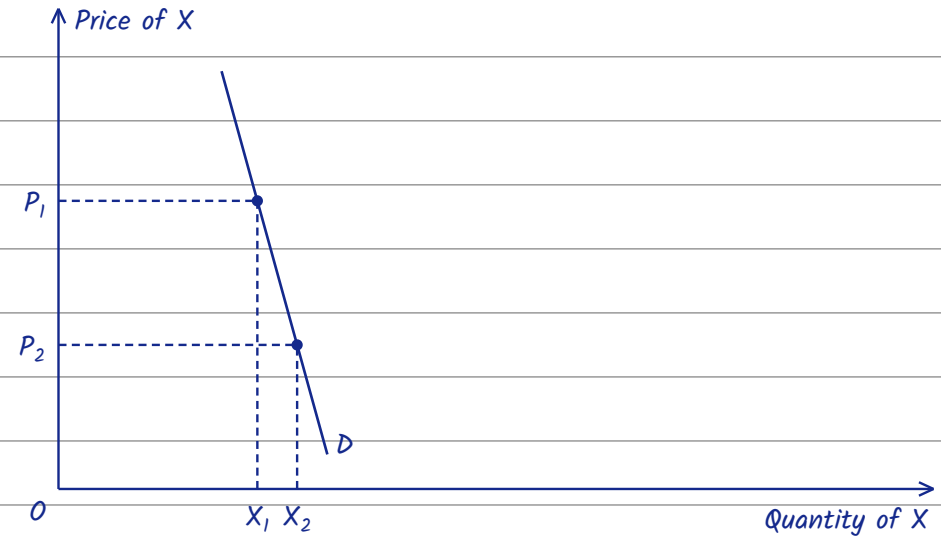
An inferior good is one where demand falls as income rises, such as cheap own-brand food. To derive the demand curve, consider a fall in the price of X. The budget line pivots outwards from B1 to B2 about the Y intercept, and the total effect can be split into a substitution effect and an income effect.

Question number



The consumer starts at E1 on I1, buying X1 at the original price P1. The compensated budget line Bc is drawn parallel to B2, at the new relative prices, but tangent to the original curve I1 at E3. The movement from E1 to E3 is the substitution effect: X is now relatively cheaper, so the consumer buys more, from X1 to X3. The movement from E3 to E2, the parallel shift from Bc to B2, is the income effect: the price fall raises real income, and because X is inferior, the consumer buys less of it, from X3 to X2. The income effect is negative but smaller than the substitution effect, so overall quantity demanded rises from X1 to X2.

Question number



Plotting the original price P1 against X1 and the lower price P2 against X2, and repeating this for other prices, gives the individual's demand curve for the inferior good. It slopes downwards, but because the income effect partly offsets the substitution effect the quantity demanded responds less to price changes than for a normal good, so the curve is steeper (less price elastic). If the negative income effect were larger than the substitution effect, the good would be a Giffen good with an upward-sloping demand curve.

Alternatively, the analysis can show a change in income: a rise in income shifts the budget line outwards parallel, and for an inferior good the new tangency point is at a lower quantity of X, so the whole demand curve shifts to the left.

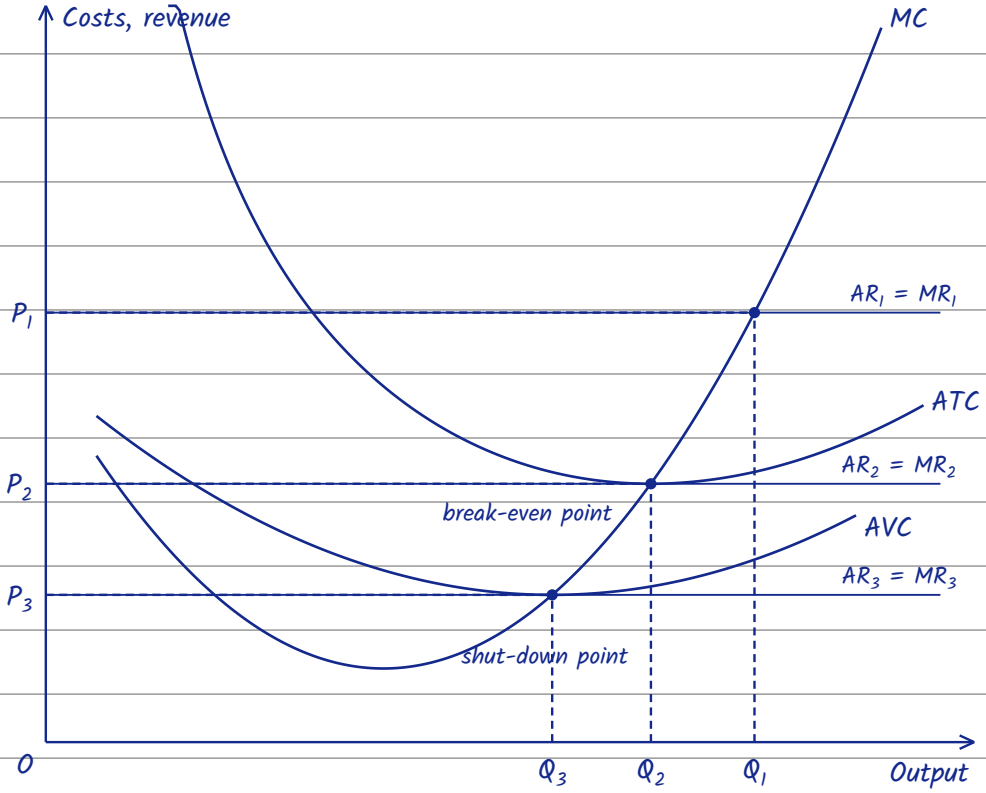
3(b)

In the short run at least one factor of production is fixed, so the firm has fixed costs that must be paid whatever the output. Marginal cost (MC) is the change in total cost from producing one more unit. Average variable cost (AVC) is total variable cost ÷ output. Both are U-shaped because of increasing and then diminishing marginal returns, and MC cuts AVC and ATC at their lowest points.

Marginal cost determines how much a firm produces. A profit-maximising firm produces where MC = MR, with MC rising. Below this output MR > MC, so each extra unit adds to profit; above it MC > MR, so each extra unit reduces profit. If

Question number

MC rises, for example because of higher raw material prices, the profit-maximising output falls.



Average variable cost determines whether the firm produces at all in the short run. At price P1 the firm produces Q1 and makes supernormal profit, since AR is above ATC. At P2 it just breaks even. At prices between P2 and P3 it makes a loss, but because price is above AVC it covers its variable costs and makes a contribution to its fixed costs, so it loses less by staying open than by shutting down. If the price falls below P3, the minimum point of AVC, it cannot even cover variable costs, so it should shut down immediately: this is the shut-down point. So in perfect competition the firm's short-run supply curve is the part of its MC curve above AVC.

So MC and AVC answer different questions: MC decides the level of output and AVC decides whether to produce. In this sense they are equally important, and both are used together, since the profit-maximising output is chosen from MC and then checked against AVC. Average total cost is relevant only to the long-run

Question number

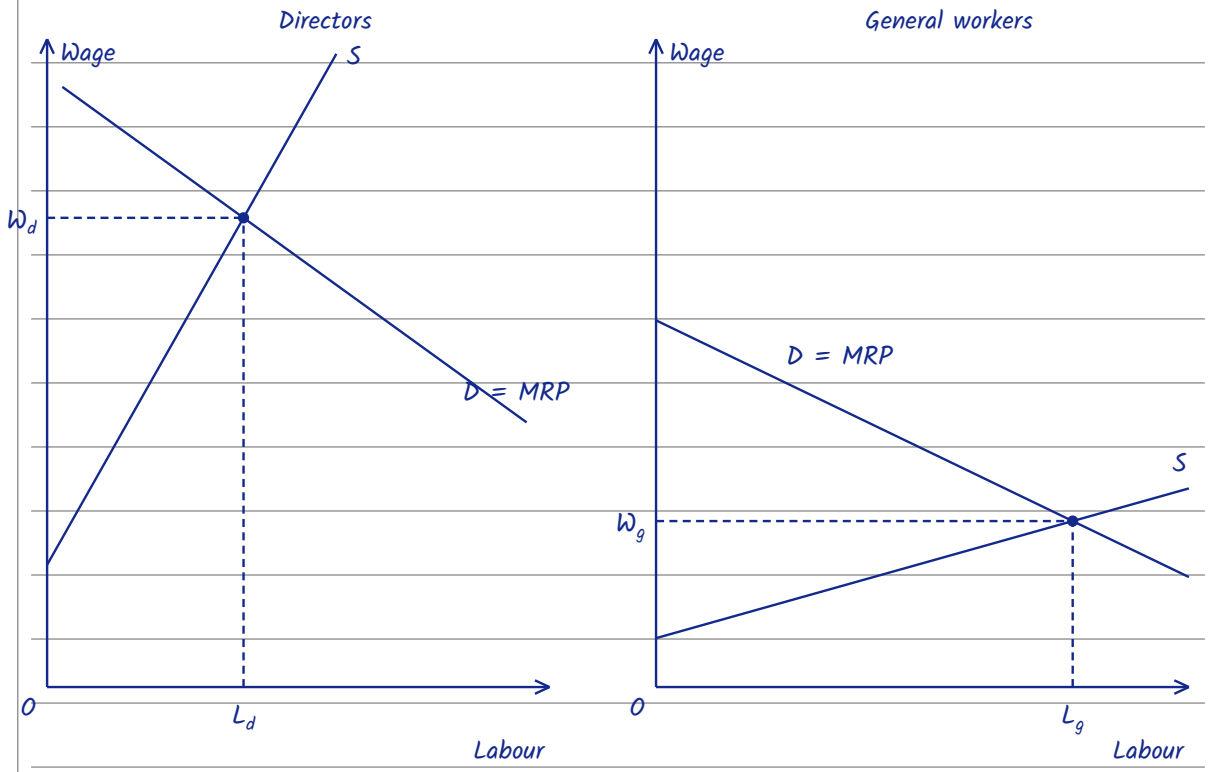
decision about whether to stay in the industry.

In practice, however, firms often do not know their MC precisely and may set prices using average cost plus a mark-up. Firms with other objectives, such as sales maximisation or survival, may also produce where price is below AVC for a time to keep customers or workers. In conclusion, MC is the more important cost for the day-to-day output decision, but AVC is crucial for the decision to produce, so both are needed for short-run production decisions.

4 The theory of wage determination states that in a competitive labour market the wage is set by the demand for and supply of labour. Demand for labour is derived from the demand for the product. According to marginal productivity theory, a profit-maximising firm employs workers up to the point where the wage equals the marginal revenue product of labour (MRP = marginal physical product × marginal revenue), so the MRP curve is the firm's demand curve for labour. The supply of labour to an occupation depends on the wage, non-monetary benefits, the skills and training needed and the working population. For an individual, a higher wage has a substitution effect (leisure becomes more costly, so they work more) and an income effect (they can afford more leisure); at high wages the income effect may dominate, giving a backward-bending supply curve.

(i) Director and general worker in the same company. The theory explains this differential well.

Question number



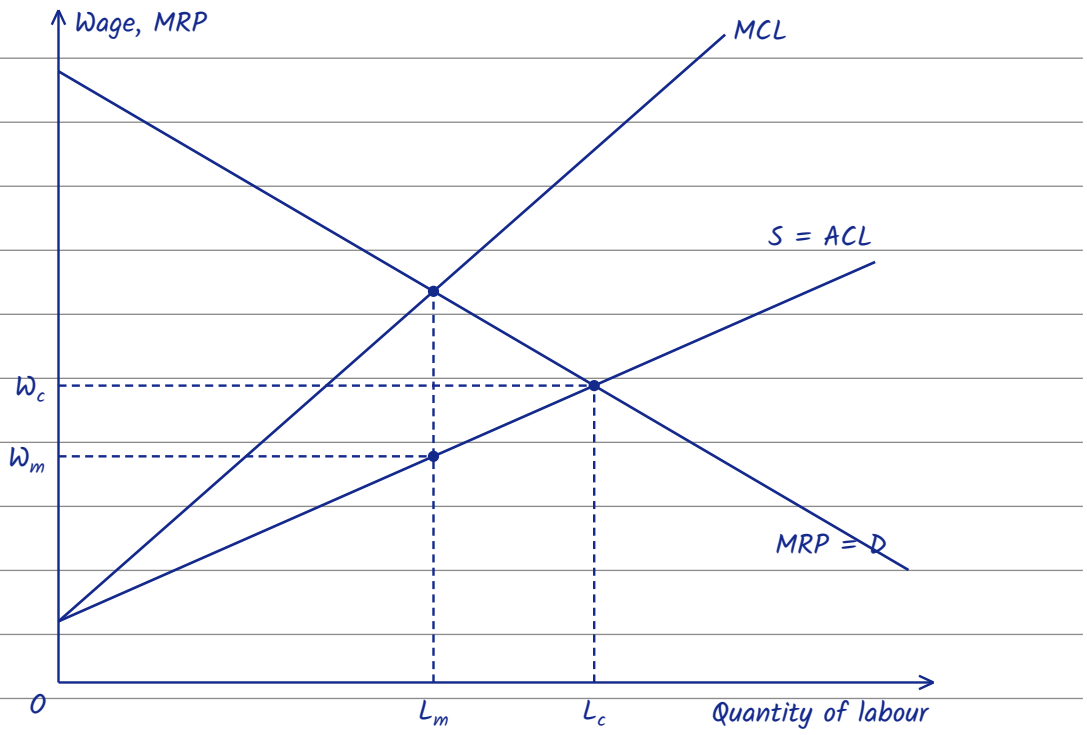
A director's decisions affect the whole company, so their MRP is very high, while a general worker's MRP is much lower. The supply of people with the ability, education and experience to be a director is small and inelastic, because these take many years to develop, whereas the supply of general workers is large and elastic, as little training is needed. So, as the diagram shows, the director's wage Wd is far higher than the general worker's wage Wg. Because supply is so inelastic, most of the director's pay is economic rent, the surplus over transfer earnings. The demand for general workers is also more elastic, since they can be replaced by machines, which keeps their wages low.

However, the theory does not fully account for the size of the gap. The MRP of a director is very hard to measure, and directors' pay is often set by remuneration committees made up of other directors, so it may be higher than MRP (a principal-agent problem). Bonuses and share options may reward short-term share prices. On the other side, general workers' wages may be raised by trade unions or a national minimum wage, or lowered by monopsony power, so the differential reflects institutions as well as market forces.

Question number

(ii) Two workers doing the same job in different companies. In a perfectly competitive labour market with perfect information and perfect mobility, the theory predicts that this differential should disappear, because workers would leave the low-wage firm for the high-wage firm until wages were equal. So differentials for the same job must be explained by imperfections in the labour market.

First, MRP can differ between companies even for the same job: a worker using more modern capital has a higher marginal physical product, and a firm that can sell its output at a higher price has a higher MR, so it can pay more. Second, there are compensating differentials for non-monetary differences: a firm offering poor conditions, unsocial hours or a city location with a high cost of living must pay more, while a firm offering job security or a pension can pay less. Third, labour is immobile: geographical immobility (housing costs, family ties) and imperfect information about vacancies mean workers do not move to the better-paid job. Fourth, market power differs. A firm that is a monopsony employer pays a lower wage than a competitive firm.



The monopsonist employs where MCL = MRP, at Lm, and pays only Wm, below

Question number

the competitive wage Wc. A company whose workers are in a strong trade union may pay above the market wage. Firms may also pay efficiency wages above the market rate to motivate workers and reduce labour turnover, and discrimination can cause different pay for the same work.

Elasticity matters in both cases: where labour supply is inelastic, a rise in demand causes a large rise in wages, and where the demand for labour is inelastic (labour is a small share of total cost or hard to replace), unions can raise wages with little loss of jobs.

In conclusion, the theory of wage determination, based on MRP and labour supply, accounts well for the differential between a director and a general worker, although the size of directors' pay also reflects their power to influence it. The basic competitive theory cannot explain why the same job is paid differently in different companies, since it predicts equal wages; this needs the theory to be extended to include monopsony, trade unions, immobility and imperfect information. So the theory accounts for differentials only when its assumptions of perfect competition are relaxed.

5

Gross National Income (GNI) is the total income earned by a country's residents in a year: GDP plus net income from abroad. To measure living standards it is converted to real GNI per head (adjusted for inflation and divided by the population), and to compare countries it is converted at purchasing power parity (PPP) exchange rates, which allow for differences in the cost of living. A higher real GNI per head means more goods and services are available per person, so material living standards are usually higher.

However, GNI has important limitations. It is an average, so it says nothing about the distribution of income: a rise may go mainly to the rich. It excludes non-market activity such as housework and subsistence farming, and the informal (hidden) economy, which are large in developing countries, so their living standards

Question number

are understated. It ignores leisure time, working conditions, the quality of goods and negative externalities such as pollution and the depletion of resources, and it counts 'defensive' spending, such as on crime prevention, as an increase in welfare. In developing countries, data are often inaccurate because collection is costly and many activities are not recorded.

The Human Development Index (HDI), published by the United Nations Development Programme, combines three dimensions of human development: a long and healthy life (life expectancy at birth), knowledge (mean years of schooling of adults and expected years of schooling of children) and a decent standard of living (GNI per head at PPP). Each is turned into an index, and they are combined into a single value between 0 and 1.

The HDI is a better measure in several ways. It recognises that living standards depend on health and education as well as income, so it shows that two countries with similar GNI per head, for example where one has better public healthcare, can have very different living standards. Because it uses the logarithm of income, extra income adds less at high incomes, reflecting diminishing marginal utility. Its data are relatively easy to collect and compare. However, it still includes GNI and so shares many of its weaknesses; it ignores inequality (although the inequality-adjusted HDI tries to correct this), the environment, political freedom and the quality of education and healthcare, and the equal weighting of the three dimensions is a value judgement. Data on schooling may also be unreliable in developing countries.

The Measure of Economic Welfare (MEW), proposed by Nordhaus and Tobin, starts from national income but adjusts it to measure welfare. It adds the value of leisure, of unpaid household work and of the services of consumer durables, and subtracts 'regrettables' such as defence, commuting costs and police spending, as well as the cost of environmental damage such as pollution and congestion. This corrects many of the weaknesses of GNI, so in principle it is a better measure of

Question number

welfare, especially of the costs of growth.

However, the MEW is very difficult to calculate. There is no agreed money value for leisure, housework or pollution, so the estimates are subjective, and deciding which spending is 'regrettable' is a value judgement. It is not published regularly by most countries, so it cannot easily be used to compare countries or to track change over time, especially in developing countries where basic data are lacking.

The extent to which they are better depends on the purpose. For comparing countries at different levels of development, HDI is more useful than GNI because it captures health and education, which matter most to living standards in poorer countries. For understanding whether growth in a rich country is raising welfare, the MEW is more relevant because it considers leisure and the environment. For short-term measurement and policy, GNI is still the most practical, since it is measured regularly and objectively.

In conclusion, HDI and MEW provide broader and therefore usually better measures of living standards than GNI, because they include non-monetary aspects of welfare. But neither is perfect: HDI still relies on GNI and omits inequality and the environment, and MEW is subjective and rarely measured. They are best used alongside real GNI per head at PPP, not as complete replacements.

6(a)

Unemployment is when people who are willing and able to work, and are actively seeking work, cannot find a job. There are several causes.

Frictional unemployment is short-term unemployment while workers move between jobs; it is caused by imperfect information about vacancies and the time taken to search. Seasonal unemployment occurs when demand for labour falls at certain times of the year, as in agriculture, tourism and construction. Structural unemployment results from long-term changes in the structure of the economy, such as the decline of an industry because of competition from abroad or new technology, when the skills or location of workers do not match the new jobs

Question number

(occupational and geographical immobility). Cyclical or demand-deficient unemployment is caused by a lack of aggregate demand in a recession: firms cut output, so the derived demand for labour falls. Classical (real wage) unemployment occurs when wages are held above the market-clearing level, for example by a minimum wage or trade unions, so the supply of labour exceeds demand.

In a developing country, structural unemployment is likely to be the most significant. As the economy develops, labour moves from agriculture to manufacturing and services, and rapid rural-urban migration means many more people arrive in cities than there are formal jobs. Many workers lack the education and skills that modern industries need, and new industries are often capital intensive. Developing countries also have high population growth, so the labour force grows faster than the number of jobs.

Seasonal unemployment is also important, because a large share of employment is in agriculture. Much unemployment in developing countries is hidden as underemployment or disguised unemployment: workers in family farms or the informal sector work fewer hours than they want, or their marginal product is close to zero. Cyclical unemployment can be severe when the economy depends on one or two commodity exports and world prices fall. Frictional unemployment is less significant, since few people can afford to search for long without benefits.

In conclusion, structural unemployment, together with underemployment, is most likely to occur in a developing country, because it is caused by the process of development itself, while cyclical unemployment can add to it in commodity price downturns.

6(b)

A government's main macroeconomic objectives are low unemployment, economic growth, low and stable inflation, a sustainable balance of payments and a fair distribution of income. Developing countries also aim to reduce poverty and achieve sustainable development.

Question number

There is a strong case for making unemployment the main objective.

Unemployment wastes scarce resources, so the economy produces inside its PPF; it causes poverty, especially where there are no unemployment benefits; and with a young and fast-growing population, high youth unemployment can lead to social unrest and emigration of skilled workers (brain drain). Creating jobs is also one of the most direct ways to reduce poverty.

However, economic growth may be a more important objective, because sustained growth is the main way of creating jobs in the long run. Growth raises real GNI per head, provides tax revenue for education and healthcare, and through rising demand creates employment. Targeting unemployment directly, for example by employing people in the public sector, may be expensive and inefficient. On the other hand, growth that is capital intensive may create few jobs (jobless growth), so growth does not always reduce unemployment.

Low inflation is also important. High inflation in a developing country hurts the poor most, since they spend most of their income on food and have no inflation-protected savings, and it discourages investment. Reducing unemployment through demand-side policies may create a trade-off, shown by the Phillips curve: lower unemployment comes with higher inflation, which may not be acceptable.

The balance of payments is another key objective. Many developing countries have current account deficits and foreign debt, and expanding demand to reduce unemployment would increase imports and the deficit. Earning foreign currency through exports may be essential to pay for imported capital goods and to avoid a currency crisis. Reducing inequality may also matter, since a very unequal distribution of income means growth does not reduce poverty.

Measurement problems also matter: in developing countries much unemployment is hidden in the informal sector or as underemployment, so official unemployment figures may be low and a poor guide to policy.

In conclusion, reducing unemployment should be a major objective, but it should

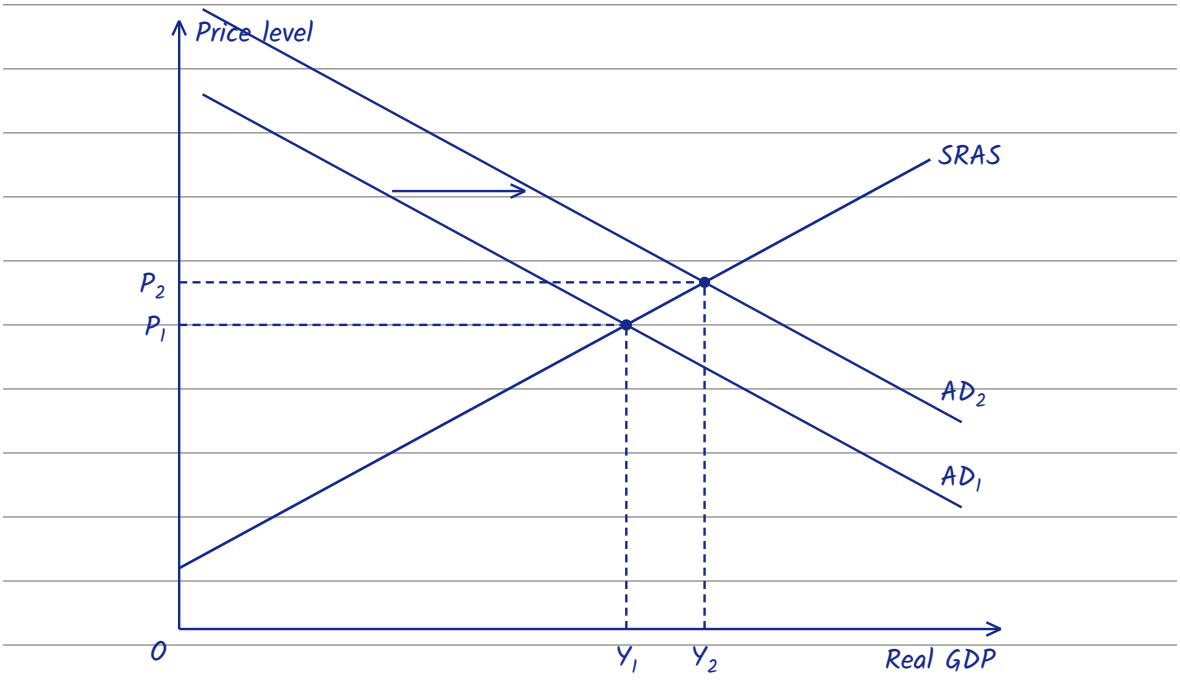
Question number

not necessarily be the main one. Its importance depends on the country: where youth unemployment is very high it may be the top priority, but in many developing countries sustained and inclusive economic growth, combined with low inflation and a sustainable balance of payments, is the best way of reducing unemployment and poverty in the long run.

7(a) Quantitative easing (QE) is a form of monetary policy in which the central bank creates new money electronically and uses it to buy financial assets, mainly government bonds but also corporate bonds, from banks and other financial institutions. It is used when interest rates are already close to zero, so they cannot be cut further.

QE works through several channels. The large purchases increase the demand for bonds, so bond prices rise and their yields (long-term interest rates) fall. This reduces the cost of borrowing for firms and households, encouraging investment and consumption. The sellers of bonds receive money, which they may use to buy other assets such as shares and property, so asset prices rise and there is a wealth effect that boosts spending. Banks have more reserves, so they may lend more. Lower interest rates may also cause the exchange rate to depreciate, making exports more competitive.

Question number



So QE should shift AD to the right from AD1 to AD2, raising real GDP from Y1 to Y2 and reducing unemployment, and preventing deflation by raising the price level from P1 to P2.

It can be effective in a recession. After the financial crisis of 2008 and during the pandemic in 2020, central banks such as the Bank of England and the US Federal Reserve used QE on a large scale, and long-term interest rates fell to very low levels, which helped to prevent a deeper recession and deflation.

However, its effectiveness is limited. In a recession, confidence is low, so firms may not want to borrow even at low interest rates, and banks may hold the extra reserves rather than lend them; the economy may be in a liquidity trap. The extra money may go into buying assets rather than spending on goods and services, so it raises asset prices and increases wealth inequality, since the rich own most shares and property. There are risks too: QE may lead to inflation later, as in 2021 to 2022, and it is difficult to reverse. Its effects are also uncertain and take time.

In conclusion, QE can help to support demand in a recession when interest rates cannot be cut further, mainly by lowering long-term interest rates, but on its own it is unlikely to be fully effective, so it works best alongside expansionary fiscal

Question number

policy.

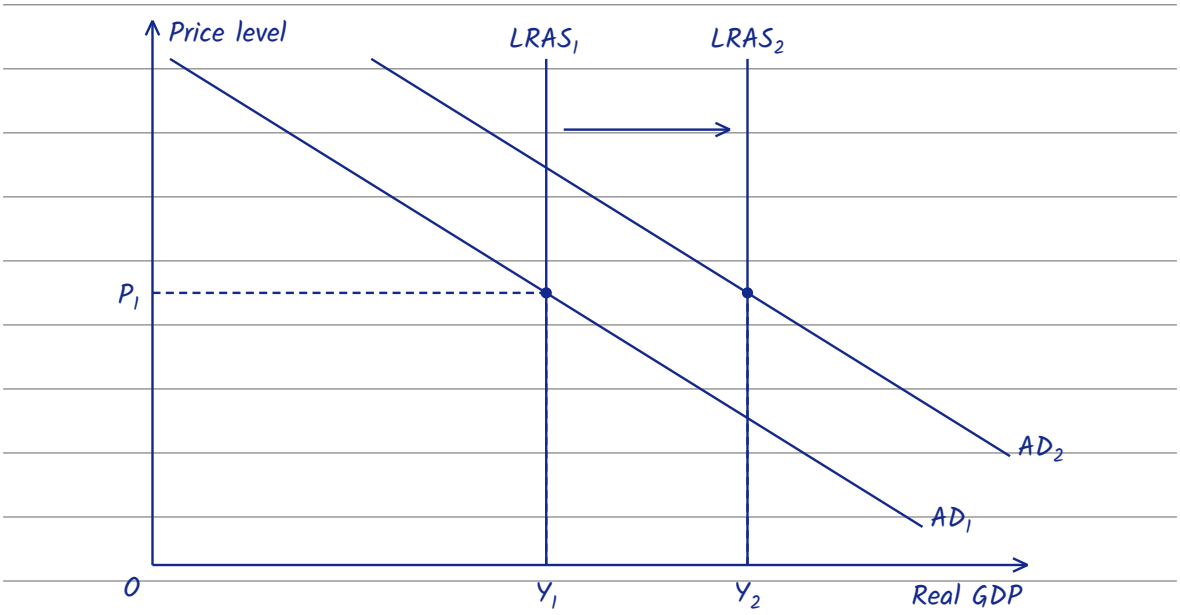
7(b) A government budget deficit is when government spending is greater than government revenue in a year. Economic growth is an increase in real GDP. Short-run (actual) growth is an increase in output using spare capacity, caused by a rise in AD; long-run (potential) growth is an increase in productive capacity, shown by an outward shift of LRAS.

Expansionary fiscal policy, increasing government spending or cutting taxes, directly increases the budget deficit, at least at first. Higher spending raises AD and, through the multiplier, national income rises by more than the initial injection. However, as incomes rise, tax revenue increases and spending on unemployment benefits falls, so the deficit increases by less than the original change. If the multiplier is large, growth can be partly self-financing.

Monetary policy can achieve growth without a budget deficit. Lower interest rates or quantitative easing reduce the cost of borrowing, so investment (through the marginal efficiency of capital) and consumption rise and AD shifts right. Lower interest rates also reduce the cost of servicing the national debt, and the resulting growth increases tax revenue, so the deficit may actually fall.

Supply-side policies aim at long-run growth.

Question number



Interventionist supply-side policies, such as spending on education, training and infrastructure, increase the quantity and quality of resources and shift LRAS to the right, from LRAS1 to LRAS2, raising potential output from Y1 to Y2 without inflation. These policies do increase spending and may increase the deficit in the short run. But market-based supply-side policies, such as deregulation, reforming labour markets and privatisation, cost little and privatisation even raises revenue. Tax cuts intended to improve incentives increase the deficit at first, but may raise revenue later if they increase work and enterprise.

Growth also tends to reduce a deficit over time. A large part of most deficits is cyclical: in a recession tax revenues fall and benefit spending rises through automatic stabilisers, so when the economy grows the cyclical deficit falls automatically. Faster growth also reduces the national debt as a percentage of GDP. Long-run growth also depends on factors such as population growth, private saving and investment and technology, which do not need government spending.

In conclusion, I disagree that policies for growth inevitably cause a budget deficit. Expansionary fiscal policy and some interventionist supply-side policies are likely to increase the deficit in the short run, but monetary policy and market-based supply-side policies need not, and in the long run growth itself increases tax

Question number

revenue and reduces the deficit. Whether a deficit results depends on the policies chosen, the size of the multiplier and the time period considered.

Cambridge Assessment International Education

Cambridge International AS & A Level

ECONOMICS **9708/43**

Paper 4 Data Response and Essays **May/June 2021**

2 hours 15 minutes

5207815665

You must answer on the enclosed answer booklet.

You will need: Answer booklet (enclosed)

INSTRUCTIONS

- Answer **three** questions in total:
  Section A: answer Question 1.
  Section B: answer **two** questions.
- Follow the instructions on the front cover of the answer booklet. If you need additional answer paper, ask the invigilator for a continuation booklet.
- You may use a calculator.
- You may answer with reference to any economy you have studied where relevant to the question.

INFORMATION

- The total mark for this paper is 70.
- The number of marks for each question or part question is shown in brackets [ ].

This document has **4** pages. Any blank pages are indicated.

DC (NF) 212777
© UCLES 2021

[Turn over

stemup.co.uk

PapaCambridge

Section A

Answer this question.

1 Public services – better or worse?

A government usually has macroeconomic policy aims that it hopes will enable some success in raising the quality of life, or well-being, of the population. In doing this the government also hopes to achieve efficiency in the use of resources.

Efficiency is measured by relating inputs to outputs. Inputs are relatively easy to count: they are financial costs of public services. Output can also be counted, but it is not necessarily a good measure of the outcome. The outcome is much harder to calculate. It is broader and more subjective – how do we assess whether a public service is 'better' or 'worse'?

This is where a knowledge of well-being makes an enormous difference. It could provide a much clearer view of the trade-offs that have to be made in allocating taxpayers' money to public services. Take the example of healthcare. To maximise the impact of expenditure on well-being, the budget may need to be adjusted to give more to mental health services and less to building general hospitals. For older people it could mean giving priority to programmes that would keep them out of hospital.

A focus on well-being should lead to better outcomes. This is where policymakers need a better understanding of behavioural economics. Governments have established Behavioural Insights teams, or Nudge Units. They have had some success. A small change in the wording of a letter to people who owed tax demonstrated how more behaviourally sensitive language sped up payments. The unit also found that jobseekers were nearly twice as likely to turn up for a job fair if the text message from the job centre used their names, and nearly three times as likely if the person sending the text message added 'good luck'.

Do tax reliefs persuade people to save? No. So enrol them instead automatically in a pensions programme as a default position, with the possibility of opting out. Allowing people to learn from mistakes is good: it reduces dependency on the public sector and helps people make better decisions for themselves. But some errors, such as failing to save anything until you are too old to earn, cannot be reversed. Then an early 'nudge' is justified. It has proved successful in spreading the habit of saving for retirement into groups not persuaded by tax reliefs alone.

In the long term this, and similar behavioural changes, may well have more influence on well-being than can be represented by concentrating on a monetary calculation of GDP. Other economic indicators could be used to assess this change in well-being.

Source: RSA Issue 1, 2017

(a) The article refers to macroeconomic policy aims. Identify and explain **two** such macroeconomic policy aims. [4]

(b) Is there evidence in the article that a knowledge of behavioural economics can help public policy? [4]

(c) The article says that 'efficiency is measured by relating inputs to outputs'. Is this how economic theory states that efficiency is determined? [5]

(d) The article deals with an improvement in well-being. Discuss whether there are any economic indicators that could be used to assess whether well-being has become better or worse. [7]

© UCLES 2021

stemup.co.uk

PapaCambridge

Section B

Answer **two** questions.

2 Assess the suggestion that a free market economy is neither possible nor desirable. [25]

3 **(a)** Explain what is meant by the concept of the 'equilibrium position of a consumer' and how the concept might be used to construct a demand curve for a good. [12]

(b) Distinguish between the income and substitution effects of a change in a good's price and analyse why the effect of a change in price is not always the same for different goods. [13]

4 **(a)** What costs ought a profit-maximising firm take into consideration when making decisions about price and output? [12]

(b) 'Price discrimination is always possible but never desirable.'

Do you agree with this opinion? [13]

5 'Wage differentials can be explained by economic theory. They are a sign of the power of a firm to exploit its workforce and are unjust.'

Do you agree with this view? [25]

6 **(a)** Distinguish between a country's national debt and its public sector budget deficit and consider which is the more important. [12]

(b) Discuss the effectiveness of alternative macroeconomic policies used to reduce a public sector budget deficit. [13]

7 **(a)** Explain and critically evaluate the quantity theory of money. [12]

(b) Monetary policy relies heavily on the theory of a monetary transmission mechanism.

Explain how a monetary transmission mechanism works and discuss its effectiveness. [13]

© UCLES 2021

stemup.co.uk

PapaCambridge

BLANK PAGE

Permission to reproduce items where third-party owned material protected by copyright is included has been sought and cleared where possible. Every reasonable effort has been made by the publisher (UCLES) to trace copyright holders, but if any items requiring clearance have unwittingly been included, the publisher will be pleased to make amends at the earliest possible opportunity.

To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced online in the Cambridge Assessment International Education Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download at www.cambridgeinternational.org after the live examination series.

Cambridge Assessment International Education is part of the Cambridge Assessment Group. Cambridge Assessment is the brand name of the University of Cambridge Local Examinations Syndicate (UCLES), which itself is a department of the University of Cambridge.

© UCLES 2021

stemup.co.uk

PapaCambridge

Surname	StemUp Model Solution
Answer booklet	Write the question number in the left-hand margin at the start of each answer.

Question number

Answers for paper 9708/43

1(a) Low and stable inflation: the government aims to keep the general price level rising slowly and predictably, for example a target of 2% a year. This protects the purchasing power of incomes and savings, reduces uncertainty for firms planning investment and keeps exports internationally competitive.

Low unemployment (full employment): the government aims to keep the percentage of the labour force who are willing and able to work but have no job as low as possible. This means scarce labour resources are not wasted, output and incomes are higher, and spending on unemployment benefits falls, which raises well-being.

1(b) Yes, there is some evidence. A small change in the wording of a letter to people who owed tax sped up payments, so tax revenue was collected sooner, which helps to reduce the government's budget deficit. Jobseekers were nearly twice as likely to attend a job fair when a text used their name, and nearly three times as likely when it added 'good luck', which could help reduce unemployment.

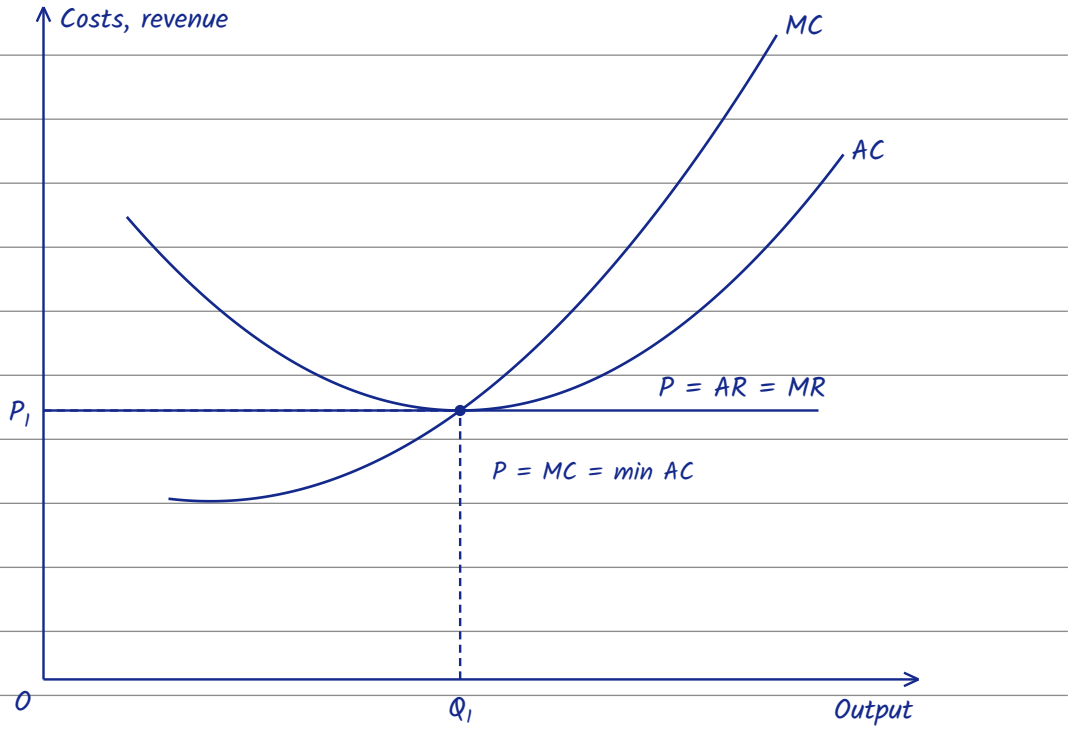
Automatic enrolment in a pension scheme, with the option to opt out, spread the habit of saving to groups not persuaded by tax reliefs. Higher saving reduces consumption, which could help to limit demand-pull inflation, and reduces dependency on the state in old age.

However, the evidence is limited: the article only says the units had 'some success', it gives no figures for the actual increase in tax paid, jobs found or savings, and going to a job fair is not the same as getting a job. So the evidence suggests behavioural economics can help, but its overall impact is not proven.

Question number

1(c)

Only partly. Relating inputs to outputs is how economic theory measures productive (technical) efficiency: a firm or public service is productively efficient when it produces its output at the lowest possible average cost, the minimum point of the average cost curve, so the fewest scarce resources are used for each unit of output. However, economic theory also defines allocative efficiency, which occurs where price equals marginal cost (P = MC). Here resources are allocated to producing the goods and services that consumers value most, so the value of the last unit to society equals the cost of the resources used to make it. This depends on the outcome for consumers, not just on the quantity of output, which is exactly the problem the article raises: output is 'not necessarily a good measure of the outcome'. A hospital could treat more patients cheaply and still be allocatively inefficient if money would raise well-being more in mental health services.



The diagram shows a firm in long-run perfect competition achieving both: P = MC (allocative) and production at minimum AC (productive). Theory also uses Pareto efficiency, where no one can be made better off without making someone else

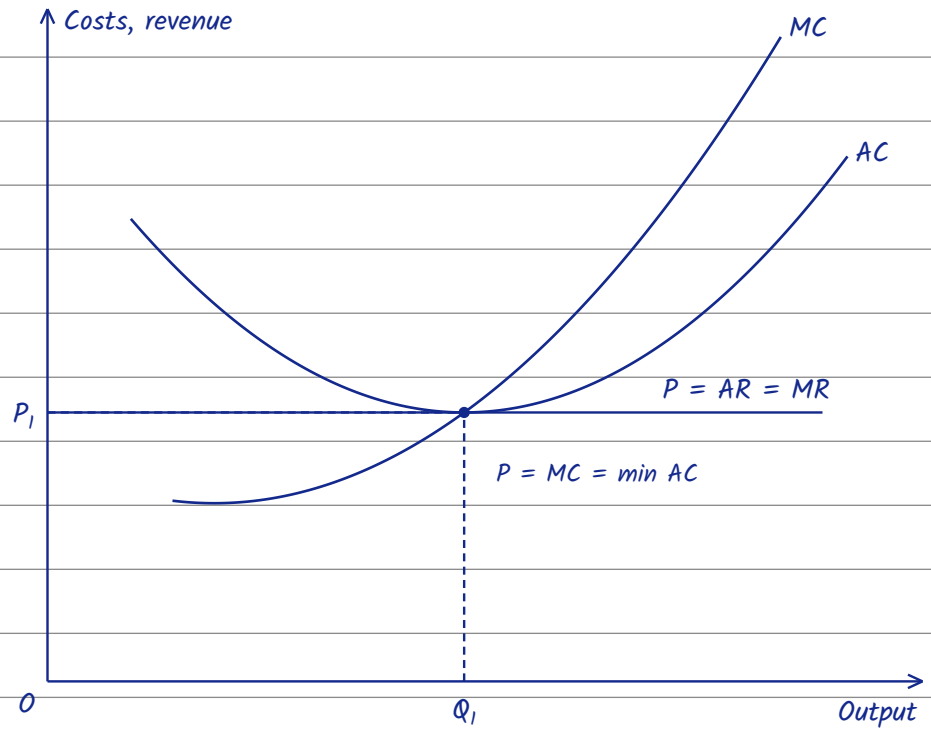
Question number	
	worse off, and dynamic efficiency over time. So the article's measure covers only one part of what economists mean by efficiency.
1(d)	Real GDP per head is the most widely used indicator. A rise means more goods and services per person, so material living standards usually improve. However, it ignores the distribution of income, leisure time, unpaid work, the informal economy and environmental damage, and the article itself warns that a 'monetary calculation of GDP' may not represent changes in well-being. The Human Development Index (HDI), published by the UNDP, is broader. It combines three dimensions: health (life expectancy at birth), education (mean years and expected years of schooling) and income (GNI per head at purchasing power parity), into an index from 0 to 1. A rising HDI shows that people are living longer and are better educated as well as richer, which is closer to well-being. But it uses only three measures, gives them equal weight and ignores inequality, mental health and political freedom. The Measure of Economic Welfare (MEW), developed by Nordhaus and Tobin, adjusts GDP by adding the value of leisure and of unpaid household work and subtracting 'regrettables' such as defence and commuting costs and the cost of environmental damage. This makes it more relevant to well-being, but the values of leisure and pollution are subjective estimates, so it is hard to calculate and rarely published. Other indicators include subjective measures of happiness and life satisfaction from surveys, and the Index of Sustainable Economic Welfare. These measure well-being directly, but answers depend on expectations and culture, so they are hard to compare between countries. In conclusion, there are useful indicators, but no single one fully measures well-being. A combination of real GDP per head with HDI and well-being surveys gives the best assessment of whether it is getting better or worse.

Question number

2

A free market economy is one in which resources are allocated entirely by the price mechanism, through the interaction of demand and supply, with private ownership of resources and no government intervention. Prices act as signals, incentives and rationing devices: if consumers want more of a good, its price rises, profits rise, and firms move resources into producing it.

Supporters argue it is desirable because it leads to an efficient allocation of resources. Productive efficiency is producing at the lowest point on the average cost curve, so no resources are wasted. Allocative efficiency is where P = MC, so resources produce exactly the goods consumers value most. In perfect competition, freedom of entry and exit means that in the long run supernormal profits are competed away and both are achieved.

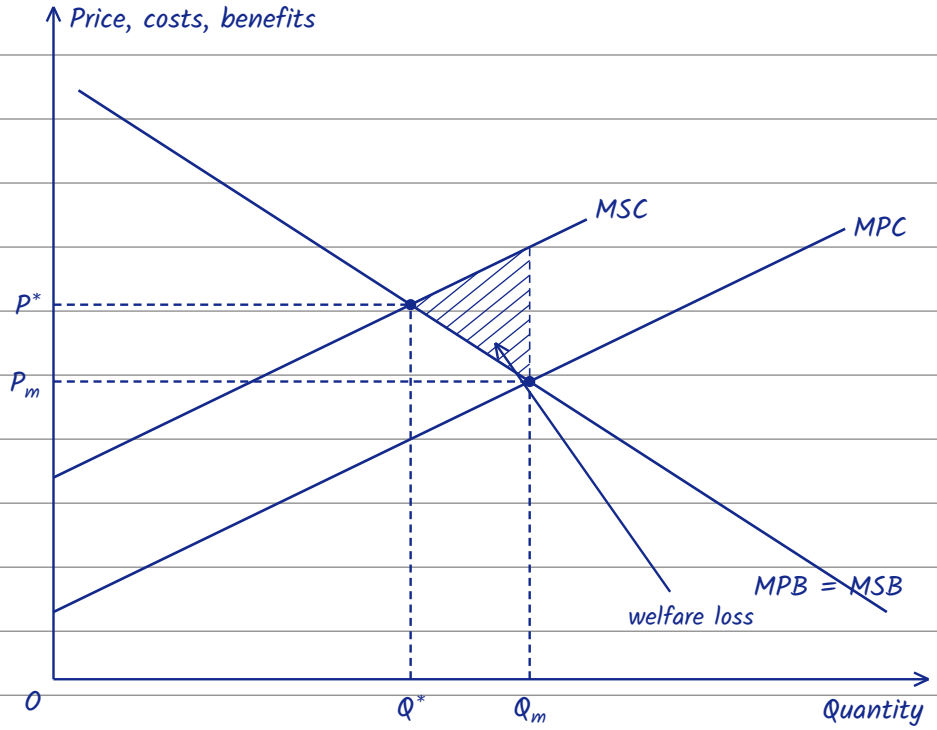


The diagram shows long-run equilibrium at Q1, where P = MC and P = minimum AC. Competition also gives consumer sovereignty and choice, and the profit motive creates incentives to innovate, which brings dynamic efficiency. Decisions are made by millions of individuals who have better information about their own wants than

Question number

a government planner, so a free market avoids the bureaucracy and shortages of command economies.

However, a free market is not desirable because individual actions are not always best for society as a whole. When production creates negative externalities such as pollution, firms consider only their private costs, so the market produces where MPC = MPB at Qm, more than the socially optimal output Q^* where MSC = MSB. The shaded triangle is the welfare loss, so resources are misallocated.



Merit goods such as education and healthcare have positive externalities and consumers lack information about their full benefits, so they are under-consumed, while demerit goods such as tobacco are over-consumed. A free market also tends to produce monopolies: a monopolist restricts output and sets price above MC, so it is allocatively inefficient. Finally, the market distributes income according to the ownership of factors of production, so it can create great inequality and leave people without access to basic needs, which many would see as unfair.

A pure free market is also not possible. Public goods such as defence, street lighting and flood defences are non-excludable and non-rival, so the free rider

Question number

problem means no private firm can charge for them and they would not be provided at all; this is a missing market. Markets themselves depend on government: they need a legal system to define and enforce property rights and contracts, a police force, and a stable currency, all provided by the state. This is why every real economy, even those close to a free market such as Hong Kong or the USA, is in practice a mixed economy.

On the other hand, government intervention can lead to government failure. Governments lack full information about the size of external costs and benefits, so taxes and subsidies may be set at the wrong level; intervention has administrative costs, may be influenced by political pressures and can create unintended consequences, such as black markets after maximum prices. Nationalised industries may lack the incentive to cut costs. So intervention will only improve efficiency if the welfare gain exceeds the cost of the government failure.

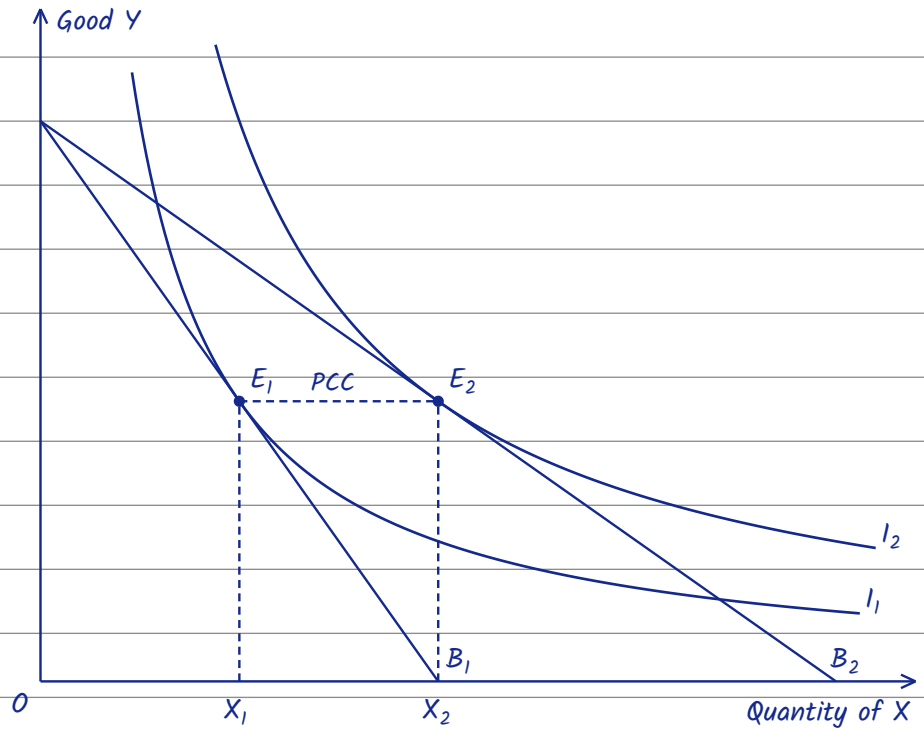
In conclusion, the suggestion is largely correct. A completely free market is not possible, because public goods and the legal framework that markets rely on must be provided by the state, and it is not desirable, because of externalities, merit goods, monopoly power and inequality. However, this does not mean the market mechanism itself is undesirable: it is the most efficient way to allocate most goods, so the best outcome is a mixed economy in which the market allocates most resources and the government intervenes only where there is clear market failure.

3(a) A consumer is in equilibrium when they have allocated their limited income between goods so that total utility (satisfaction) is maximised, so there is no incentive to change their spending. Using marginal utility theory, and the law of diminishing marginal utility, this happens when the marginal utility per $ spent is the same for every good (the equi-marginal principle): $MUx \div Px = MUy \div Py$, and all income is spent.

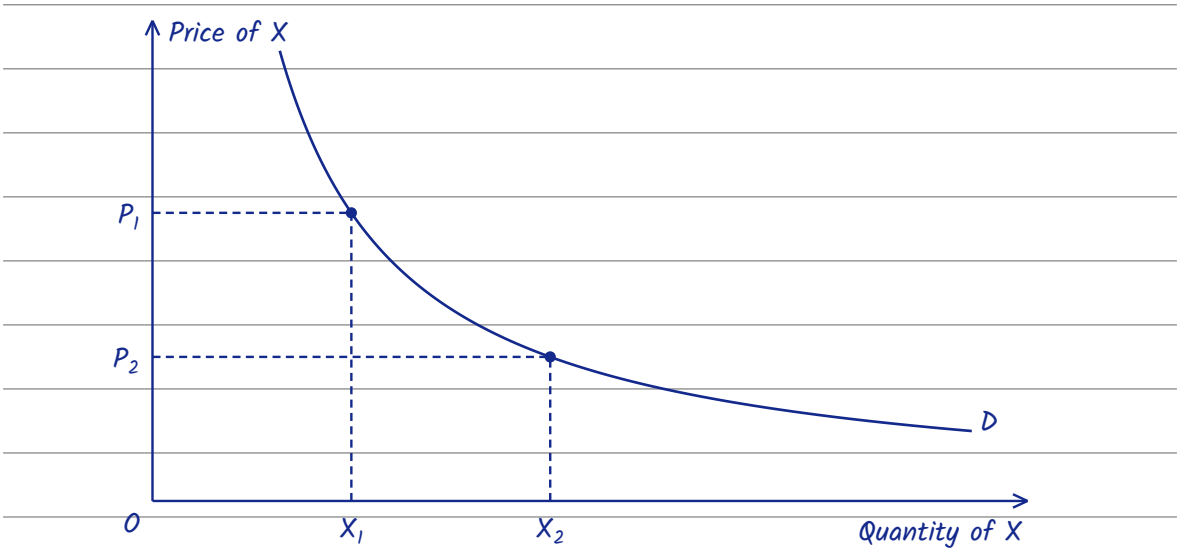
Question number

This can be used to build a demand curve. Suppose the price of X falls. Now MUx ÷ Px > MUy ÷ Py, so the last $ spent on X gives more satisfaction than the last $ spent on Y. The consumer buys more X and less Y. As they buy more X, its marginal utility falls, until MUx ÷ Px = MUy ÷ Py again at a larger quantity of X. So a lower price gives a larger quantity demanded, and plotting each price against the equilibrium quantity gives a downward sloping demand curve.

Using indifference curve analysis: an indifference curve shows combinations of two goods that give equal satisfaction; it is convex to the origin because of a diminishing marginal rate of substitution. The budget line shows the combinations the consumer can afford, with slope equal to the price ratio Px ÷ Py. The consumer is in equilibrium where the budget line is tangent to the highest attainable indifference curve, so MRS = Px ÷ Py.



Question number



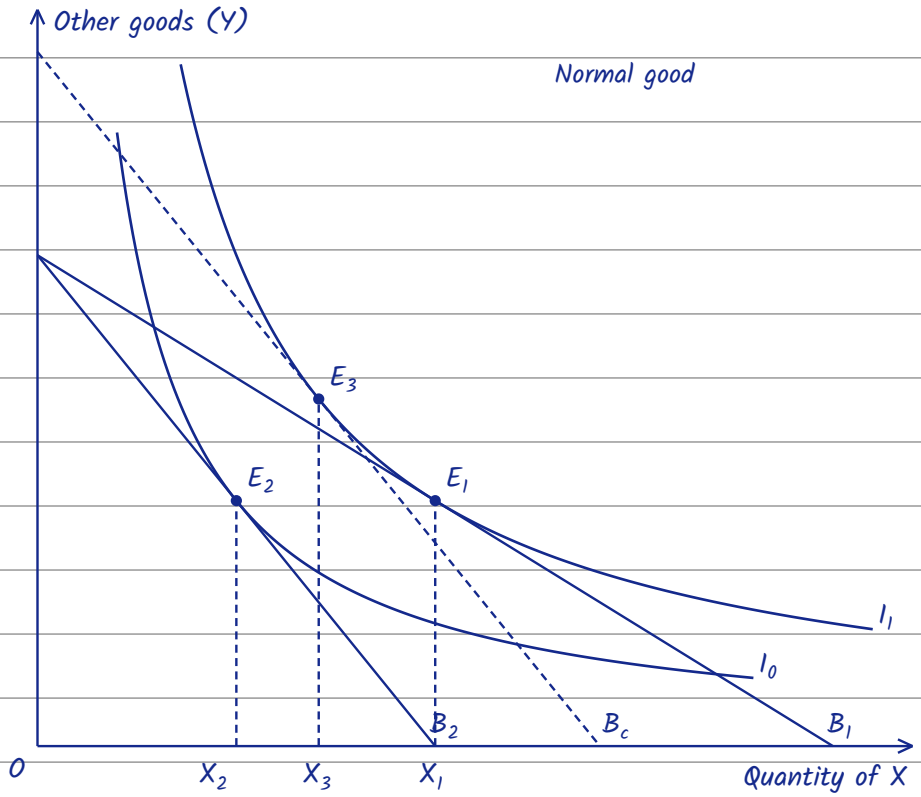
In the first diagram, with price P1 the budget line is B1 and the consumer is in equilibrium at E1 on I1, buying X1. When the price of X falls to P2, the budget line pivots outwards about the Y intercept to B2, because the same income now buys more X. The new equilibrium is E2 on the higher indifference curve I2, buying X2. Joining the equilibrium points gives the price consumption curve (PCC). Each price and the equilibrium quantity is then plotted in the second diagram: P1 with X1 and P2 with X2. Repeating this for every possible price traces out the individual's demand curve D, which slopes downwards for a normal good.

3(b) When the price of a good changes, the change in quantity demanded can be split into two effects. The substitution effect is the change caused only by the change in relative prices, holding real income constant: a good that becomes relatively more expensive is replaced by substitutes, so the substitution effect is always in the opposite direction to the price change. The income effect is the change caused by the change in real income (purchasing power): a price rise makes the consumer worse off in real terms. Its direction depends on the type of good.

In the diagrams, a rise in the price of X pivots the budget line inwards from B1 to B2. The compensated budget line Bc is drawn parallel to B2 (new relative prices) but tangent to the original curve I1. The movement E1 to E3 is the substitution

Question number

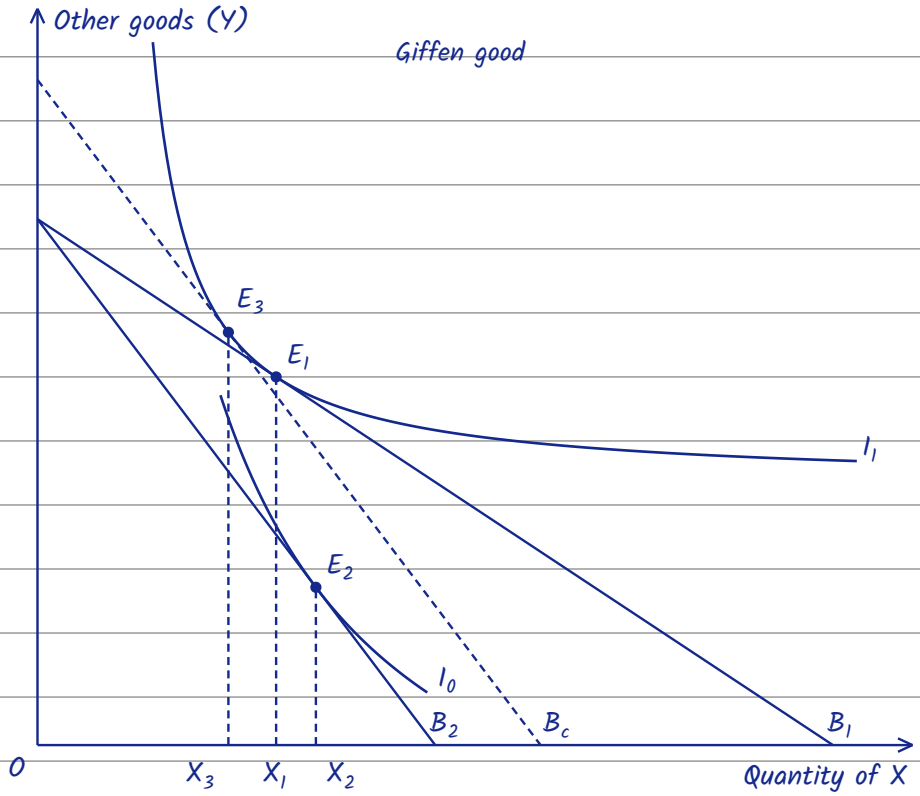
effect and the parallel shift from Bc to B2, E3 to E2, is the income effect.



For a normal good, demand falls when real income falls, so the income effect reinforces the substitution effect: demand falls from X1 to X3 and then further to X2. The demand curve slopes downwards. How much demand falls depends on the price elasticity of demand: if there are close substitutes, the substitution effect is large and demand is elastic; for a necessity with few substitutes, demand changes only a little.

For an inferior good, demand rises when real income falls, so the income effect works against the substitution effect. If, as usual, the substitution effect is larger, demand still falls when price rises, but by less than for a normal good, so demand is less price elastic.

Question number



A Giffen good is an inferior good where the positive income effect is larger than the substitution effect. In the diagram, the substitution effect reduces demand from X1 to X3, but the fall in real income increases demand from X3 to X2, so demand ends up higher than the original X1. A rise in price therefore increases the quantity demanded, and the demand curve slopes upwards. This happens with a staple food that takes a large share of a very poor household's income, because when its price rises the household can no longer afford better foods and buys more of the staple.

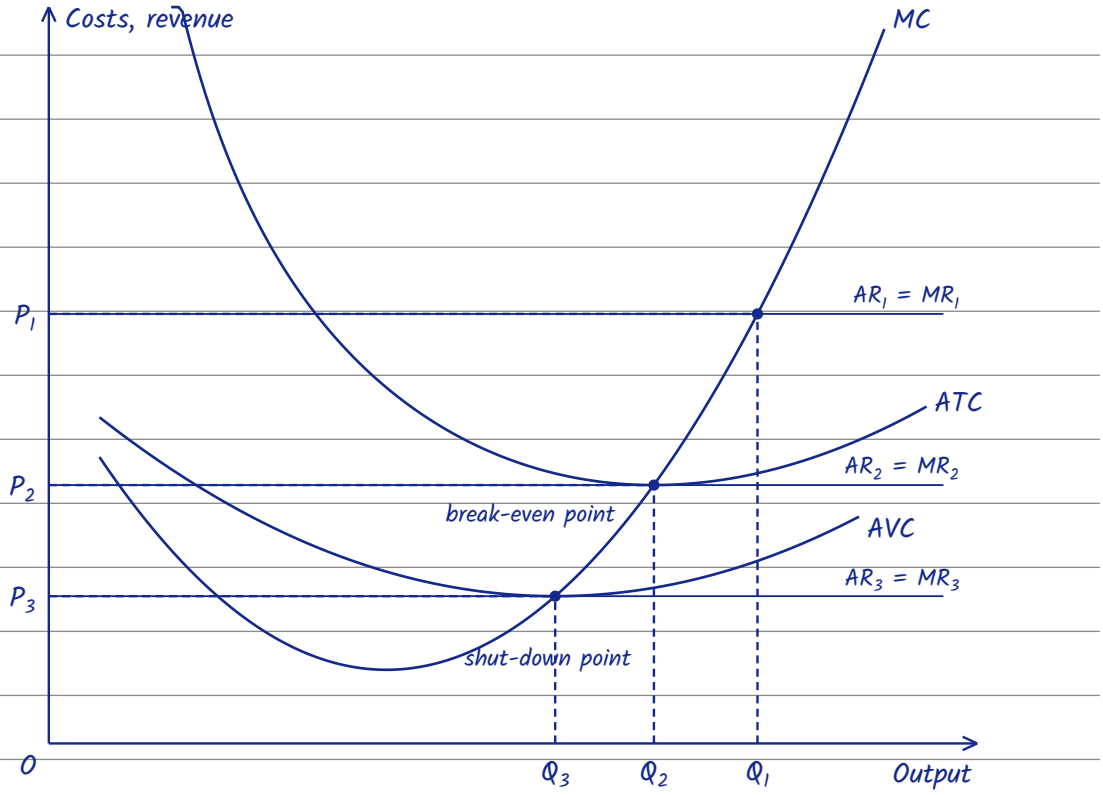
So the effect of a price change is not the same for all goods because the size and direction of the income effect differ between normal, inferior and Giffen goods, and the size of the substitution effect depends on the availability of substitutes. The difference is also greater when the good takes a large share of income, since the income effect is then larger. In practice Giffen goods are very rare, so for almost all goods a price rise reduces demand, differing only in how much.

Question number

4(a)

A firm's costs can be divided into fixed costs, which do not change with output in the short run (rent, interest on loans), and variable costs, which rise with output (raw materials, wages of hourly workers). Total cost = fixed cost + variable cost. Average total cost (ATC) is total cost ÷ output, average variable cost (AVC) is variable cost ÷ output, and marginal cost (MC) is the change in total cost from producing one more unit. Economists also include opportunity cost, so normal profit is counted as a cost.

Marginal cost is the key cost for the output decision. A profit-maximising firm produces where MC = MR, with MC rising and cutting MR from below. If MR > MC, producing one more unit adds more to revenue than to cost, so profit rises if output increases; if MC > MR, profit rises if output is cut. The price is then set from the average revenue (demand) curve at that output; in perfect competition the firm takes the market price.



Average total cost determines whether the firm makes a profit and is key in the long run. At price P1, the firm produces Q1 where MC = MR1, and because AR is

Question number

above ATC it makes supernormal profit. At P2, AR = ATC at the break-even point and only normal profit is made. In the long run all costs are variable, so the firm must cover ATC: if price stays below ATC it will leave the industry.

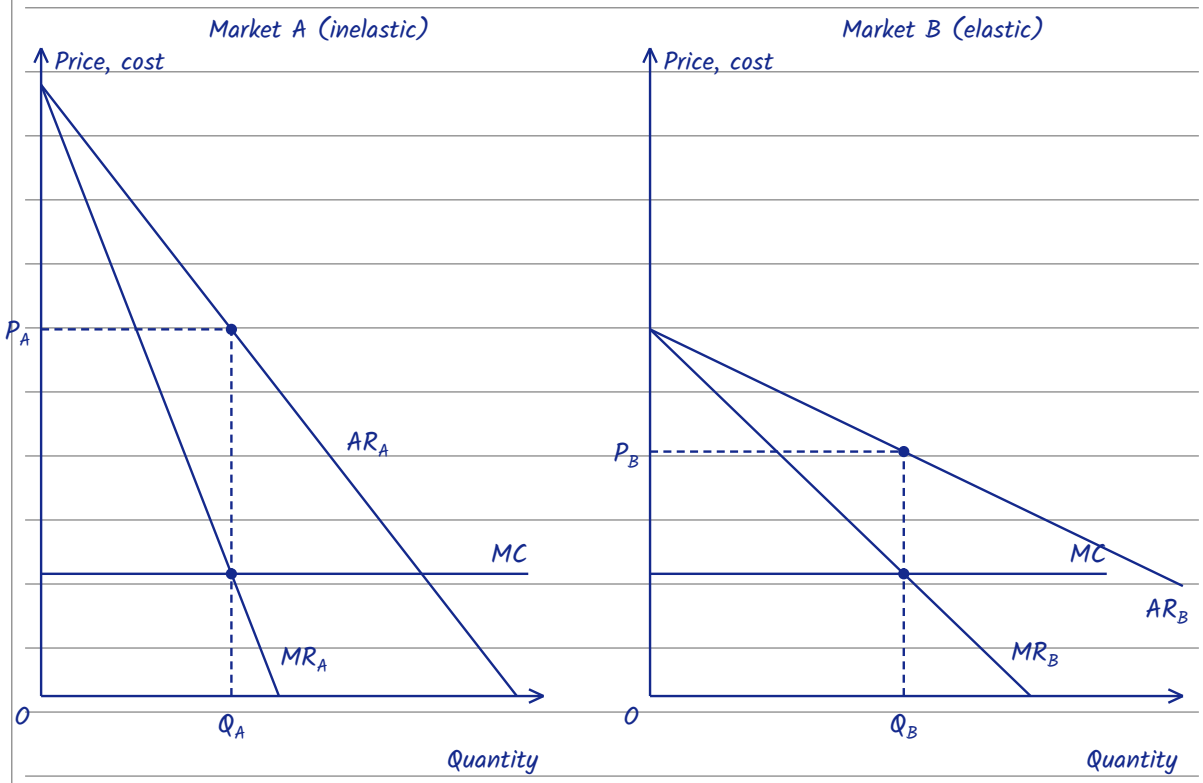
Average variable cost is key in the short run. Fixed costs must be paid whether or not the firm produces, so they do not affect the decision to produce in the short run. As long as price is at least equal to AVC, as at P3, the firm covers its variable costs and makes some contribution to fixed costs, so it should keep producing even though it makes a loss. If price falls below the minimum of AVC, the shut-down point, it loses less by closing down and paying only its fixed costs. So a profit-maximising firm should consider MC to choose output, AVC to decide whether to produce in the short run and ATC to decide whether to stay in the industry in the long run. In practice firms may not know their MC precisely, so many use average cost plus a mark-up when setting prices.

4(b)

Price discrimination is when a firm sells the same product to different consumers at different prices for reasons not related to differences in cost, for example peak and off-peak rail fares or student discounts.

It is not always possible. Three conditions are needed. First, the firm must have price-setting power, so it cannot happen in perfect competition, where firms are price takers. Second, the markets must be separated by time, age, location or income, and consumers must not be able to resell the product from the cheap market to the expensive one; this is easier for services such as haircuts and travel than for goods. Third, the price elasticity of demand must differ between the markets, and the cost of keeping them separate must be lower than the extra revenue gained.

Question number



The diagram shows third-degree price discrimination with constant MC. The firm maximises profit where MR = MC in each market. In market A demand is price inelastic, so it charges a high price PA; in market B demand is elastic, so it charges a lower price PB. Total profit is higher than with a single price.

It is often seen as undesirable. Consumers in the inelastic market pay more and lose consumer surplus, which becomes extra producer surplus, so income is redistributed from consumers to producers. Under first-degree discrimination all consumer surplus is taken. Higher profits may also be used to strengthen barriers to entry.

However, it is not 'never desirable'. Consumers in the elastic market, often those on lower incomes such as students and pensioners, pay a lower price than they would with a single price, and some who would otherwise be priced out can now buy the product, so output may increase. Price discrimination can also allow a firm to survive: a rural bus service or a rail company may find that at any single price AR is below ATC, but by charging different prices its total revenue covers total cost, so the service is provided rather than shut down. Extra profit can finance

Question number

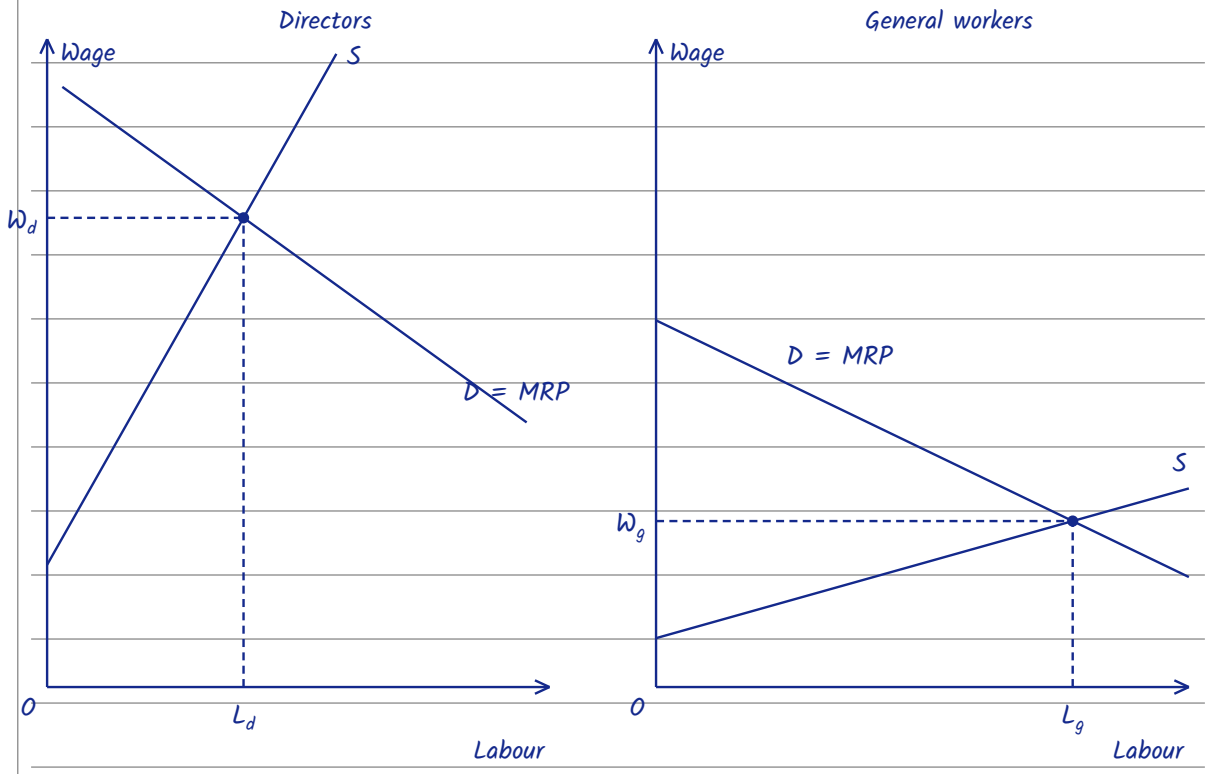
research and development, for example drug firms charging less in low-income countries. Under first-degree discrimination the firm produces up to the point where P = MC, the allocatively efficient output.

In conclusion, I disagree with the opinion. Price discrimination is not always possible, as it depends on market power, separate markets and different elasticities. It is not always undesirable either: whether it is desirable depends on whether output increases, who gains and loses, and whether it keeps a valuable service running. It is most harmful when a monopoly uses it simply to take consumer surplus.

5 A wage differential is the difference in wage rates between groups of workers. It can mean different wages for different occupations, such as doctors and cleaners, different wages for the same job in different firms, regions or sectors, or differences between workers with different skills, experience, gender or ethnicity. The view makes three claims: that theory explains differentials, that they reflect the power of firms to exploit workers, and that they are unjust. Each needs to be assessed.

Economic theory explains wages through demand and supply. The demand for labour is a derived demand, based on marginal revenue product (MRP = marginal physical product × marginal revenue). A profit-maximising firm employs workers up to the point where the wage equals MRP, so the MRP curve is the demand curve for labour. The supply of labour depends on the wage, the skills and training required, non-monetary benefits and the size of the working population. In a competitive labour market the wage is set where demand equals supply.

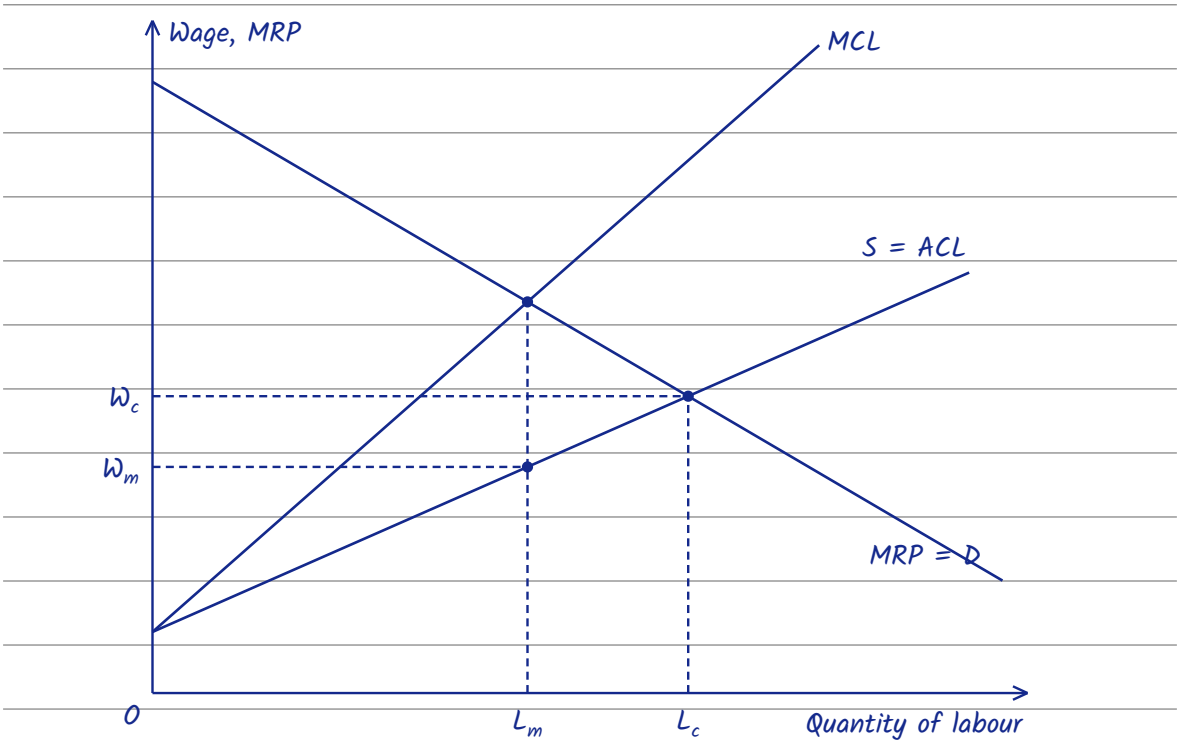
Question number



The diagram shows how this explains differentials between occupations. For skilled jobs such as company directors, MRP is high, and supply is low and inelastic because few people have the ability, qualifications and experience, so the wage Wd is high. For general workers MRP is lower and supply is large and elastic, so the wage Wg is low. The high wage includes a large element of economic rent, the payment above transfer earnings. Theory also explains compensating differentials, where dangerous or unpleasant jobs pay more, and differences between firms when productivity or the price of the product differ, since these change MRP.

The second claim is that differentials show the power of a firm to exploit its workforce. This is true where there is monopsony, a single or dominant buyer of labour, such as a large employer in a small town. To employ more workers the monopsonist must raise the wage for all of them, so the marginal cost of labour (MCL) is above the average cost of labour (the supply curve).

Question number



The monopsonist employs Lm where MCL = MRP, but pays only Wm, read from the supply curve. In a competitive market the wage would be Wc and employment Lc. So the monopsonist pays a lower wage and employs fewer workers, and workers are paid less than their MRP; this gap is what economists mean by exploitation of labour. So some differentials, for example between workers doing the same job for a monopsony employer and a competitive one, are due to the power of the firm. However, many differentials are not caused by firms' power. They come from differences in MRP and in labour supply, which reflect differences in skills, productivity and the demand for the product. Workers can also have power: strong trade unions can raise wages above the competitive level, and when a union faces a monopsonist (bilateral monopoly) the wage depends on bargaining strength. Wages in the public sector are often set by governments for political reasons, and minimum wage laws limit how low wages can go. Immobility of labour, imperfect information and discrimination also cause differentials that the simple model does not predict, so the theory explains differentials well in principle but needs these imperfections to explain real-world wages.

Question number

Whether differentials are unjust is a normative judgement. Differentials perform a useful economic function: they act as signals and incentives that move labour into jobs where it is most valued and reward people for the cost and time of training, so removing them would reduce efficiency. Differentials that reflect differences in productivity, skill or unpleasant conditions can therefore be seen as fair. But differentials caused by monopsony exploitation, discrimination by gender or ethnicity, or directors effectively setting their own pay are hard to justify, and very large differentials create inequality of income and wealth.

In conclusion, I only partly agree. Economic theory, using MRP and labour supply, does explain most wage differentials. Some, where firms have monopsony power, do show exploitation and may be unjust, but many result from differences in skills and productivity, or from the power of unions and governments, and these play an important role in allocating labour. So differentials are not always a sign of exploitation, nor always unjust.

6(a) The public sector budget deficit is a flow: it is the amount by which government spending exceeds government revenue (mainly taxation) in a particular time period, usually one year. The national debt is a stock: it is the total accumulated amount the government owes, the sum of all past borrowing not yet repaid, for example through selling government bonds to the private sector, banks and foreign investors.

The two are linked, because each year's deficit must be financed by borrowing, so it is added to the national debt. A budget surplus allows part of the debt to be repaid. Both are usually compared using a percentage of GDP, because a larger economy can afford to borrow more.

The budget deficit is important because it shows the current fiscal stance. A deficit injects demand into the economy, so it can help reduce unemployment in a recession, partly through automatic stabilisers. But a large deficit in a boom can

Question number

cause demand-pull inflation, financing it by borrowing may raise interest rates and crowd out private investment, and financing it by creating money can be inflationary. A deficit financed from abroad may also worsen the balance of payments.

The national debt is important because interest has to be paid on it every year. Debt interest has an opportunity cost, since the money could instead be spent on education or healthcare, and it shifts the burden to future generations, who face higher taxes. A high debt may lead to a lower credit rating, so lenders demand higher interest rates, and in extreme cases a debt crisis, as in Greece after 2010. Which is more important depends on the situation. A large deficit for one or two years in a recession is not serious if the debt is low; the UK ran a large deficit in 2020 during the pandemic. In the long run, the debt matters more, since persistent deficits build up debt, and its sustainability depends on whether the interest rate is below the growth rate of GDP, so that the debt-to-GDP ratio can fall. It also matters whether the debt is owed in domestic currency to domestic lenders (as in Japan, with debt over 200% of GDP) or in foreign currency. Overall, the national debt relative to GDP is the more important indicator of long-term risk, while the deficit matters more for short-term macroeconomic management and shows the direction the debt is moving.

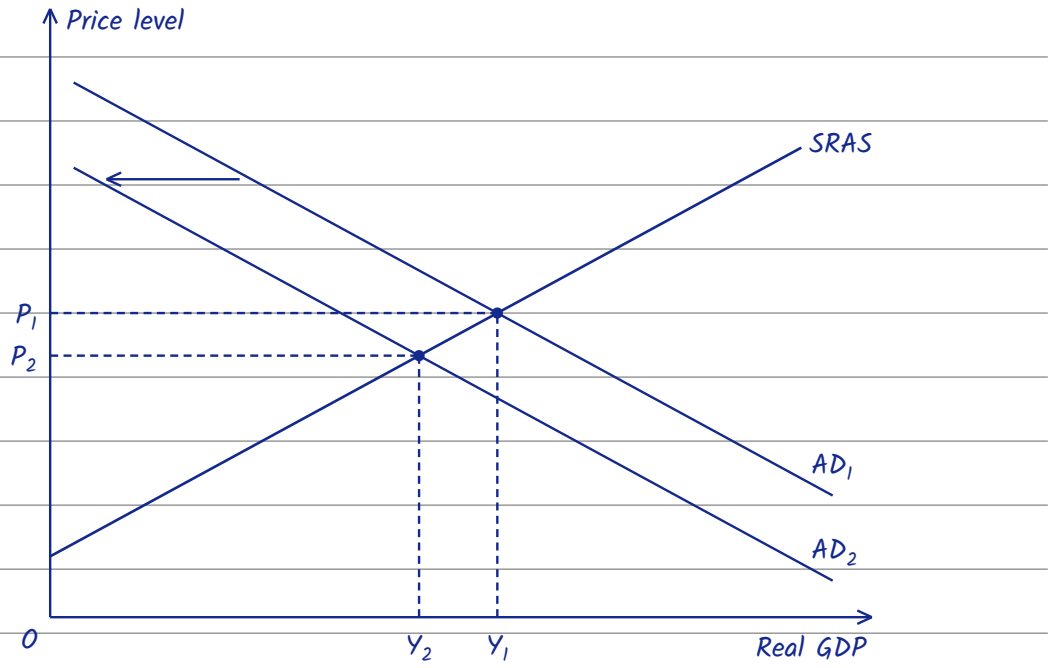
6(b) A government can reduce a budget deficit in the short run by raising taxes or cutting public spending, known as fiscal consolidation or austerity, and in the long run by policies that raise economic growth.

Raising taxes, such as income tax or VAT, increases revenue directly. Indirect taxes are quick to raise and hard to avoid. However, higher taxes reduce disposable income and consumption, so AD falls and, through the multiplier, national income falls by more than the initial change. Lower incomes reduce tax revenue and raise spending on benefits, so the deficit falls by less than expected. Very high income

Question number

tax rates may reduce the incentive to work and increase avoidance, and VAT is regressive, so it worsens inequality.

Cutting government spending reduces the deficit directly and may avoid the disincentive effects of higher taxes. Cuts to capital spending are politically easier but harm long-run growth, while cuts to current spending, such as public sector wages and benefits, are unpopular. Again, lower spending reduces AD.



The diagram shows that austerity shifts AD left from AD1 to AD2, so real GDP falls from Y1 to Y2 and unemployment rises. If the economy is already in recession, this is self-defeating, since automatic stabilisers increase the cyclical deficit. If the fiscal multiplier is small, or the economy is near full capacity, the cost in lost output is smaller and the policy is more effective.

Monetary policy can support fiscal consolidation: lower interest rates reduce the cost of servicing the national debt and encourage investment and consumption, offsetting the fall in AD. However, this is limited when interest rates are already close to zero.

In the long run, supply-side policies such as spending on education and training, infrastructure, and tax reforms that encourage enterprise can increase productive

Question number

capacity and shift LRAS to the right. Faster growth raises incomes and employment, so tax revenue rises and benefit spending falls without increasing tax rates. This is the least damaging way to reduce the deficit, but it is slow, its effects are uncertain and some policies initially increase spending.

In conclusion, the most effective approach depends on whether the deficit is cyclical or structural. A cyclical deficit falls automatically when growth returns, so cutting it in a recession is costly. A structural deficit needs tax rises or spending cuts, best introduced gradually once the economy is recovering, combined with supply-side policies for long-run growth and supportive monetary policy.

7(a) The quantity theory of money is based on the Fisher equation of exchange, $MV = PT$, where M is the money supply, V is the velocity of circulation (the number of times each unit of money is spent in a period), P is the general price level and T is the number of transactions (or real output, Y). MV is total spending and PT is the value of what is sold, so the equation is an identity: it is always true by definition.

The quantity theory turns this identity into a theory of inflation by making two assumptions. First, V is constant, or stable and predictable, because it depends on slowly changing institutional factors such as how often people are paid. Second, T is constant, because the economy tends towards full employment, so real output is determined by supply-side factors and does not depend on the money supply. It follows that any increase in M leads to a proportional increase in P: if the money supply rises by 10% faster than output, prices rise by 10%. Monetarists such as Milton Friedman used this to argue that inflation is always caused by excessive growth of the money supply, so it should be controlled by limiting money supply growth.

There is some evidence for the theory in the long run: in hyperinflations, such as Zimbabwe in 2008, very rapid money creation was followed by very rapid inflation.

Question number

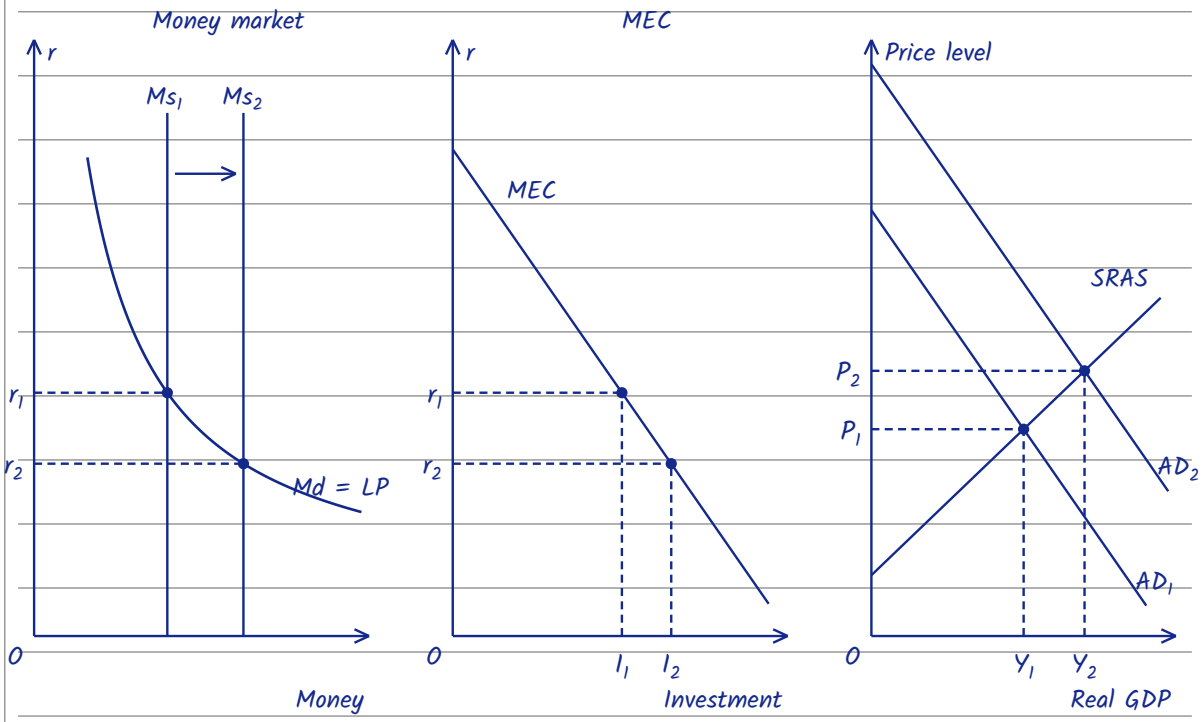
However, the assumptions can be criticised. Velocity is not constant: Keynesians argue that when the money supply rises, people may hold extra money as idle balances (the speculative demand for money), so V falls and offsets the rise in M. In many countries velocity fell sharply after quantitative easing from 2009, so huge increases in the money supply caused little inflation. Velocity also changes with financial innovation such as card and online payments.

T is not constant when there is spare capacity: if the economy is below full employment, extra spending increases real output rather than prices. Causation may also run the other way, from P to M: rising prices and wages lead firms and households to borrow more, so banks create more money in response.

Finally, the money supply is hard to measure, since there are several measures, narrow and broad, and they often move differently, and it is hard to control, because commercial banks create most money through credit creation. So the theory is a useful explanation of long-run inflation, but it is weak in the short run.

7(b) The monetary transmission mechanism is the process by which a change in the money supply or interest rates affects real variables such as output, employment and the price level. In the Keynesian version it links the money market to the goods market through the rate of interest.

Question number



Suppose the central bank increases the money supply from Ms1 to Ms2, for example by buying bonds. People now hold more money than they want, so they use the excess to buy bonds. This raises bond prices and, because bond prices and interest rates are inversely related, the interest rate falls from r1 to r2 along the liquidity preference (Md) curve. The lower interest rate reduces the cost of borrowing, so firms invest in projects whose expected return (marginal efficiency of capital) is now above the interest rate: investment rises from I1 to I2 along the MEC curve. Investment is an injection, so AD shifts from AD1 to AD2, and through the multiplier real GDP rises from Y1 to Y2 by more than the initial rise in investment, raising employment, with some rise in the price level from P1 to P2. Lower interest rates also raise consumption of durables bought on credit, raise asset prices (a wealth effect) and depreciate the exchange rate, increasing net exports.

However, each link can be weak. In a liquidity trap, when interest rates are already very low, the Md curve is almost horizontal, so extra money is held as idle balances and the interest rate does not fall; many economies faced this after 2008 with rates close to zero. If the MEC curve is interest inelastic, because

Question number

investment depends mostly on business expectations ('animal spirits'), a fall in interest rates causes only a small rise in investment; in a deep recession firms may not invest at all. The multiplier may be small if leakages, such as saving, taxes and imports, are high, as in a very open economy.

There is also a negative feedback effect: as income rises, the transactions demand for money increases, shifting Md to the right, which pushes interest rates back up and partly offsets the original fall. Commercial banks may not pass on lower rates or may be unwilling to lend, and the whole process has long and variable time lags, often up to two years.

In conclusion, the transmission mechanism explains how monetary policy can affect the real economy, and it works reasonably well in normal times when interest rates are well above zero and confidence is stable. It is much less effective in a deep recession, when a liquidity trap, low confidence and an inelastic MEC weaken the links, which is why governments have used quantitative easing and fiscal policy alongside it.