

Cambridge International A Level Economics (9708)

Lecture Notes: A2 Level

Teacher: Abdur Rehman Shahzaib

Contents

1. Perfect Competition
2. Efficiency (with real-world examples)
3. Monopoly, Imperfect Competition and Price Discrimination
4. Monopolistic Competition and Oligopoly

Cambridge International A Level Economics (9708)

Perfect Competition

A2 Lecture Notes with Diagrams and a Model Essay

Market structures | The price taker | Revenue: MR-DARP | Profit maximisation ($MC = MR$) | The four short-run positions | The long run | Past paper essay 9708/44/O/N/25

Prepared by: Abdur Rehman Shahzaib

1. Where Perfect Competition Fits: Market Structures

Different markets are organised in different ways. The way a market is organised is called its **market structure**, and two questions decide which structure a market has:

- **The number and size of firms:** is the market shared by thousands of tiny sellers, a handful of giants, or just one firm?
- **The nature of the product:** do all firms sell an identical product, or does each firm sell something slightly different?

Perfect competition sits at one extreme end of this spectrum: the perfect market. The word “perfect” appears three times in its logic, and this is a useful way to introduce it to the class: the products are **perfect substitutes** for each other, each firm faces **perfectly elastic demand**, and buyers and sellers have **perfect knowledge** of prices and products. At the other extreme sits monopoly, with one firm and no substitutes; every real market lies somewhere between the two.

2. Characteristics of Perfect Competition

Perfect competition: a market structure with a very large number of small firms selling an identical (homogeneous) product, where no single firm is big enough to influence the market price.

- **Large number of small firms.** Thousands of producers, each with the smallest possible market share, so no firm's output decision can move the market price.
- **Homogeneous (standardised) products.** The output of different firms cannot be told apart; one farmer's wheat is exactly like another's. Buyers therefore choose on price alone.
- **Same price, price taker.** Every firm charges the market price for every unit. A firm takes the price the market sets; it cannot choose its own.
- **Perfect knowledge.** Buyers know every seller's price; sellers know the technology and profits available. Nobody can be overcharged, because a cheaper identical unit is always known and available.
- **No advertising.** With identical products and perfect information there is nothing to advertise; branding cannot exist.
- **No barriers to entry or exit.** New firms can enter freely, and leaving costs nothing (sunk costs are zero). This single assumption drives everything that happens in the long run.
- **Objective: survival through efficiency.** With no power over price, a firm's only route to profit is producing at the lowest possible cost.

Close real-world examples. No real market is perfectly competitive, but some come close: markets for wheat and other grains, fresh vegetable and fish markets where dozens of stalls sell the same produce, and foreign currency exchange, where the product (a dollar) is perfectly identical whoever sells it.

3. The Assumptions Versus the Real World

Perfect competition is a hypothetical scenario, a benchmark that real markets are measured against, not a description of them. Each assumption breaks down somewhere in reality, and this table becomes evaluation material in essays:

Assumption in the model	What happens in reality
-------------------------	-------------------------

Identical products	Firms differentiate their products through quality, design and branding
Infinite (very many) producers	Not possible in practice; most markets support a limited number of firms
Perfect knowledge	Information is asymmetric: the seller usually knows more than the buyer
No barriers to entry or exit	Sunk costs exist, so entering and leaving a market is never free
No advertising	Advertising is everywhere and clearly does shape consumer preferences

4. The Firm as a Price Taker

Price taker: a firm that must accept the price determined by the whole market, because its own output is far too small a share of total supply to influence that price.

The price is set where the **industry** demand curve (all consumers together) meets the industry supply curve (all firms together). Each individual firm then faces that price as a fact of life, so the demand curve for the **individual firm** is a horizontal line at the market price. This two-panel diagram is the signature diagram of the topic; students should practise drawing the pair side by side.

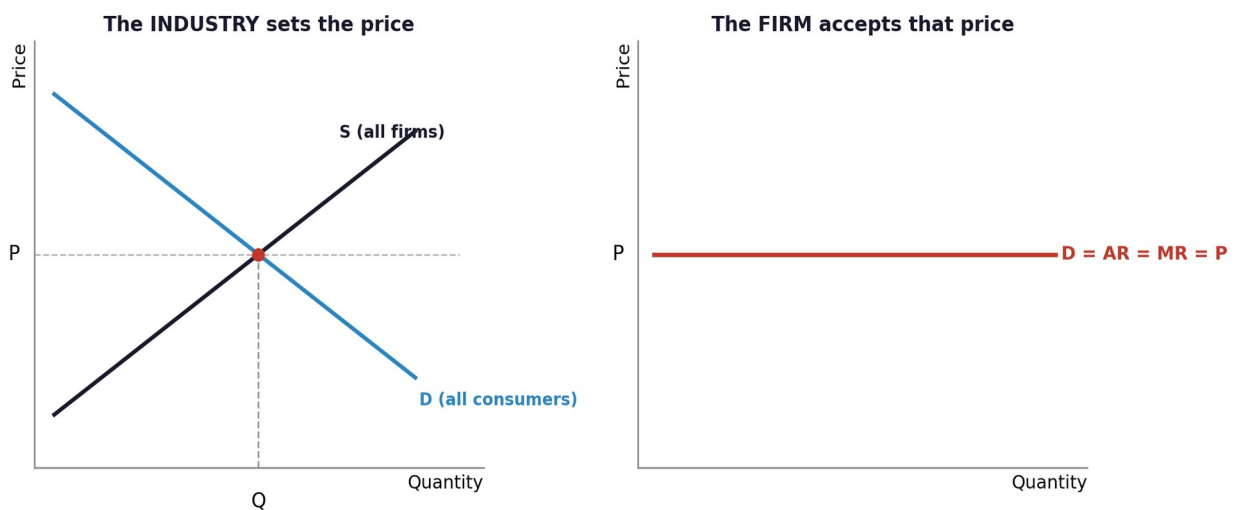


Figure 1: The industry's demand and supply set the price P . Each firm can sell any quantity it likes at P , so the firm's demand curve is horizontal: perfectly elastic.

Why is the firm's demand curve horizontal?

- **If the firm raises its price above P :** buyers switch instantly to the thousands of identical rivals (perfect knowledge, perfect substitutes). It sells nothing and total revenue collapses to zero, so no firm ever prices above the market.
- **If the firm cuts its price below P :** it gains nothing, because it could already sell every unit it produces at the full market price. Cutting the price only sacrifices revenue on every unit.

- **The result:** the firm sells everything at exactly P, so its total revenue rises in a straight line with output: sell 10% more units and total revenue rises by exactly 10%.

5. Revenue in Perfect Competition: MR-DARP

Total revenue (TR): the total amount a firm earns from selling its output. $TR = \text{price} \times \text{quantity sold}$.

Average revenue (AR): revenue earned per unit sold. $AR = TR \div Q = (P \times Q) \div Q = P$, so average revenue always equals the selling price.

Marginal revenue (MR): the additional revenue from selling one more unit. $MR = \Delta TR \div \Delta Q$.

In perfect competition every unit sells at the same market price, so each extra unit adds exactly the price to total revenue. Work through this table with a market price fixed at \$5:

Price (\$)	Quantity	TR (\$)	AR (\$)	MR (\$)
5	1	5	5	5
5	2	10	5	5
5	3	15	5	5
5	4	20	5	5
5	5	25	5	5

Every column tells the same story: AR and MR are both stuck at \$5, the market price. So one horizontal line on the diagram carries four names at once, and the class mnemonic for it is **MR-DARP**:

$$D = AR = MR = P \quad (\text{one horizontal line, four labels})$$

Every diagram in this topic is built on that line. Label it fully in the exam, writing $D = AR = MR$, because each label carries meaning: it is the firm's demand curve, its revenue per unit, and its revenue from the next unit, all equal to the market price.

6. Profit Maximisation: $MC = MR$

A firm maximises profit at the output where **marginal cost equals marginal revenue ($MC = MR$)**. The logic works from both sides of that point, and the arrows in Figure 2 are worth reproducing on the board:

- **Below Q^* ($MR > MC$):** the next unit brings in more revenue than it costs to make. Producing it adds to profit, so expand output.
- **Beyond Q^* ($MC > MR$):** the last unit cost more than it earned; it created a loss. Cut output back.
- **At Q^* ($MC = MR$):** no unit can be added or removed without reducing profit. This is the profit-maximising output.

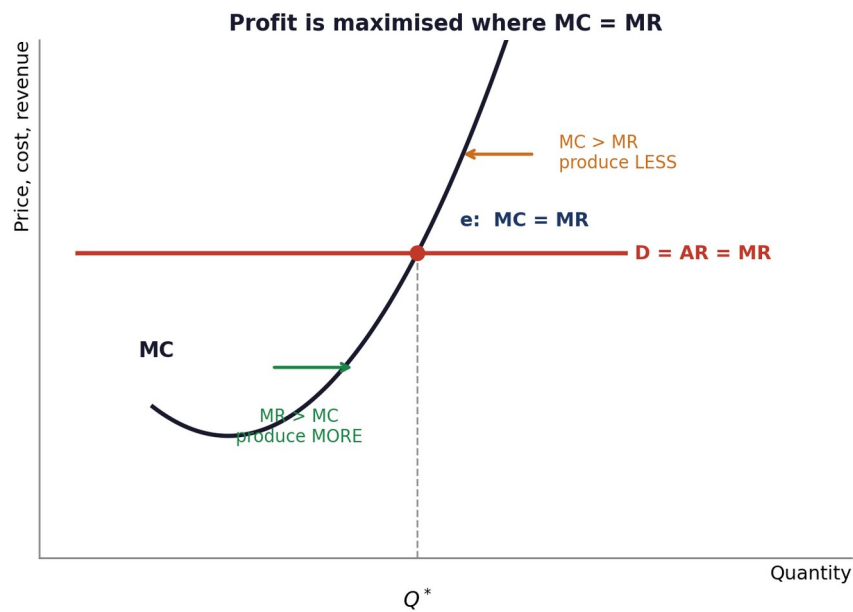


Figure 2: At point e the rising MC curve cuts the horizontal MR line. Left of Q^* each unit adds profit; right of Q^* each unit adds loss.

Two remarks that help in every later diagram. First, in this topic the cost curves stay the same in every scenario; it is the revenue line that moves up or down over time. Second, a firm's survival task in the short run is not to cover total cost but to cover its **average variable cost**, a point that becomes the shutdown rule below.

7. The Four Short-Run Positions

In the short run the firm has both fixed and variable costs, and the market price can sit at four different levels relative to the firm's cost curves. The output choice is always $MC = MR$; what changes is the profit or loss at that output, read by comparing AR with ATC .

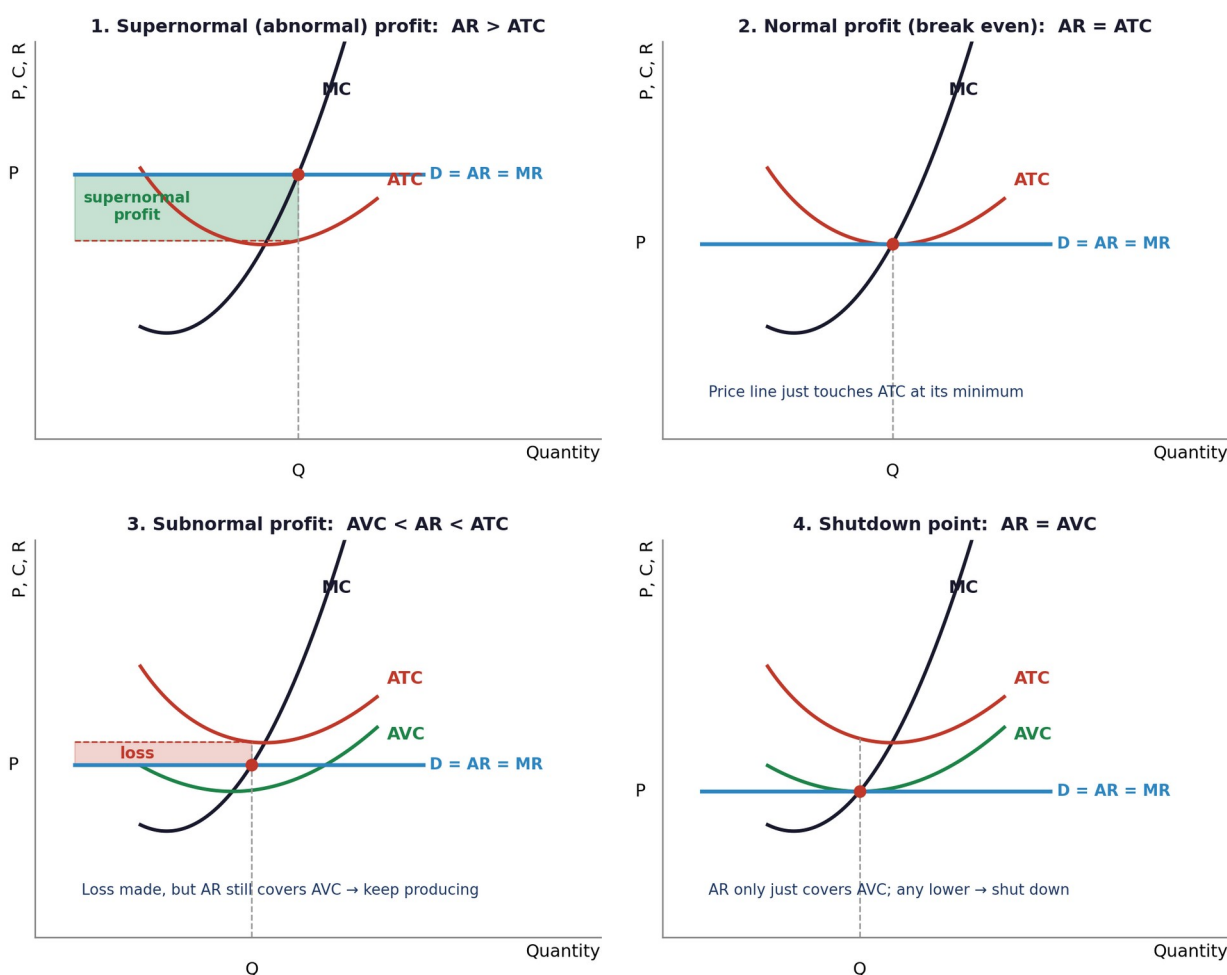


Figure 3: The four short-run positions. In each panel output is where $MC = MR$; the gap between AR and ATC at that output gives the profit or loss per unit.

Position	Condition	What it means
1. Supernormal (abnormal) profit	$AR > ATC$	The price more than covers all costs; profit above normal is earned (shaded rectangle)
2. Normal profit	$AR = ATC$	Break even: revenue just covers all costs including the reward needed to keep the entrepreneur in this industry
3. Subnormal profit (loss)	$AVC < AR < ATC$	A loss is made, but price still covers variable costs and contributes something towards fixed costs, so the firm keeps producing
4. Shutdown point	$AR = AVC$	Price only just covers variable cost; below this the firm loses less by closing at once

The shutdown rule and the idea of contribution

Why does a loss-making firm keep going in position 3? Because in the short run the fixed costs (rent, machinery) must be paid whether the firm produces or not. As long as the selling price is above variable cost per unit, every unit sold makes a **contribution**: $\text{contribution} = \text{selling price} - \text{variable cost}$. That contribution goes towards the fixed cost bill, so producing loses less money than shutting down. Once price

falls below AVC, each unit no longer even pays for its own materials and labour, and closing immediately becomes the cheaper option. In short: **operate while $AR \geq AVC$; shut down when $AR < AVC$.**

8. The Long Run: Entry, Exit and Normal Profit

The long-run story rests on one characteristic: no barriers to entry or exit. Profits and losses are signals, and firms are free to act on them.

- **If firms earn supernormal profit:** the profit attracts new entrants. Industry supply shifts right (S_0 to S_1), the market price falls, and each firm's horizontal AR line slides down. Entry continues until the price just touches the minimum of ATC, where only normal profit remains and the incentive to enter is gone.
- **If firms make subnormal profit (losses):** the opposite happens. Loss-making firms exit, industry supply shifts left, the price rises, and the AR line moves up until the survivors again earn exactly normal profit.

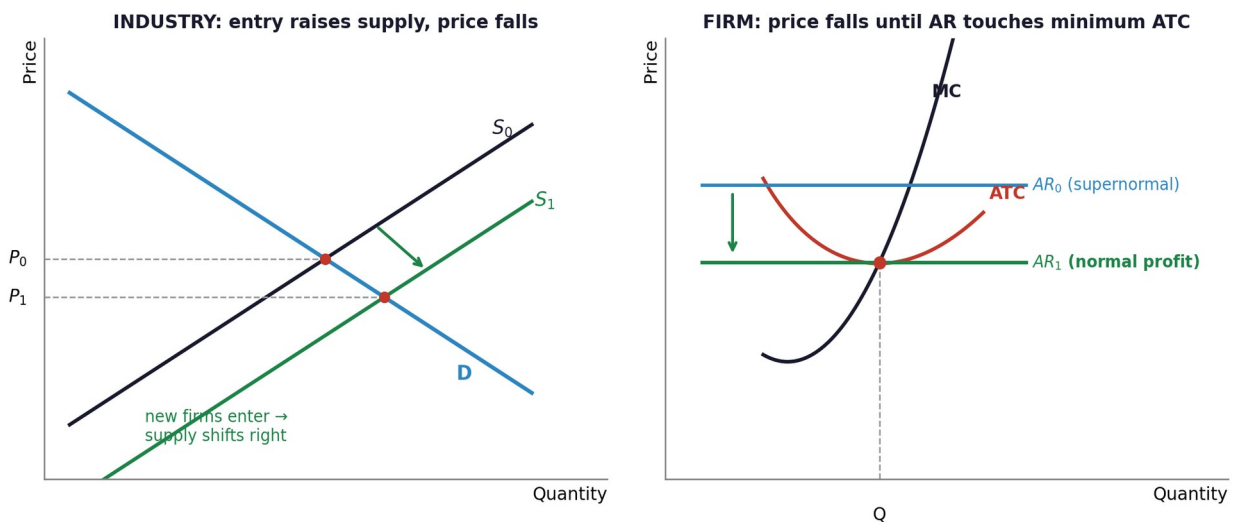


Figure 4: Supernormal profit attracts entry: industry supply shifts right, price falls from P_0 to P_1 , and the firm's AR line falls until it is tangent to ATC at its minimum, leaving normal profit only.

Time period	Positions the firm can experience
Short run	Supernormal profit, normal profit, subnormal profit, or shutdown
Long run	Normal profit only: entry competes away supernormal profit, exit eliminates losses

This long-run equilibrium is also the source of the efficiency claims made for perfect competition: the firm ends up producing at the lowest point of its average cost curve (productive efficiency) at a price equal to marginal cost (allocative efficiency). Both are worth flagging to students now, because evaluation questions on efficiency build directly on this diagram.

9. Past Paper Question with Model Answer

9708/44/O/N/25: "Subnormal and supernormal profits are only experienced in the short run and only by firms in perfect competition." With the help of diagrams, evaluate this statement. [20 marks]

Introduction and definitions (AO1)

Normal profit is the minimum reward required to keep an entrepreneur in their present industry; it is treated as a cost and is earned when average revenue equals average total cost ($AR = ATC$). Supernormal (abnormal) profit is any profit above this level, earned when $AR > ATC$, while subnormal profit means revenue fails to cover total cost, $AR < ATC$, though a firm will continue in the short run provided variable costs are covered. The statement makes two separate claims: that these profit levels exist only in the short run, and that only firms in perfect competition experience them. Each claim must be tested in turn.

Analysis: profits in perfect competition (AO2)

In perfect competition, a price-taking firm faces a horizontal demand curve where $D = AR = MR$ at the market price and produces where $MC = MR$. In the short run the market price can sit above, on, or below the firm's average total cost, so the firm may earn supernormal profit ($AR > ATC$), normal profit, or subnormal profit while price still covers average variable cost (see Figure 3, panels 1 to 3 above). Both profit positions named in the statement are therefore possible in the short run.

These positions cannot last, because perfect competition has no barriers to entry or exit. Supernormal profit attracts new firms: industry supply shifts right, price falls, and each firm's AR line slides down until it is tangent to the minimum of ATC, restoring normal profit (Figure 4). Subnormal profit causes exit, supply shifts left, and price rises until normal profit returns. So within perfect competition, the statement's first claim holds: subnormal and supernormal profits are strictly short-run experiences, and long-run equilibrium delivers normal profit only.

Evaluation: is either claim fully correct? (AO3)

The first claim, "only in the short run", is true for perfect competition but not as a general rule, because whether abnormal profit survives into the long run depends entirely on barriers to entry, not on the calendar. A monopoly protected by strong barriers, such as patents, ownership of key resources, or very large economies of scale, produces where $MC = MR$ with price above average cost and no entrant can compete the profit away. Its supernormal profit persists in the long run, as Figure 5 shows.

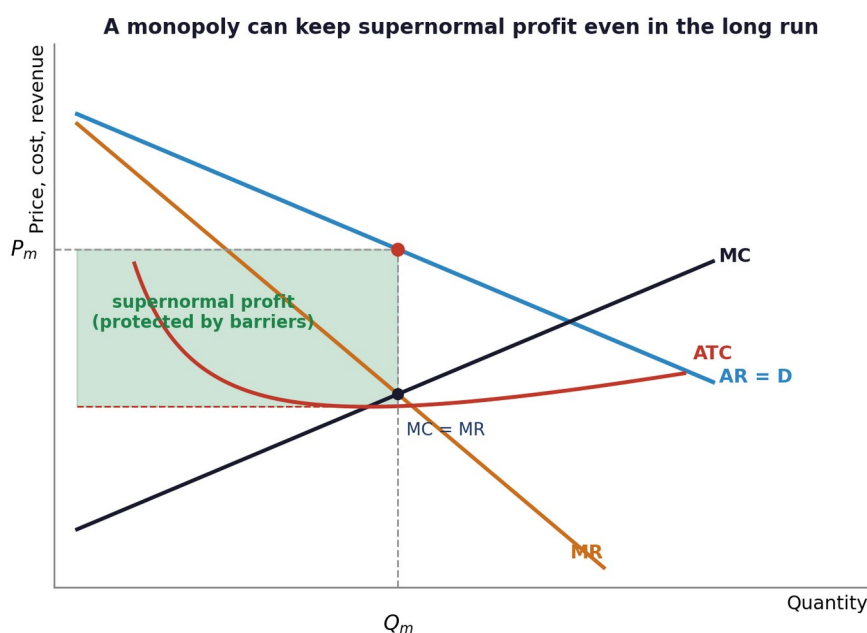


Figure 5: A monopolist earning supernormal profit (shaded). Barriers to entry stop new firms from entering, so the profit can persist into the long run, contradicting the statement.

The second claim, “only by firms in perfect competition”, is clearly wrong in both directions. Firms in every market structure can experience these profit levels in the short run: a monopolistically competitive restaurant can earn supernormal profit before rival restaurants open nearby, and any firm in any structure, including a monopoly facing collapsing demand, can make subnormal profit. Indeed, monopolistic competition behaves exactly like perfect competition on this point, since its weak barriers mean entry erodes supernormal profit in the long run too. What separates the structures is not whether abnormal profits can appear, but whether they can last.

A further qualification is the idea of contestable markets: even a monopolist may be unable to sustain supernormal profit in the long run if the threat of entry is strong enough, while, in the other direction, a perfectly competitive firm’s “short run” can in practice persist for years if the adjustment of entry and exit is slow. The clean short-run/long-run split is an analytical device rather than a fixed length of time.

Conclusion

The statement is at best half true. Within perfect competition the analysis supports it: free entry and exit guarantee that subnormal and supernormal profits are temporary, and only normal profit survives into the long run. But as a general statement it fails twice over. Short-run abnormal profits are experienced by firms in every market structure, not only perfect competition, and supernormal profit can persist into the long run wherever barriers to entry protect it, most clearly under monopoly. The decisive factor is the strength of barriers to entry rather than the market structure’s name or the time period itself.

Exam tip for this essay

A statement with two claims must be split before it is evaluated; candidates who treat it as one claim usually agree or disagree with the whole sentence and lose the evaluation marks. Diagrams carry heavy weight here: draw at least the perfect competition short-run profit and loss diagrams, the long-run adjustment, and one monopoly diagram with persistent supernormal profit. Label every curve, and make the final judgement name the real dividing line: barriers to entry.

Cambridge International A Level Economics (9708)

Economic Efficiency

A2 Lecture Notes with Diagrams and Two Model Essays

Productive efficiency | Allocative efficiency ($MSB = MSC$) | The PPC and Pareto optimality | Efficiency in perfect competition | Dynamic efficiency | Two 20-mark model essays

Prepared by: Abdur Rehman Shahzaib

1. What Does Efficiency Mean?

Efficiency is about making the most of limited resources: producing the maximum output from the minimum resources, and pushing the cost per unit down as far as it will go. Because resources are scarce, an inefficient economy is throwing away output that people wanted. The syllabus splits the idea into two branches you must keep apart:

	Productive efficiency	Allocative efficiency
Whose perspective?	the producer's: how things are made	society's: what mix of things is made
Condition	producing at the lowest point of average cost (AC at its minimum, where $MC = ATC$)	$MSB = MSC$, which is the same as price (AR) = MC
The question it answers	"Are we wasting resources in production?"	"Are resources going to the goods society values most?"
On the PPC	any point ON the curve	one single point on the curve

Real-world examples

Budget airlines show the two branches side by side. Filling every seat, fast turnarounds and one aircraft type push the cost per passenger down (productive efficiency); deciding which routes get the planes, guided by which fares passengers will pay, is about allocating resources to the flights people value most (allocative efficiency).

2. Productive Efficiency

Productive efficiency: producing output at the lowest possible average cost, using the minimum resources for the output produced. The firm operates at the minimum point of its average cost curve, where $MC = ATC$.

In the long run the benchmark is the **long-run average cost curve (LRAC)**. Each short-run average cost curve shows costs with a fixed factory size; the LRAC is the envelope beneath them all. Productive efficiency means operating at the very bottom of the LRAC, the output level called the **minimum efficient scale (MES)**, where all economies of scale have been exploited and diseconomies have not yet set in.

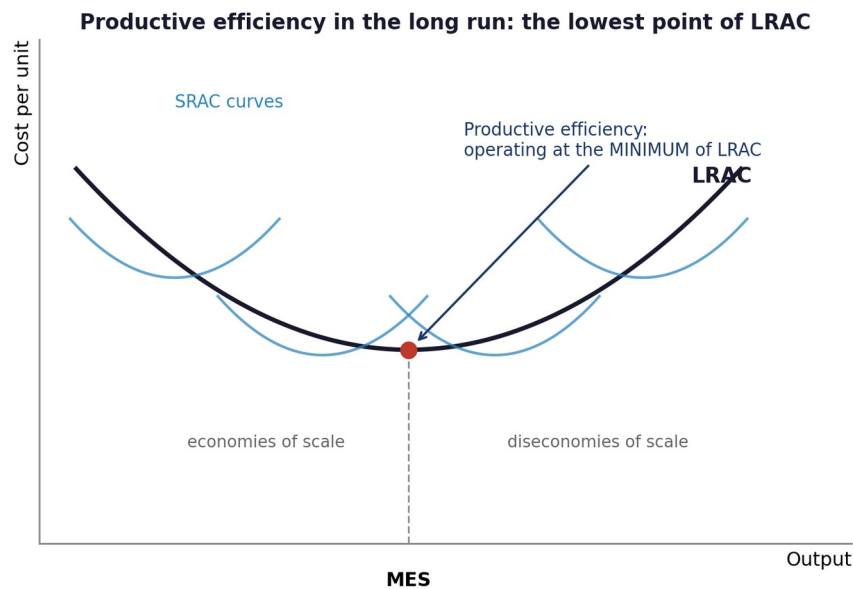


Figure 1: Productive efficiency in the long run means operating at the minimum of LRAC, the minimum efficient scale, where every economy of scale has been used up.

There is also a least-cost rule for choosing the mix of factors. A firm is combining capital and labour efficiently when the last dollar spent on each factor adds the same output: $MPK \div P_k = MPL \div P_l$. If a dollar of capital were adding more output than a dollar of labour, the firm could cut costs by switching towards capital, so the original mix was not efficient.

Real-world examples

Toyota's lean production system is the textbook case: just-in-time parts delivery, minimal stocks and continuous waste elimination pushed its cost per car below rivals for decades, and manufacturers worldwide copied the method.

Minimum efficient scale explains why some industries have few, huge producers. A modern container ship, steel mill or cement plant only reaches the bottom of its LRAC at enormous output, so small producers simply cannot match the cost per unit of the giants.

3. Allocative Efficiency

Allocative efficiency: resources are allocated so that the goods and services society values most are produced, at the output where marginal social benefit equals marginal social cost ($MSB = MSC$), equivalently where price (AR) = MC . There is no welfare or deadweight loss.

Point e in Figure 2 is the **socially optimum level of output**. Away from it, in either direction, society loses welfare:

- **Over-allocation (output beyond Q^*):** $MSC > MSB$, so the last units cost society more than the benefit they give. This is the situation with external costs, and the triangle to the right of e is the welfare loss.
- **Under-allocation (output short of Q^*):** $MSB > MSC$, so society values extra units above their cost yet they are not produced. This is the situation with external benefits, and the triangle to the left of e is the welfare loss.

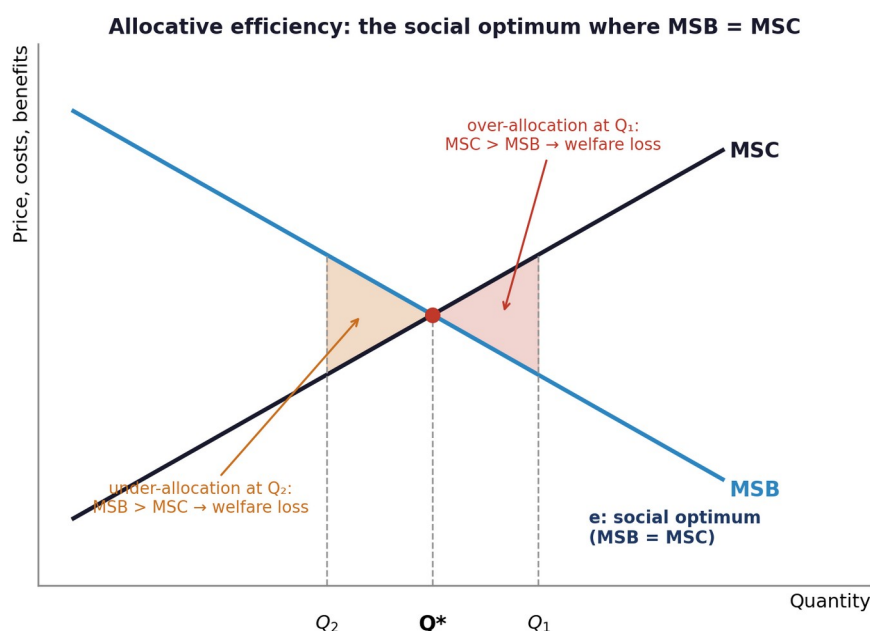


Figure 2: Allocative efficiency at e , where $MSB = MSC$. Producing beyond Q^* (over-allocation) or short of Q^* (under-allocation) both create a welfare loss, shown by the shaded triangles.

Real-world examples

Over-allocation: coal-fired electricity and petrol vehicles. Because producers and drivers ignore the pollution costs falling on third parties, MSC exceeds MSB at the market output, exactly the right-hand triangle in Figure 2.

Under-allocation: vaccinations. Each jab protects people beyond the person vaccinated, so MSB exceeds MPB and the free market delivers too few, which is why governments provide them cheaply or free.

Pricing the externality: London's congestion charge and Singapore's electronic road pricing make drivers pay towards the delay and pollution costs they impose, pushing traffic back towards the socially optimal level.

4. The PPC, Both Efficiencies, and Pareto Optimality

The production possibility curve ties the two ideas together, and the relationship between them is a favourite examiner's point. Every point on the PPC is productively efficient, since no more of one good can be made without giving up some of the other. But only one of those points matches the combination of goods that society actually wants, and only that point is allocatively efficient. So the allocatively efficient point is necessarily productively efficient, while a productively efficient point may or may not be allocatively efficient.

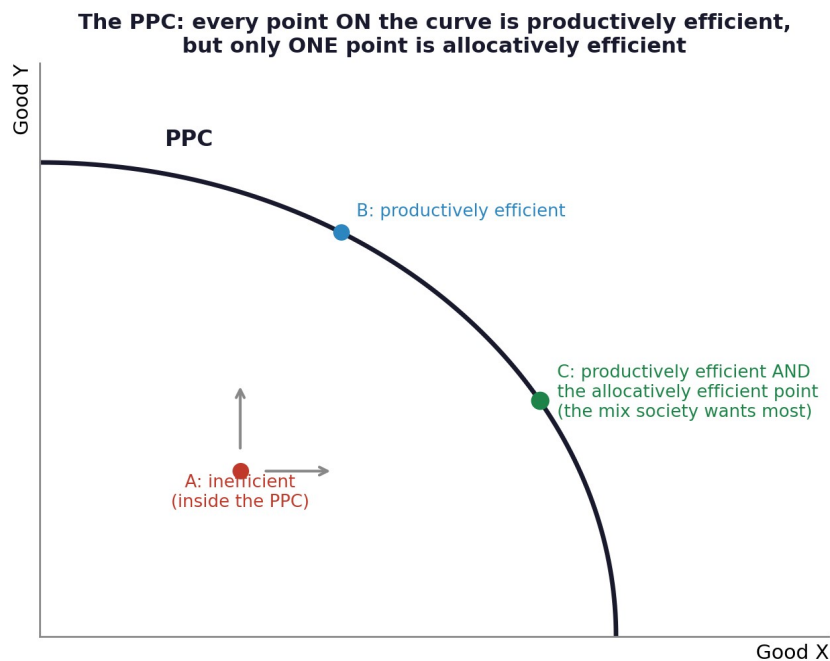


Figure 3: Point A wastes resources. Points B and C are both productively efficient, but only C, the mix society values most, is also allocatively efficient.

Pareto optimality: a situation where no one can be made better off without making somebody else worse off. It is the formal way of describing an efficient allocation and is closely related to allocative efficiency.

Inside the PPC, at a point like A, the economy can move right or up and give people more of one good at no cost in the other, so A cannot be Pareto optimal. On the curve, every gain in Good X now costs some Good Y, so nobody can be made better off without someone losing: points on the frontier are Pareto optimal.

Real-world examples

The COVID-19 lockdowns of 2020 pushed whole economies deep inside their PPCs: aircraft parked, factories dark, workers idle at home, point A on the diagram. The recovery that followed was a movement back towards the frontier, output regained without giving anything up, because the resources had simply been unemployed.

A choice ALONG the frontier appeared in the same period: manufacturers switched assembly lines from cars and clothing to ventilators, masks and sanitiser, more of one good at the cost of another, a move from one efficient point to another such as B to C.

5. Efficiency in Long-Run Perfect Competition

The long-run equilibrium of perfect competition is the standard example of both efficiencies achieved at once, which is why this diagram belongs in almost every efficiency essay. Free entry and exit push every firm to the point where its horizontal demand line just touches ATC at its minimum, and at that same point price equals marginal cost:

- **PE test:** output is at the minimum of ATC, so production wastes nothing: productively efficient.
- **AE test:** $MC = AR = \text{price}$, so the value of the last unit to consumers equals its resource cost: allocatively efficient.

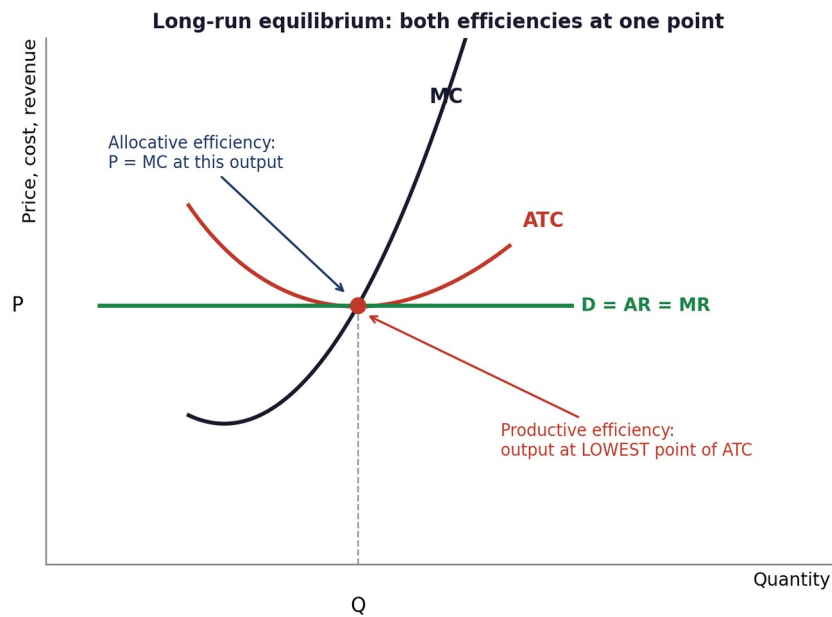


Figure 4: The long-run equilibrium of a perfectly competitive firm. Both tests are passed at the single tangency point: minimum ATC and $P = MC$.

Real-world examples

Global grain markets come closest to this outcome in practice. A wheat farmer in Punjab, Australia or Kansas sells an identical commodity at the world price, cannot influence that price, and survives only by adopting the lowest-cost seeds, machinery and methods available. Margins stay thin, close to normal profit, and any farmer whose costs stay above the best practice is eventually forced out, which is the entry-and-exit discipline of Figure 4 working in real life.

6. Dynamic Efficiency

Dynamic efficiency: efficiency achieved over time: through investment, innovation and new production processes, a firm produces more output relative to its inputs as time passes, so its whole cost curve falls.

Productive and allocative efficiency are snapshots at a moment in time; dynamic efficiency asks what happens to costs across years. A dynamically efficient firm meets the changing needs of its market by investing in research and development, better machinery and new processes. The investment often raises costs at first, and the payback arrives later, but the reward is that the entire LRAC curve shifts downward, as Figure 5 shows. A firm that never invests drifts the other way: it gradually becomes less efficient than its rivals and may be forced out of the market.

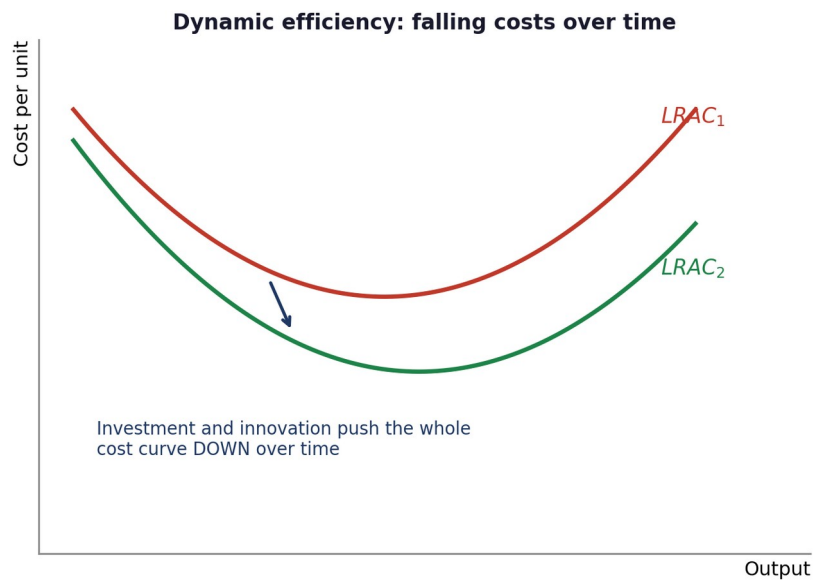


Figure 5: Dynamic efficiency. Investment and innovation shift the whole long-run average cost curve down from $LRAC_1$ to $LRAC_2$ over time.

Dynamic efficiency matters most when comparing market structures. Monopolies and oligopolies retain supernormal profits, and those retained profits are exactly what finances research, development and product innovation, which can eventually deliver lower prices and new technology to consumers. Perfectly competitive firms earning only normal profit have no such fund. So a market that fails the static efficiency tests may still beat perfect competition on efficiency over time, a point that earns evaluation marks in almost any efficiency essay.

Real-world examples

Lithium-ion batteries are the decade's clearest LRAC-shifting-down story: sustained R&D and learning-by-doing cut the cost per kilowatt-hour by roughly ninety percent between 2010 and the early 2020s, which is what made affordable electric cars possible at all. Solar panels followed the same path.

Pharmaceutical firms show the profit-to-innovation link directly: patent-protected profits on existing medicines fund the research pipelines that produce the next generation of drugs, innovation that an industry of tiny normal-profit firms could not finance.

7. Model Essay 1: Encouraging Electric Vehicles

“The increased use of electric vehicles (EVs) is encouraged as part of governments’ climate change policies because they create fewer negative externalities than diesel and petrol (gas) vehicles. Evaluate, with the help of a diagram(s), two policies that a government may use to encourage the use of EVs.” [20 marks]

Introduction and definitions (AO1)

Allocative efficiency is achieved at the socially optimum level of output where $MSB = MSC$ (equivalently $MC = AR$), with no welfare loss. Markets fail when they miss this level. Diesel and petrol vehicles carry large negative externalities, costs such as emissions imposed on third parties, so they are over-consumed. Electric vehicles, by contrast, behave like a **merit good**: their full social benefit, cleaner air and slower climate change, exceeds the private benefit drivers consider, so the free market under-consumes them, producing at Q_1 where $MSB > MSC$. Panel 1 of Figure 6 shows the under-consumption and its welfare loss. Two policies to correct this are a subsidy and an advertising campaign.

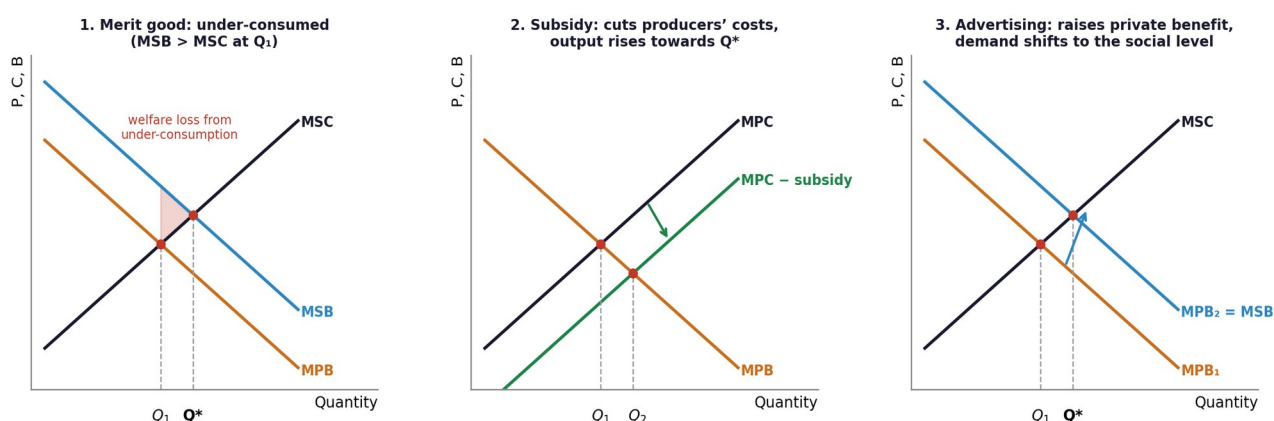


Figure 6: The market under-consumes EVs at Q_1 (panel 1). A subsidy cuts production costs and expands output towards Q^* (panel 2); advertising raises perceived private benefit and shifts demand towards the social level (panel 3).

Policy 1: a subsidy to EV producers (AO2)

A subsidy is a payment to producers that lowers the cost of making each EV. As panel 2 shows, the effective cost curve shifts down from MPC to $MPC - \text{subsidy}$, the price to consumers falls, and the quantity bought rises from Q_1 towards the social optimum. Beyond correcting the externality gap, cheaper EVs strengthen exports and economic growth in countries that build them, and the growing market encourages new businesses, charging networks, battery plants, servicing, into existence.

Policy 2: an advertising and information campaign (AO2)

Under-consumption of a merit good partly reflects imperfect information: drivers underestimate running-cost savings and the environmental gain. Government advertising attacks this cause directly. As panel 3 shows, better information raises the benefit consumers perceive, shifting MPB_1 towards MSB , so demand and quantity rise towards Q^* . Advertising can be effective and tends to have a long-term impact, since a changed attitude to EVs persists after the campaign ends.

Evaluation (AO3)

Each policy has real weaknesses. The subsidy carries an **opportunity cost**: money spent subsidising EVs is not spent on hospitals or schools. Setting its size is also guesswork, because the external benefit of an EV cannot be measured precisely, and consumers' lack of knowledge means even a well-priced EV may not

sell. Advertising is costly and time consuming, its persuasive power is uncertain, and reaching the right target market is hard, a campaign may preach to people who cannot afford a new car at all. It also does nothing to cut the price, which for most buyers is the real barrier.

The two policies are stronger together than alone, and other tools exist: an indirect tax on petrol and diesel cars attacks the rival product's over-consumption, and state provision of charging infrastructure removes a practical obstacle. The best mix depends on the government's budget and how far consumers' information, rather than price, is the binding constraint.

Conclusion

Both policies push the EV market towards the socially optimal level, the subsidy through price, advertising through information. Since the market failure has both a price cause and an information cause, government intervention will be needed on both fronts, and more than one set of policies used together, ideally financed partly by taxing the polluting substitutes, gives the best prospect of reaching allocative efficiency.

Real-world examples

Application that lifts this essay: Norway used generous tax exemptions and perks for EV buyers and by the mid-2020s electric cars made up the overwhelming majority of its new car sales, the strongest evidence anywhere that these incentives shift Q_d towards Q^* . The United States attached large EV tax credits to its Inflation Reduction Act, the UK ran plug-in grants for years, and China's long subsidy programme built the world's biggest EV market. On the other side of the market, many of the same governments tax petrol heavily and have announced future bans on new petrol and diesel car sales, taxing the substitute while subsidising the merit good.

8. Model Essay 2: Removing Subsidies from Private Education

"To improve allocative efficiency economists frequently advise governments to remove existing subsidies to the private sector providers of education. With the help of a diagram, evaluate this advice." [20 marks]

Introduction and definitions (AO1)

Allocative efficiency exists where $MSB = MSC$ ($MC = AR$), the output with no welfare loss. To judge the advice, education's nature matters. Education is not a public good: a school place is excludable, rivalrous and rejectable, so it is a **private good** that markets can provide, unlike a pure public good, which is non-excludable, non-rivalrous and non-rejectable and which markets fail to provide at all. Education is, however, a **merit good**: it generates external benefits, a more skilled, productive and informed society, that individuals ignore when deciding how much schooling to buy. Left alone, the market therefore under-consumes education, producing at Q_1 in panel 1 of Figure 7, where $MSB > MSC$ and the shaded welfare loss exists.

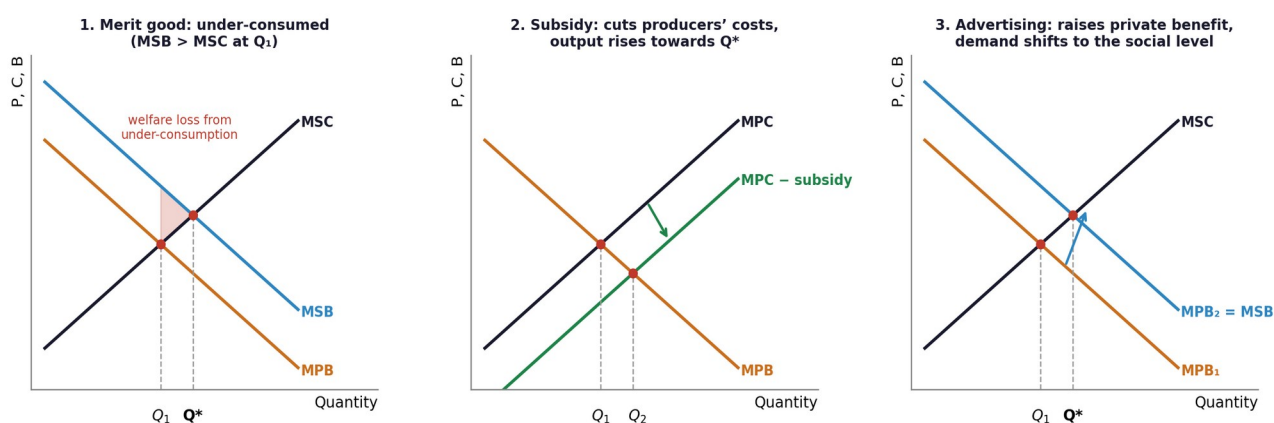


Figure 7: Education is under-consumed at Q_1 (panel 1). The existing subsidy moves provision from Q_1 towards the optimum (panel 2); removing it reverses that movement. Advertising is a partial alternative (panel 3).

Analysis: what the subsidy does, and what removing it would do (AO2)

The existing subsidy lowers private schools' costs, shifting MPC down to $MPC - \text{subsidy}$ in panel 2, cutting fees and raising enrolment from Q_1 towards Q^* . Its benefits reach beyond the classroom: lower fees make schooling affordable for poorer families so skills are enhanced, new schools are encouraged, which creates employment, and the standard of living rises with a more educated workforce. Removing the subsidy runs this film backwards: fees rise, enrolment falls back towards Q_1 , and the welfare loss from under-consumption widens. Far from improving allocative efficiency, removal moves output further from $MSB = MSC$.

Evaluation (AO3)

The economists' advice is not baseless, and a fair answer concedes its logic. Subsidies carry an opportunity cost, the funds could build state schools or hospitals, and because the external benefit of education cannot be measured, the subsidy may overshoot the optimum; an output pushed beyond Q^* creates its own welfare loss, which is presumably the economists' fear. Subsidised private providers may also pocket part of the payment rather than cut fees, muting the effect on enrolment.

Yet the case for keeping support remains stronger. If the subsidy is withdrawn, an alternative policy such as an advertising campaign (panel 3) could raise demand by reducing families' lack of knowledge about education's returns, and its effects are long term. But advertising is expensive, slow, and unlikely to help

families who understand education's value perfectly well and simply cannot pay higher fees, exactly the group the subsidy protects. The two instruments treat different causes: advertising treats information failure, the subsidy treats affordability.

Conclusion

The advice should be rejected in its strong form. Education is a merit good already under-consumed by the market, and the subsidy is what holds provision near the socially optimal level, so removing it would widen the very allocative inefficiency the economists want to cure. The sensible refinement is not removal but better design, reviewing the subsidy's size against the best available estimate of the external benefit, and supplementing it with information campaigns, so that both the affordability cause and the information cause of under-consumption are addressed at once.

Real-world examples

Application that lifts this essay: Pakistan's Punjab Education Foundation pays low-cost private schools a per-student subsidy so that children who could never afford fees attend free, a direct real-world version of panel 2, and a policy whose removal would visibly push enrolment back towards Q_1 . Chile has run a nationwide voucher system funding private school places for decades, and several countries subsidise private universities' research. Citing a scheme like these, and asking what would actually happen to enrolment if it were withdrawn, turns an abstract diagram into applied evaluation.

Exam tips for efficiency essays

Open every efficiency essay with both definitions and their conditions (minimum AC; $MSB = MSC$ or $P = MC$), then anchor each policy or claim to a labelled diagram: the socially optimum point e , the market output, and the welfare loss triangle. Both model essays above follow the same skeleton: define → diagnose the market failure with a diagram → analyse each policy with its own diagram → evaluate with opportunity cost, measurement problems and information problems → conclude with a direct judgement.

Keep the dynamic efficiency card ready for evaluation: static efficiency is a snapshot, and structures that fail it (monopoly, oligopoly) may deliver falling costs and innovation over time out of retained profits. One developed paragraph on this lifts almost any efficiency essay into the top band.

Cambridge International A Level Economics (9708)

Monopoly and Imperfect Competition

A2 Lecture Notes with Diagrams

*Revenue under imperfect competition | Monopoly: barriers, equilibrium, objectives and efficiency |
Price discrimination and its three degrees | Contestable markets | The principal-agent problem*

Prepared by: Abdur Rehman Shahzaib

1. Revenue When the Firm Faces a Downward-Sloping Demand Curve

In perfect competition the firm could sell any amount at the market price, so AR and MR were one horizontal line. The moment competition becomes imperfect, the firm faces a **downward-sloping demand curve**: to sell more units it must cut the price, and the cut applies to every unit sold, not just the extra one. That single fact reshapes all the revenue curves, and this table is the cleanest way to see it. Work through it column by column with the class:

Price (\$)	Quantity	TR (\$)	AR (\$)	MR (\$)	What is happening
10	1	10	10	10	price cuts raise TR:
9	2	18	9	8	demand ELASTIC here,
8	3	24	8	6	PED > 1, MR positive
7	4	28	7	4	
6	5	30	6	2	TR peaks: PED = 1,
5	6	30	5	0	MR = 0, TR maximum
4	7	28	4	-2	price cuts lower TR:
3	8	24	3	-4	demand INELASTIC,
2	9	18	2	-6	PED < 1, MR negative
1	10	10	1	-8	

Four rules fall straight out of the table, and they hold for every imperfectly competitive diagram in this topic. First, AR always equals price, so the demand curve IS the AR curve. Second, MR lies below AR at every output, because the price cut needed to sell one more unit drags down the revenue on all earlier units too. Third, for a straight-line demand curve MR falls exactly twice as fast and reaches zero at the midpoint of demand. Fourth, the demand curve splits into an elastic upper segment (MR positive), a unitary point (MR zero, TR maximum), and an inelastic lower segment (MR negative).

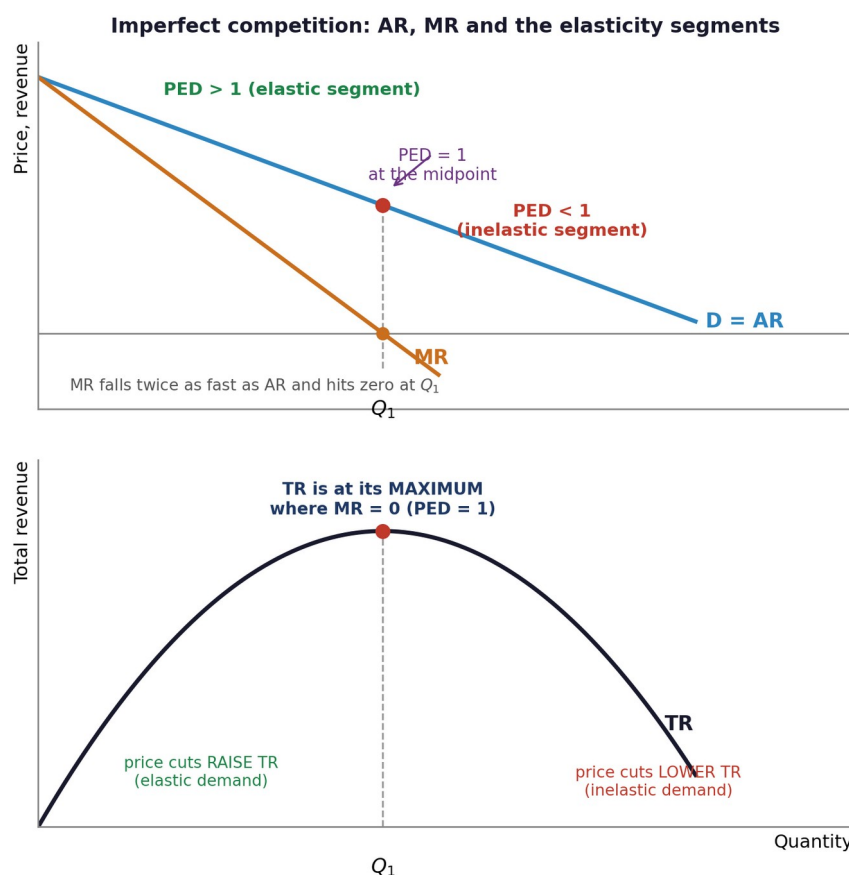


Figure 1: The AR and MR curves with the elasticity segments (top), and the total revenue hill beneath them (bottom). TR peaks exactly where $MR = 0$ and $PED = 1$.

An exam nuance worth planting early: a profit-maximising firm always ends up on the ELASTIC segment of its demand curve. Profit maximisation needs $MC = MR$, and since marginal cost is always positive, MR must be positive at that output, which only happens on the elastic upper half. A firm found operating on the inelastic segment could raise price, sell fewer units, and increase both revenue and profit at once.

2. Monopoly

Monopoly: a market structure with a single (or completely dominant) seller of a product with no close substitutes, protected by high barriers to entry. A pure monopoly is the industry's only firm; in practice a firm with a dominant share is treated as having monopoly power, and in the UK a 25% market share is enough for legal (working) monopoly status.

2.1 Characteristics of a monopoly

- **One dominant seller with the highest market share.** The firm IS the industry, so the downward-sloping industry demand curve is the firm's own AR curve.
- **Price maker.** Unlike the perfectly competitive price taker, the monopolist chooses its position on the demand curve: it sets either the price or the output, though never both independently.
- **No close substitutes, so demand is relatively inelastic.** Buyers have nowhere comparable to switch, so a price rise loses relatively few customers and tends to raise total revenue. (Keep the earlier nuance: at the chosen profit-maximising output, the firm still sits on the elastic segment.)

- **High barriers to entry and exit.** These are the source of the monopolist's power and the reason its profits survive into the long run.
- **Abnormal (supernormal) profit in the short run AND the long run.** No entrant can compete the profit away, which is the single biggest contrast with perfect competition.
- **Can price discriminate and may advertise.** Advertising here defends the brand and widens the moat rather than fighting head-to-head rivals.

2.2 Barriers to entry: natural versus artificial

Barriers divide into those that arise from the nature of the industry (natural) and those the firm deliberately builds (artificial or strategic). The distinction is exam-relevant because natural barriers can justify leaving a monopoly alone, while artificial ones invite regulation.

Natural barriers	Artificial (strategic) barriers
Economies of scale: a large incumbent produces at unit costs no small entrant can match; in the extreme (a natural monopoly such as a water network) one firm serving the whole market is the cheapest possible arrangement	Patents and copyright: legal sole rights over an invention, process or content for a period, blocking imitation
Control of key resources: owning a large share of a raw material's supply shuts rivals out of the input market	Predatory pricing: deliberate deep price cuts that a new entrant cannot survive, with prices restored once it leaves
Huge capital and investment requirements: power grids, railways and networks cost too much for entrants to duplicate	Brand loyalty and heavy advertising: decades of branding make demand for the incumbent's product inelastic and entry expensive
Location: a single shop or port may enjoy a local monopoly simply because of where it is	Mergers and takeovers: buying up actual or potential rivals; legal protection, where the state itself grants sole status to utilities and nationalised industries

Real-world examples

Utilities are the classic natural monopolies: one electricity distributor typically serves an entire city, K-Electric in Karachi is the example your students see on every bill, because duplicating the wires would be pure waste.

Patent-protected medicines show artificial barriers at work: for the patent's life the inventor is the sole legal seller, and prices fall sharply only when generics enter on expiry.

Resource control: OPEC's coordination of a large share of world oil supply gives its members collective price-making power no entrant can bypass.

2.3 Equilibrium: abnormal profit that lasts

The monopolist maximises profit exactly as any firm does, at the output where $MC = MR$, then charges the highest price the demand curve allows for that output. Because AR exceeds ATC at this output, the firm earns abnormal profit, the shaded rectangle in Figure 2. In perfect competition this profit would summon entrants; here the barriers hold them out, so the same diagram describes the short run AND the long run. That persistence is what makes monopoly profit an exam-favourite contrast with perfect competition.

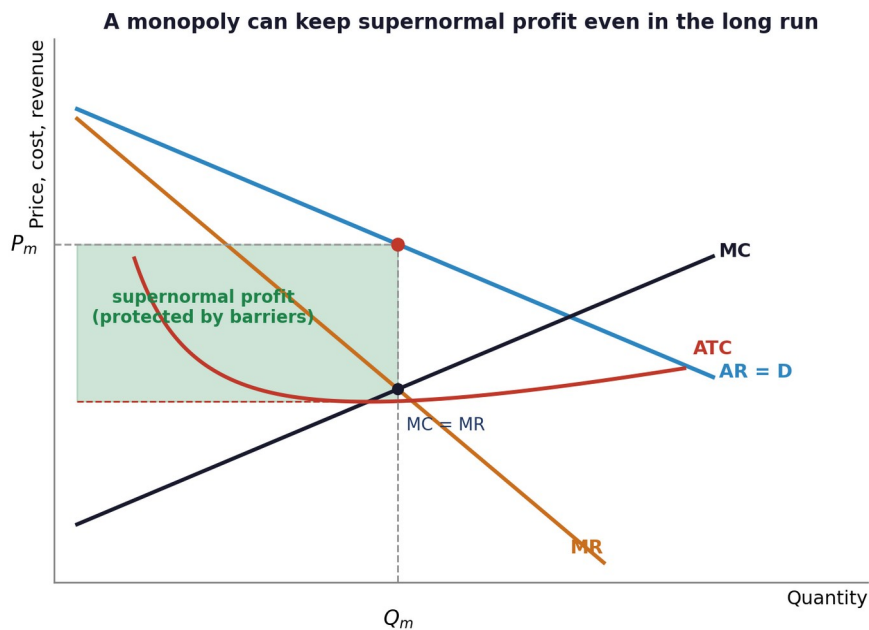


Figure 2: Monopoly equilibrium. Output where $MC = MR$, price read off the AR curve, and abnormal profit (shaded) protected by barriers to entry in the long run.

2.4 One diagram, four objectives

Profit maximisation is only one possible goal. Because ownership and control are often separated (Section 5), managers may pursue other objectives, and each objective picks a different point on the same diagram. Learning Figure 3 as one picture answers a whole family of exam questions:

- **Profit maximisation ($MC = MR$):** the smallest output and highest price of the four; maximises the gap between TR and TC.
- **Revenue maximisation ($MR = 0$):** managers judged on sales figures push output to where TR peaks, at the unitary-elastic midpoint of demand.
- **Sales (output) maximisation ($AR = ATC$):** the largest output consistent with survival, breaking even with normal profit; often chosen to build market share or deter entry.
- **Allocative efficiency ($MC = AR$):** the socially best output, where price equals marginal cost; a public-sector monopoly may be instructed to price here.

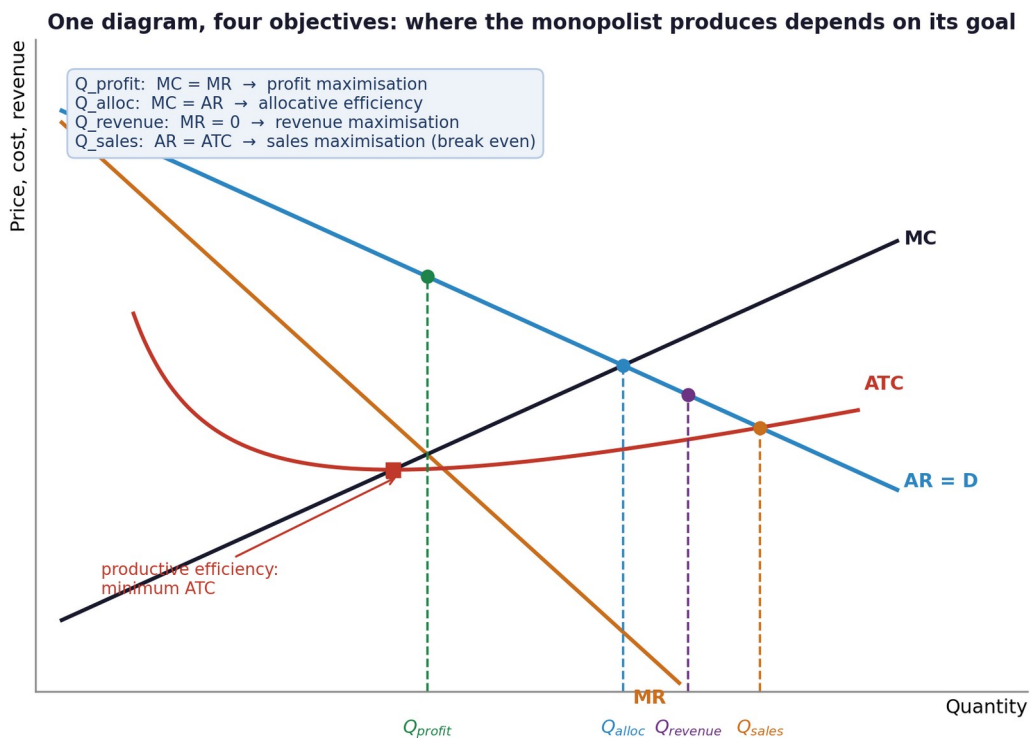


Figure 3: The same monopoly, four different chosen outputs. Profit maximisation gives the least output and highest price; sales maximisation the most output a break-even firm can sustain. The square on ATC marks productive efficiency at minimum average cost.

2.5 Is monopoly efficient?

Efficiency test	Verdict for a private profit-maximising monopoly
Allocative efficiency ($P = MC$)	FAILS. Price exceeds marginal cost at the chosen output, so too little is produced: under-allocation and a deadweight welfare loss. Consumer surplus is transferred to the producer.
Productive efficiency (minimum ATC)	Not guaranteed. The $MC = MR$ output rarely coincides with the bottom of ATC; it is possible, but only by coincidence or instruction.
X-efficiency (no organisational slack)	At risk. With no competitive pressure, costs may drift above their minimum, the quiet-life problem of Section 5.
Dynamic efficiency (innovation over time)	Monopoly's strongest defence. Retained abnormal profit finances R&D and new products; entry threats cannot strip the reward for innovating.

The balanced verdict, and the one examiners reward, is that monopoly fails the static tests but can beat competitive markets dynamically, and where economies of scale are large, a single producer's costs (and even its price) can be lower than a fragmented industry's would be. The case against monopoly is strongest when barriers are artificial and the market is not contestable; the case for it is strongest for natural monopolies and innovation-driven industries.

2.6 Advantages and disadvantages at a glance

Against monopoly	For monopoly
Higher price and lower output than competition; welfare loss	Economies of scale can mean lower costs, and possibly lower prices, than many small firms

Allocative inefficiency ($P > MC$) and possible productive inefficiency	Abnormal profit funds R&D and innovation: dynamic efficiency
X-inefficiency from the absence of competitive pressure	Stability and resources for long-term investment in networks and infrastructure
Consumer choice restricted; consumer surplus transferred to producer	Natural monopoly avoids wasteful duplication of infrastructure

3. Price Discrimination

Price discrimination: charging different prices to different consumers for the SAME product, for reasons not associated with differences in cost, based on differences in their price elasticity of demand.

Three conditions must hold before a firm can discriminate successfully. It must have **price-making power** (a price taker has no prices to vary). It must be able to **separate consumers into markets with different PEDs**, by age, time of use, income group or geography. And it must be able to **prevent resale** (seepage) between the markets, or buyers in the cheap market would simply resell to the dear one and collapse the price gap.

3.1 The three degrees

First-degree discrimination charges every individual consumer the maximum they are willing to pay, unit by unit. If it could be achieved perfectly, the entire consumer surplus would be converted into producer revenue, as Figure 4 shows. It is rare in pure form, think of bargaining for concert tickets or a bazaar negotiation, because discovering each buyer's maximum is costly.

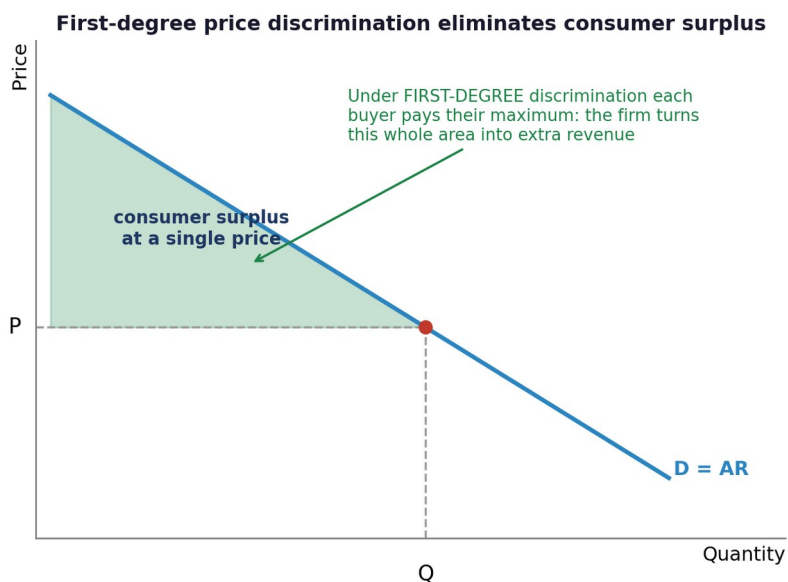


Figure 4: First-degree discrimination. Charging each buyer their maximum converts the whole consumer surplus triangle into producer revenue.

Second-degree discrimination charges different prices for successive blocks of consumption rather than for different people. Utility tariffs are the standard case: the first block of electricity units is priced high, later blocks cheaper, so heavy users pay a lower average price per unit than light users.

Third-degree discrimination is the form the exam usually wants: the same product sold to different GROUPS at different prices, peak versus off-peak travellers, adults versus students, domestic versus foreign buyers. The firm splits its output so that the marginal revenue in each market equals marginal cost, then charges what each market will bear: a high price where demand is inelastic, a low price where it is elastic. Figure 5 shows the logic.

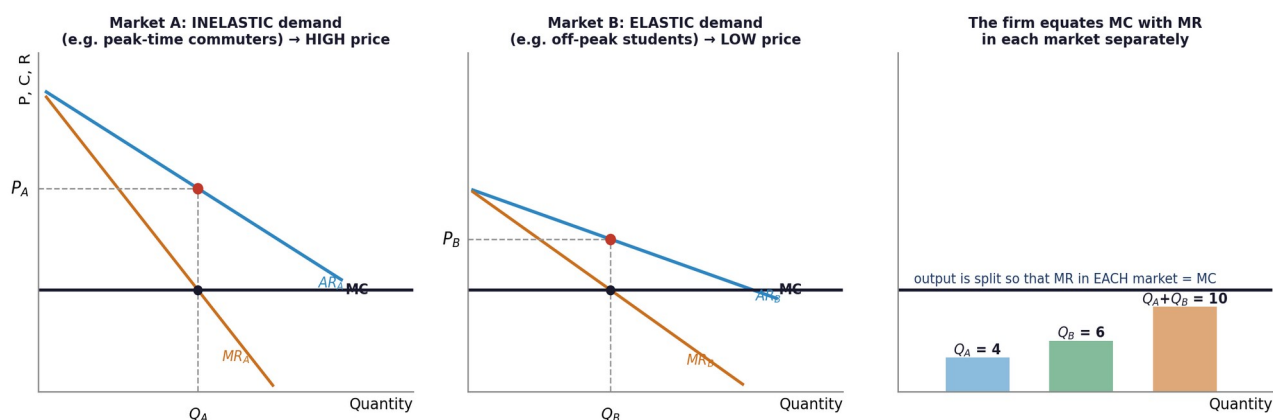


Figure 5: Third-degree discrimination. With MC common to both markets, the firm sets $MR = MC$ in each one separately: the inelastic market pays P_A , the elastic market the lower P_B .

3.2 Who wins and who loses?

No definitive judgement is possible on whether price discrimination serves the public interest, which is exactly what makes it good essay material. The gains and losses split as follows:

Costs of price discrimination	Benefits of price discrimination
Producers capture consumer surplus, a welfare loss for individuals	Producers earn higher revenue and profit, which can finance R&D and better products
Consumers in the inelastic market now pay a HIGHER price than before	Consumers in the elastic market pay a LOWER price than before
Can entrench monopoly power by funding predatory behaviour	Low-income groups may afford a product previously out of reach (student and off-peak fares)
Feels unfair: identical product, different prices	Higher charges on some users can fund provision that benefits all, e.g. high private fees enabling a hospital to buy equipment that also serves non-paying patients

Real-world examples

Transport is the textbook case: railways and airlines charge peak commuters (inelastic, must travel) far more than off-peak leisure travellers (elastic, flexible), and student and senior concessions are third-degree discrimination by age group.

Electricity companies' block tariffs are second-degree discrimination your students' families experience monthly.

Cinemas charging students less on weekday afternoons, and medicines priced differently in domestic and foreign markets, both show group-based pricing with resale kept difficult.

4. Contestable Markets

Contestable market: a market in which the existing firm earns only normal profit, because it cannot set price above average cost without attracting immediate entry, owing to the absence of barriers to entry and of sunk costs. The threat of competition, not competition itself, does the disciplining. (Theory developed by William Baumol.)

A market is contestable when new firms face no barriers to entry or exit, incur no sunk costs (costs that cannot be recovered on leaving, such as specialised equipment or past advertising), suffer no competitive disadvantage against the incumbent, can use the same technology, and can enter and leave rapidly. Under those conditions any abnormal profit invites **hit-and-run entry**: a rival enters, takes a share of the profit, and exits costlessly the moment the incumbent responds.

The implication overturns the usual structure-based analysis: what matters is not how many firms a market HAS but how easily it COULD have more. Even a monopoly behaves like a competitive firm if its market is perfectly contestable. To deter entry the incumbent must abandon the profit-maximising price P_1 in Figure 6 and price at average cost, P_2 , earning only normal profit; the safest entry-proof position of all is P_3 at the minimum of ATC, which also forces the firm to eliminate X-inefficiency, since bloated costs would leave room for a leaner entrant.

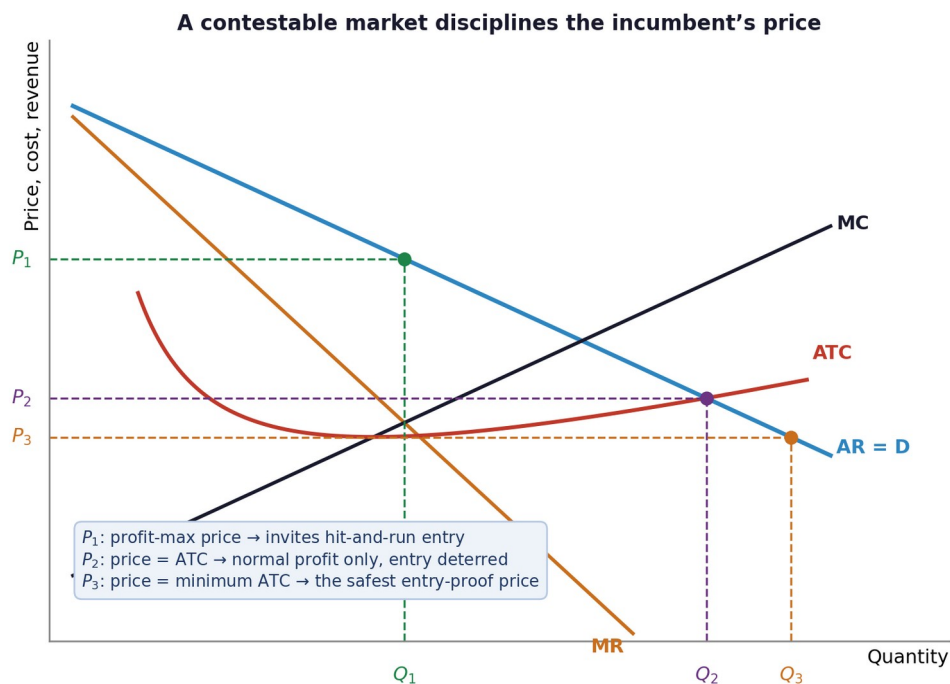


Figure 6: Contestability disciplines price. At P_1 the incumbent invites hit-and-run entry; pricing at average cost (P_2), or at minimum ATC (P_3), removes the profit that entry would chase.

Evaluation. Perfect contestability is as rare as perfect competition: almost every real market involves some sunk cost, advertising can never be recovered, and incumbents rarely share their technology or brand advantages. The useful idea is the SPECTRUM: the more contestable a market, the closer even a dominant firm's behaviour moves towards the competitive outcome. This makes contestability a first-rank evaluation point in any monopoly essay: the welfare verdict on a monopoly depends less on its market share than on how open its market is to attack.

Real-world examples

Budget airlines made European routes contestable: aircraft can be leased and switched between routes, so an incumbent charging fat margins on a route quickly finds a low-cost carrier flying it.

Ride-hailing apps made city taxi markets contestable almost overnight, and incumbent fares and service quality responded exactly as the theory predicts.

5. The Principal-Agent Problem

Principal-agent problem: a problem arising from conflict between the objectives of principals and those of the agents who take decisions on their behalf. In a public limited company the shareholders (principals) delegate day-to-day control to managers (agents).

The problem is born from the **divorce of ownership from control**. Shareholders own the firm and typically want maximum profit; the managers who actually run it may want salary, status, empire-building, or simply a quiet life, doing just enough profit to keep shareholders satisfied rather than the most profit possible (profit *satisficing*). Accountability is weakest when shareholders are a large, scattered group of small investors, none of whom finds it worthwhile to monitor closely.

The root cause is **asymmetric information**: the agents know far more about the business's day-to-day reality than the owners do, so weak decisions are hard to detect. Two consequences matter for this topic. Managers may be negligent because they are not fully accountable, producing **organisational slack**, costs above their minimum, which is the X-inefficiency met in Section 2.5. And managers may steer the firm to objectives other than profit, which is precisely why the revenue-maximising and sales-maximising outputs of Figure 3 exist: those are the outputs agents choose when sales figures, growth and market share, not profit, determine their pay and prestige.

Reducing the problem means realigning the agents' incentives with the principals' objectives: closer monitoring of managers' actions (independent directors, audits, shareholder meetings), and incentive pay that ties managers' rewards to the owners' goal, profit-related bonuses and share option schemes that make the managers part-owners themselves. Each remedy has a cost, monitoring is expensive and bonus schemes can be gamed, so the problem is managed rather than eliminated.

Real-world examples

Share options are the standard alignment tool: a manager holding options gains personally when the share price rises, so decisions tilt back towards shareholder value.

Shareholder revolts at AGMs over executive pay are the principal-agent conflict playing out in public: owners using their votes to discipline agents whose rewards outran performance.

Exam tips for this topic

Learn Figure 3 as one picture. Questions on monopoly objectives, on revenue versus profit maximisation, and on why a firm might break even on purpose are all answered by marking the right point on that single diagram and explaining the condition ($MC = MR$, $MR = 0$, $AR = ATC$, $MC = AR$).

For any price discrimination essay: definition, the three conditions, the third-degree diagram with the inelastic market paying more, then the two-sided welfare table, and a conclusion that refuses a blanket verdict, who gains depends on which market a consumer sits in and on what the extra profit funds.

Contestability is your strongest evaluation card in monopoly essays: the harm a monopoly does depends

on barriers and sunk costs, not on market share alone. One developed paragraph with Figure 6's logic lifts the AO3 band.

Connect the sections: the principal-agent problem explains WHY firms pursue the alternative objectives of Figure 3 and WHERE X-inefficiency comes from; contestability explains WHEN a monopoly cannot afford either. Examiners reward answers that link these ideas rather than reciting them separately.

Cambridge International A Level Economics (9708)

A2 Lecture Notes: Monopolistic Competition and Oligopoly

Imperfect Competition | Teacher: Abdur Rehman Shahzaib

Perfect competition and pure monopoly are the two ends of the market structure spectrum, and almost no real market sits at either end. The two structures in this lecture occupy the middle ground, together they are called imperfect competition. Monopolistic competition sits close to the competitive end: many small firms, but each with a product of its own. Oligopoly sits close to the monopoly end: a handful of large firms whose decisions depend on each other. Keep one question in mind throughout: how much market power does the firm have, and what stops it from using that power?

Part 1: Monopolistic Competition

Definition

Monopolistic competition is a market structure in which a large number of medium-sized firms compete by selling differentiated products. It is the market structure closest to the model of perfect competition, because of the large number of competing firms and buyers, but each firm's product is slightly different, so each firm faces its own downward sloping demand curve and has some influence over its price. Firms are price makers, within limits.

1.1 Characteristics

- **A large number of buyers and sellers.** There are several, usually medium-sized, firms, so each has only a small share of the market and a limited influence on it.
- **Product differentiation.** The output of each firm has common features that keep the firms in competition with one another, and some unique features that distinguish it from rivals. The differences can be physical (recipe, design, quality) or qualitative (branding, packaging, customer service, a USP, the atmosphere or “snob appeal” of a store).
- **Weak barriers to entry and exit.** New firms can enter fairly easily, each entering with its own differentiated version of the product, and it is easy for firms to recoup their capital expenditure on exit from the market, little is sunk.
- **Some price-making power.** Because its product is differentiated, each firm can raise its price without losing all its customers. But because close substitutes exist, demand for each firm's product is relatively ELASTIC, the demand curve is downward sloping but flatter than a monopolist's.
- **Price and non-price competition.** Firms compete through price, and heavily through advertising, promotion, branding and service.

Real examples

Hairdressers and salons: every salon offers a haircut, but each differs in location, stylists, atmosphere and price.

Restaurants and takeaways: hundreds compete in any large city, each with its own menu, taste and brand.

Clothing retail: countless labels sell shirts and jeans, differentiated by design, branding and image.

1.2 Short run equilibrium: abnormal profit

Like every profit-maximising firm, the firm produces where $MC = MR$ (point e in Figure 1), sells output Q at price P on its demand curve (point a), and has average total cost C (point b). Because of the differentiated product, the firm can hold price above average cost in the short run: AR exceeds ATC, and the shaded rectangle PabC is **abnormal profit**.

Notice how the demand curve compares with a monopolist's: it is flatter, because many close substitutes exist, demand for any one firm's product is relatively elastic. The more substitutes appear, the flatter and more elastic this curve becomes, which is exactly what happens next.

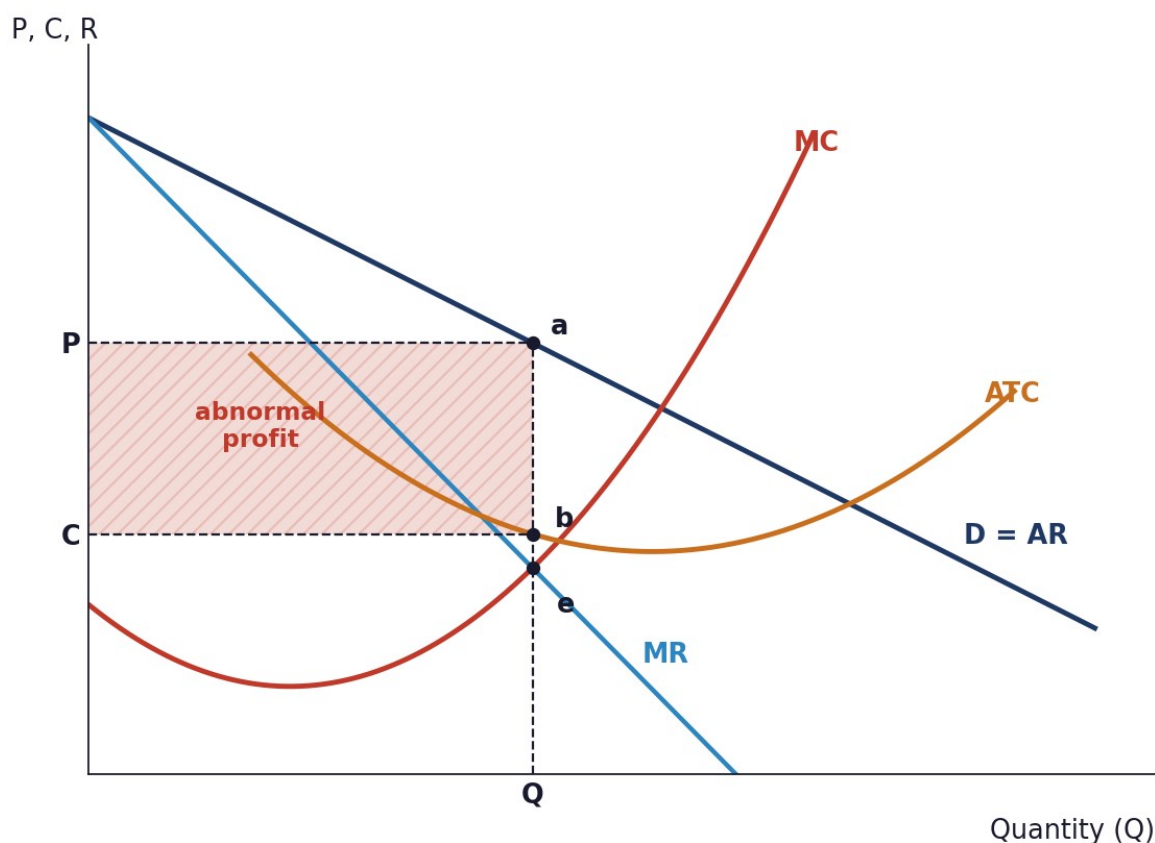


Figure 1: Monopolistic competition in the short run, abnormal profit ($P > ATC$ at output Q)

1.3 Long run equilibrium: entry competes profit away

Abnormal profits attract new firms, and the weak barriers to entry let them in. Entry has two effects on each existing firm: **demand falls** (the demand curve shifts left as the new substitutes take away customers) and **PED rises** (demand becomes more elastic, because buyers now have more alternatives). Entry continues until only **normal profit** is earned, that is, until $AR = ATC$.

In Figure 2 the demand curve has shifted in until it is exactly TANGENT to the ATC curve at point a, directly above the $MC = MR$ output Q. Price equals average cost, so the firm just covers all its costs, including the reward for staying in the industry, and there is no further incentive for entry or exit.

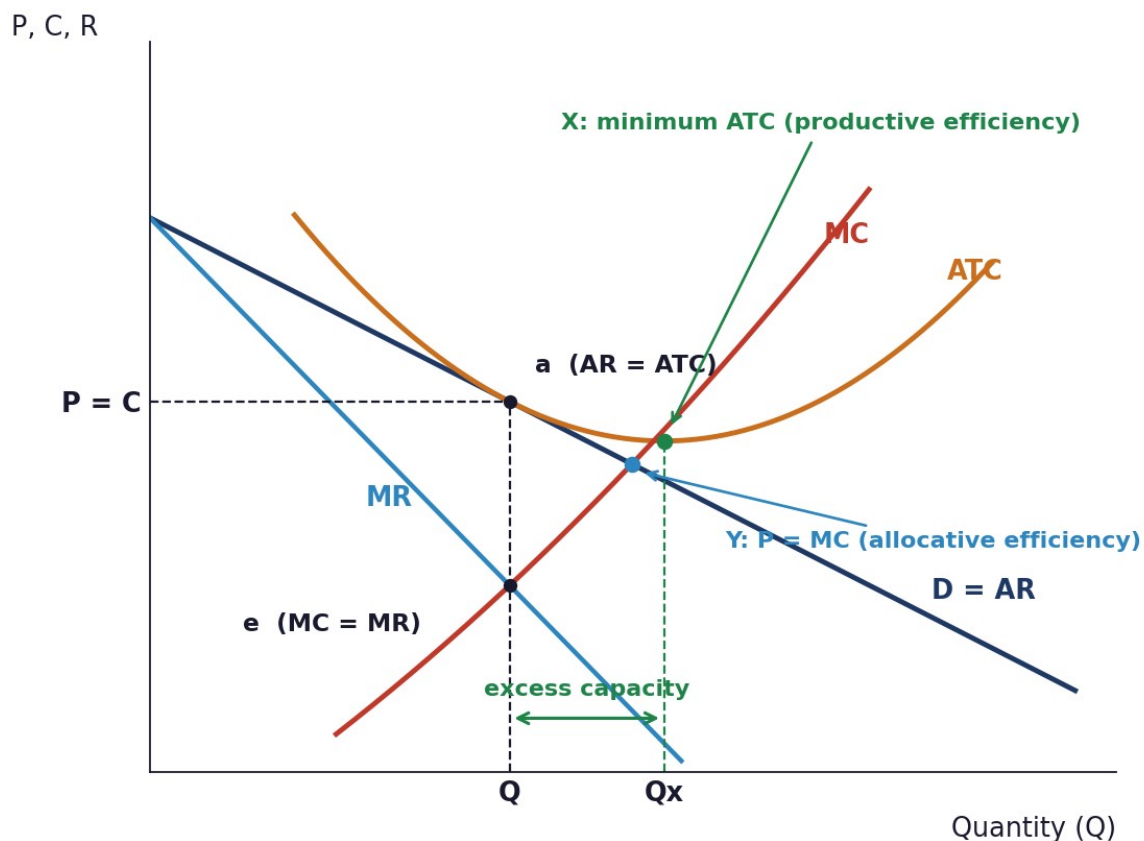


Figure 2: Monopolistic competition in the long run, normal profit (AR = ATC at the tangency point a)

1.4 Efficiency in the long run

- **Productive efficiency is NOT achieved.** Productive efficiency requires production at minimum ATC (point X at output Q_x). The firm produces at Q, to the left of Q_x , so average cost is higher than it needs to be.
- **Allocative efficiency is NOT achieved.** Allocative efficiency requires $P = MC$ (point Y). At output Q, price exceeds marginal cost, society values an extra unit at more than it costs to make, so too little is produced.
- **Excess capacity.** The gap between Q and Q_x is the problem of excess capacity in the long run: each firm operates below its efficient scale. The market has many firms, each producing a little at a cost higher than the minimum, think of a street with ten half-empty restaurants.

The firm is therefore economically inefficient in the long run, even though it earns only normal profit. The usual evaluation point: the inefficiency is the price society pays for VARIETY. Consumers get choice between differentiated products, and the price of that choice is slightly higher average cost.

1.5 Where it sits between the extremes

	Perfect competition	Monopolistic competition	Monopoly
Short run profit	normal or abnormal	abnormal possible	abnormal
Long run profit	normal	normal	abnormal
Barriers to entry	none	weak	high
Product	homogeneous	differentiated	unique

Long run efficiency	productive and allocative	neither (excess capacity)	neither
---------------------	---------------------------	---------------------------	---------

The economic assumption throughout is the same: every firm maximises profit where **MC = MR**. What changes between structures is the shape and position of the revenue curves, and what entry is allowed to do to them.

Part 2: Oligopoly (Imperfect Competition)

Definition

Oligopoly is a market situation in which total output is concentrated in the hands of a few firms. The market is dominated by a small number of large firms, typically three, four or five, each controlling a significant market share and significant capital. It is possibly the most realistic model of market structure: most large real-world industries are oligopolies.

2.1 The concentration ratio (CR)

An oligopoly is identified by its **concentration ratio**: CR_n measures the percentage share of the market held by the largest n firms. A high CR means the market is dominated by a few sellers.

Worked example

An industry contains 150 firms, which together supply 100% of the market. The largest 4 firms supply 60% of total sales, while the remaining 146 firms share the other 40%.

$CR_4 = 60\%$. Four firms control well over half the market, so despite the 150 names on paper, this industry behaves as an oligopoly.

2.2 Characteristics

- **High or strong barriers to entry.** Large capital is required, set-up costs are high, and the incumbents' economies of scale push their ATC down to levels a small entrant cannot match.
- **Identical OR differentiated products.** Oligopolists may sell near-identical products (sugar, oil, steel) or differentiated ones (cars, banking services). Both count, what defines oligopoly is the small number of large firms, not the product.
- **Interdependence.** This is THE defining feature. With only a few large rivals, decisions taken by firms are interdependent: before changing its price, each firm must anticipate the reaction of its rivals and build that reaction into its own decision making. Firms are rivalrous and strategic, in both price and non-price competition.
- **Imperfect information between firms.** Each firm knows a great deal about the market but cannot be certain what its rivals will do next, which is exactly why strategy matters.

2.3 Price rigidity

The uncertainty and risks associated with price competition may lead to **price rigidity**: prices in oligopoly do not change often, even when costs change. Firms avoid price wars, which can be mutually destructive,

and compete instead through **non-price competition**: advertising, branding, product features, loyalty schemes and service.

Real examples

Automobiles: a handful of global groups dominate; they rarely fight on list price, competing instead on features, financing and brand.

Airlines: a few carriers dominate most routes; fare structures move together and competition shifts to loyalty programmes and service.

Banking: the large banks watch each other's rates closely and change them infrequently, competition runs through branding and products.

Soft drinks: Coca-Cola and Pepsi have competed for a century mainly through advertising, not price.

2.4 Two faces of oligopoly: compete or cooperate

Everything in oligopoly follows from one choice: the few firms can COMPETE with each other (non-collusive oligopoly) or COOPERATE (collusive oligopoly, an anti-competitive action). The map below organises the whole topic.

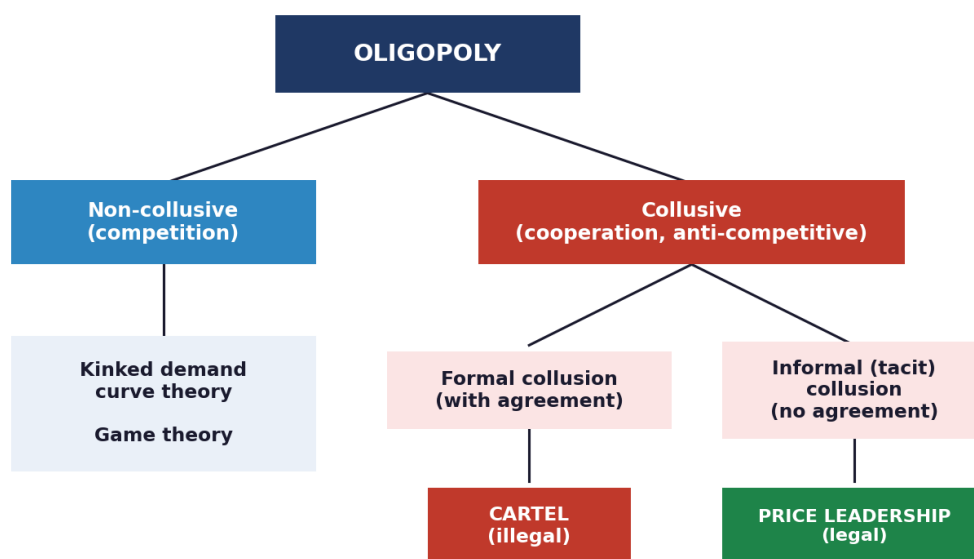


Figure 3: Classification of oligopoly behaviour

2.5 Non-collusive oligopoly 1: the kinked demand curve

The kinked demand curve model explains price rigidity. Each firm believes that if it **lowers** its price, competitors WILL follow (they must, or lose customers), but if it **raises** its price, competitors will NOT follow (they happily collect its customers).

Work through it with Pepsi and Coke. Pepsi considers a price RISE, expecting total revenue to rise. In reality Coke will not follow, Pepsi's customers switch to Coke, Pepsi loses, so the price rise is rejected. Pepsi then considers a price CUT, expecting total revenue to rise as customers switch from Coke. In reality Coke will follow the cut, both end up selling about the same quantities at lower prices, both lose, so the price cut is rejected too. Conclusion: **price tends to be rigid** at the current level.

The demand curve the firm faces is therefore **KINKED** at the current price: elastic above the kink (raise price and you lose many customers to rivals who hold their prices) and inelastic below it (cut price and you win few customers, because rivals match the cut). The kink creates a vertical **GAP** in the MR curve at the current output. Marginal cost can rise or fall anywhere within that gap, from MC_1 to MC_2 , and the profit-maximising price and output do not change at all, a second, powerful reason why oligopoly prices are rigid.

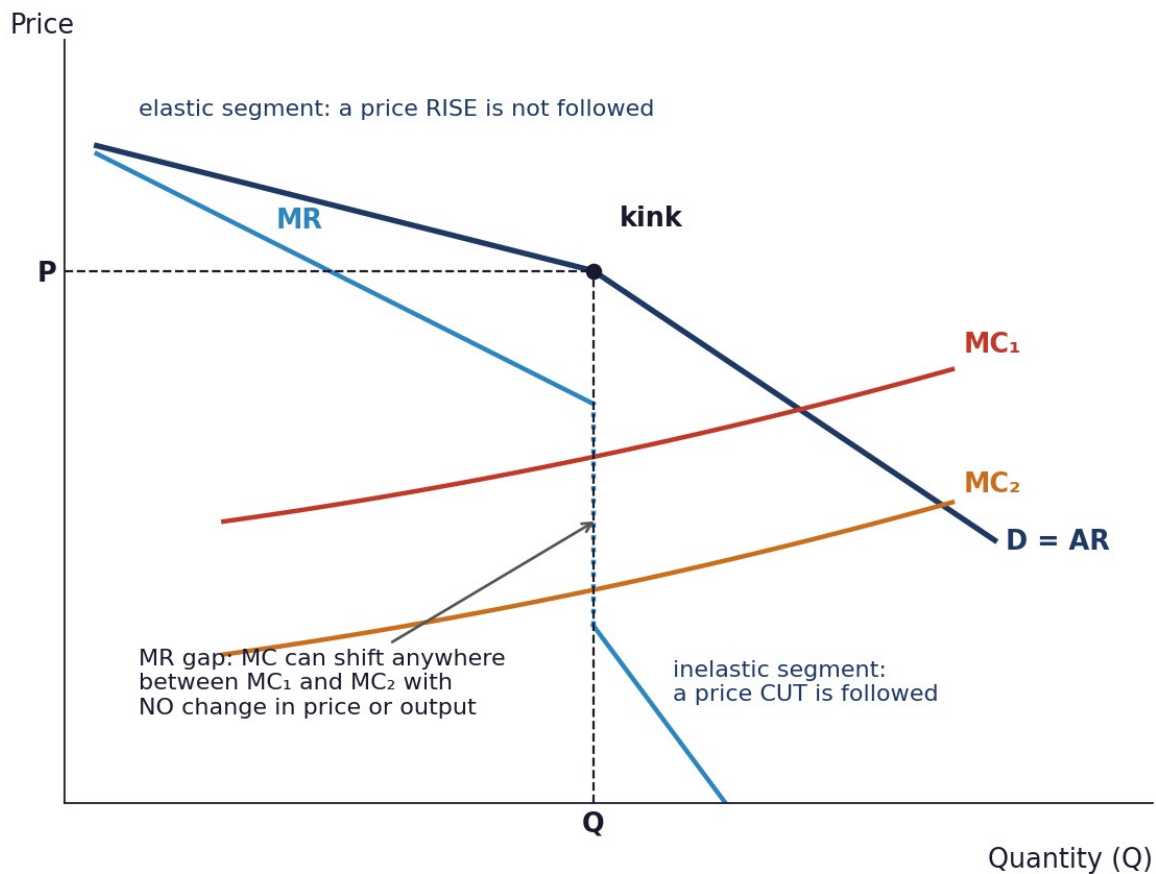


Figure 4: The kinked demand curve, elastic above the kink, inelastic below, with the MR gap at output Q

2.6 Non-collusive oligopoly 2: game theory

Game theory is a crucial tool for understanding non-collusive oligopolies, where firms must strategise based on the anticipated actions of their competitors. Its key idea is the **Nash equilibrium**: a situation in which each firm is doing the best it can **GIVEN** what its rivals are doing, so no firm has an incentive to change its strategy on its own.

The classic illustration is the **prisoner's dilemma**. Two rival firms, A and B, each choose a **HIGH** price or a **LOW** price without knowing the other's choice. The table shows profit (in \$ millions) as (A's profit, B's profit):

	B: high price	B: low price
A: high price	(10, 10)	(2, 14)
A: low price	(14, 2)	(5, 5)

Look at firm A's reasoning. If B prices high, A earns 14 by cutting price instead of 10. If B prices low, A earns 5 by cutting instead of 2. Cutting price is A's best reply **EITHER WAY**, a dominant strategy, and B reasons identically. Both cut, and the market lands at (5, 5), the Nash equilibrium, even though both would be better off at (10, 10).

This one table explains the central tension of oligopoly: firms have a strong joint incentive to COOPERATE on high prices, and each firm has a strong private incentive to CHEAT on any such agreement. That is why collusion keeps appearing, and why it keeps breaking down.

2.7 Collusive oligopoly 1: cartels (formal collusion)

In a **cartel**, firms form explicit agreements to fix prices or output in order to maximise their JOINT profits, the group behaves like a single monopolist, with each member given a production quota. This is formal collusion, collusion WITH an agreement, and it is **illegal** in almost every country because it raises prices against consumers.

Real examples

OPEC: the best-known legal exception, a cartel of oil-producing GOVERNMENTS that sets production quotas to influence the world oil price. It is beyond national competition law.

Enforcement cases: competition authorities such as the UK CMA and the European Commission have fined cartels in trucks, air cargo and construction hundreds of millions for price fixing.

Cartels are unstable for exactly the prisoner's dilemma reason: once the price is fixed high, every member can profit by secretly producing beyond its quota, and policing the agreement is hard. Cartels work best when firms are few, products are homogeneous, and cheating is easy to detect.

2.8 Collusive oligopoly 2: price leadership (informal, tacit collusion)

Tacit collusion occurs when firms in an oligopoly indirectly coordinate their actions WITHOUT any explicit agreement. The most common form is **price leadership**: one firm, usually the dominant or most established one, changes its price first, and the others follow within days. Nothing is written down and nobody meets, so it is **legal**, though competition authorities watch it closely.

Real examples

Retail banking: when a leading bank moves its mortgage or savings rates after a central bank decision, rivals typically match within the week.

Petrol retailing: pump prices in an area move almost together, each station following the local leader rather than agreeing anything explicitly.

Exam tips

Interdependence is the word examiners look for in any oligopoly definition, and the kinked demand curve is the model that turns it into a diagram. Always draw the MR gap and say that MC can shift within it without changing price or output.

Keep the classification sharp: non-collusive (kinked demand, game theory) versus collusive, and within collusive, formal (cartel, illegal) versus informal (price leadership, legal). The 8-mark favourite is to explain the difference with one example of each.

For monopolistic competition, the two must-have long run points are the TANGENCY ($AR = ATC$, normal profit) and EXCESS CAPACITY (output below minimum ATC). For evaluation, weigh the inefficiency against the benefit of product variety.