

Consumer and producer theory: costs, revenue and profit (AS & A Level notes)

These notes cover consumer behaviour (total and marginal utility, the law of diminishing marginal utility, consumer equilibrium, and the income and substitution effects of a price change) and producer theory (short-run and long-run production, the law of diminishing returns, the costs of production, economies and diseconomies of scale, revenue, and profit maximisation) for Cambridge International AS & A Level Economics, syllabus 9708. This topic sits within the price system and the microeconomy strand of the AS & A Level component. All figures, schedules and examples below are invented for teaching purposes; every calculation is shown in full so you can follow — and reproduce — the method.

Syllabus mapping note: check the exact subsection number(s) for each part of this topic against the current 9708 syllabus grid for your examination series before use, since precise AS/A Level placement can shift between syllabus versions.

1. Consumer behaviour: utility and demand

1.1 Total and marginal utility

Economists model consumer choice using the idea of **utility** — the satisfaction or benefit a consumer gets from consuming a good or service. Early utility theory treats utility as measurable in imaginary units called **utils**, which lets us build simple numerical schedules even though real satisfaction cannot literally be counted.

Utility (TU): the total satisfaction a consumer gains from consuming a given quantity of a good over a period of time.

Marginal utility (MU): the extra (additional) utility gained from consuming one more unit of a good, with consumption of all other goods held constant.

$MU = \Delta TU \div \Delta Q$, where ΔTU is the change in total utility and ΔQ is the change in quantity consumed (usually one extra unit).

Worked example 1 — a utility schedule. A student buys cups of tea. The table shows total utility (TU) and marginal utility (MU, calculated as the change in TU for each extra cup) as consumption rises from 1 to 7 cups.

Cups of tea (Q)	Total utility (TU, utils)	Marginal utility (MU, utils)
1	20	20 (20 – 0)
2	36	16 (36 – 20)
3	48	12 (48 – 36)
4	56	8 (56 – 48)
5	60	4 (60 – 56)



6	60	0 (60 – 60)
7	56	–4 (56 – 60)

Each extra cup adds less to total utility than the one before — MU falls steadily from 20 to 4 utils over the first five cups. TU reaches its maximum of 60 utils at the 5th cup and remains there through the 6th cup, since the 6th cup's MU is zero (it adds nothing further). From the 7th cup onward, MU is negative: an extra cup now makes the student feel worse (perhaps too much caffeine), so TU falls. If this good were free, a satisfied consumer would keep consuming only up to the point where MU reaches zero (maximising total utility); once a price must be paid, however, the rational stopping point comes earlier, where MU (in money terms) equals price — see Section 1.3.

1.2 The law of diminishing marginal utility

Law of diminishing marginal utility: as a consumer consumes additional units of a good within a given period, the marginal utility gained from each successive unit eventually falls, other things being equal.

This is a behavioural assumption about most goods, not a mathematical certainty — but it holds for almost all everyday goods (food, drink, clothing) once a basic level of consumption is reached.

1.3 Why diminishing marginal utility explains a downward-sloping demand curve

A rational, utility-maximising consumer will only buy an additional unit of a good if the utility gained from it is worth at least as much as the price they must pay. As MU falls with each extra unit consumed (Worked example 1), the maximum price the consumer is willing to pay for successive units also falls. Put another way: to persuade the consumer to buy more of a good, the price must fall, because the extra utility from additional units is worth less to them. Since the marginal utility schedule (converted into money terms) traces out the maximum price a consumer will pay at each quantity, it is effectively the consumer's individual demand curve for that good, provided the marginal utility of money itself is assumed constant — a simplifying assumption in this basic (cardinalist) model. Because MU diminishes, that curve slopes downward from left to right: quantity demanded rises only as price falls.

1.4 Consumer equilibrium (outline)

Consumer equilibrium (utility-maximising rule): a consumer with a fixed money income allocates spending between goods so that the marginal utility gained from the last dollar spent on each good is equal, and total income is used up:

$$MU_x \div P_x = MU_y \div P_y = \text{MU of the last dollar spent on any other good}$$

This is the **equi-marginal principle**. If $MU_x \div P_x > MU_y \div P_y$, the consumer gains more utility per dollar from X than from Y, so total utility rises if they switch spending from Y to X — and keep switching until the ratios are equalised (or as close to equal as whole units allow).

Worked example 2 — finding the equilibrium bundle. Good X costs $P_x = \$4$ and good Y costs $P_y = \$2$. A consumer has a fixed weekly income of \$20 to spend on the two goods. The marginal utility schedules are:

Unit	MU_x (utils)	$MU_x \div P_x$	MU_y (utils)	$MU_y \div P_y$
1st	40	10	22	11
2nd	32	8	18	9
3rd	24	6	14	7
4th	16	4	10	5
5th	8	2	6	3

Step 1 — rank purchases by MU per dollar, highest first: Y1 (11), X1 (10), Y2 (9), X2 (8), Y3 (7), X3 (6), Y4 (5), X4 (4), ...

Step 2 — buy in that order until the \$20 budget is used up: Y1 (\$2, cumulative \$2) → X1 (\$4, cum. \$6) → Y2 (\$2, cum. \$8) → X2 (\$4, cum. \$12) → Y3 (\$2, cum. \$14) → X3 (\$4, cum. \$18) → Y4 (\$2, cum. \$20 — budget exactly exhausted).

Equilibrium bundle: X = 3 units, Y = 4 units. Check: $(4 \times 3) + (2 \times 4) = 12 + 8 = \20 . ✓ At this bundle $MU_x \div P_x = 24 \div 4 = 6$ and $MU_y \div P_y = 10 \div 2 = 5$ — the closest the two ratios can get given that units must be bought whole; the next unit of X (X4, ratio 4) cannot be bought because the \$20 budget has already been fully spent (cumulative spend after Y4 = \$20). Note that X4's ratio (4) is actually higher than the ratio of the next unbought unit of Y (Y5, ratio 3) — so if any budget remained, X4 would be bought before Y5.

Step 3 — confirm this bundle maximises total utility. Total utility at X = 3, Y = 4: $TU_x = 40 + 32 + 24 = 96$; $TU_y = 22 + 18 + 14 + 10 = 64$; total = **160 utils**. Compare with an alternative bundle costing the same \$20, say X = 4, Y = 2 (spend = $16 + 4 = \$20$): $TU_x = 40 + 32 + 24 + 16 = 112$; $TU_y = 22 + 18 = 40$; total = 152 utils — lower. This confirms the equi-marginal bundle (160 utils) gives more total satisfaction than the alternative for the same spending.

Exam note: CIE 9708 questions on consumer equilibrium typically ask you to state and apply the equi-marginal condition ($MU_x \div P_x = MU_y \div P_y$) to a given bundle, rather than requiring a full ranked-purchase allocation from scratch. The step-by-step ranking above is included to build understanding of why the rule holds, not because you need to reproduce every step in an exam answer.

1.5 Income and substitution effects of a price change

A change in the price of a good has two separate effects on quantity demanded, together known as the **price effect**:

Substitution effect: the change in quantity demanded caused purely by the change in the good's price relative to other goods, with real income held constant. A fall in price makes the good relatively cheaper than substitutes, so consumers switch towards it.

Income effect: the change in quantity demanded caused by the change in the consumer's real purchasing power (real income) that results from the price change, with relative prices held constant. A fall in price raises real income (the same money income now buys more), which changes quantity demanded depending on whether the good is normal or inferior.

Type of good	Substitution effect of a price fall	Income effect of a price fall	Overall effect on quantity demanded
Normal good	Positive (quantity rises)	Positive (quantity rises)	Rises — both effects reinforce each other
Inferior good	Positive (quantity rises)	Negative (quantity falls, but by less than the substitution effect)	Rises — substitution effect dominates
Giffen good (special case)	Positive (quantity rises)	Negative, and larger than the substitution effect	Falls — an exception to the law of demand

A Giffen good is a strongly inferior good that takes up a very large share of a poor consumer's budget (the historic textbook example is a staple food). A price fall frees up so much real income that the consumer buys less of the staple and switches to more varied, preferred food — so quantity demanded falls as price falls, giving an upward-sloping demand curve. This is a rare, theoretical case, not typical behaviour.

2. Producer theory: production in the short run and long run

Short run: the time period in which at least one factor of production is fixed in supply (commonly capital/plant size) while other factors (such as labour) can be varied.

Long run: the time period in which all factors of production are variable, so the firm can change the scale of its entire operation; there is no fixed factor.

2.1 The law of diminishing (marginal) returns

Law of diminishing marginal returns: as successive equal units of a variable factor are added to a fixed factor of production in the short run, a point is eventually reached beyond which the marginal product of the variable factor declines.

Worked example 3 — a production schedule. A workshop has a fixed amount of capital (machinery). Workers (L) are added one at a time. Total product (TP) is output per week; marginal product (MP) is the extra output from the last worker added; average product (AP) is $TP \div L$.

Workers (L)	Total product (TP)	Marginal product (MP)	Average product (AP)
1	10	10	10.00
2	22	12	11.00
3	36	14	12.00
4	52	16	13.00
5	65	13	13.00
6	72	7	12.00
7	72	0	10.29
8	66	-6	8.25

Sample calculations: MP for the 4th worker = $TP(4) - TP(3) = 52 - 36 = 16$. AP for the 4th worker = $TP(4) \div 4 = 52 \div 4 = 13.00$. MP for the 8th worker = $TP(8) - TP(7) = 66 - 72 = -6$ (negative, because the 8th worker gets in the way of the others and total output falls).

Three stages are visible: (i) **increasing returns**, workers 1–4, where MP rises with each extra worker ($10 \rightarrow 16$) as the fixed capital is used more effectively and workers can specialise; (ii) **diminishing marginal returns**, from the 5th worker onward, where MP falls each time ($16 \rightarrow 13 \rightarrow 7 \rightarrow 0$) because each new worker has less capital and floor space to work with; (iii) **negative returns**, from the 8th worker, where MP turns negative and TP itself falls. Note that MP crosses AP exactly at AP's maximum (both equal 13.00 at the 5th worker) — a general rule: when $MP > AP$, AP is rising; when $MP < AP$, AP is falling.

3. Costs of production

Fixed cost (FC): a cost that does not change with the level of output in the short run (e.g. rent, insurance) — it must be paid even if output is zero.

Variable cost (VC): a cost that changes directly with the level of output (e.g. raw materials, piece-rate wages).

Total cost (TC): $TC = FC + VC$.

Average fixed cost (AFC): $AFC = FC \div Q$. **Average variable cost (AVC):** $AVC = VC \div Q$.

Average (total) cost (AC or ATC): $AC = TC \div Q = AFC + AVC$.

Marginal cost (MC): the extra cost of producing one more unit of output. $MC = \Delta TC \div \Delta Q$ (equivalently $\Delta VC \div \Delta Q$, since fixed costs do not change).

Worked example 4 — a full short-run cost schedule. A firm has fixed costs of \$100. Its variable costs at each output level are given below; all other columns are calculated from them.

Q	FC (\$)	VC (\$)	TC (\$)	AFC (\$)
1	100	60	160	100.00
2	100	110	210	50.00
3	100	150	250	33.33
4	100	180	280	25.00
5	100	220	320	20.00
6	100	270	370	16.67
7	100	330	430	14.29
8	100	400	500	12.50

Q	AVC (\$)	AC (\$)	MC (\$)
1	60.00	160.00	60
2	55.00	105.00	50

3	50.00	83.33	40
4	45.00	70.00	30
5	44.00	64.00	40
6	45.00	61.67	50
7	47.14	61.43	60
8	50.00	62.50	70

Working for $Q = 4$, shown in full: $TC = FC + VC = 100 + 180 = \280 . $AFC = 100 \div 4 = \$25.00$. $AVC = 180 \div 4 = \$45.00$. $AC = 280 \div 4 = \$70.00$ (check: $AFC + AVC = 25 + 45 = \$70.00$ ✓). $MC = TC(4) - TC(3) = 280 - 250 = \30 .

Notice that AFC falls continuously as output rises (a fixed sum spread over more units — this is why AFC alone is sometimes drawn as a curve that never turns upward). AVC and AC are both U-shaped: they fall as output rises from low levels (fewer workers/machines are under-used, matching the increasing returns seen in Worked example 3), reach a minimum, then rise again once diminishing returns push MC up sharply. MC is also U-shaped and reaches its own minimum (here, \$30 at $Q = 4$) before AVC or AC reach theirs, because MC responds first to changing marginal productivity.

3.1 The relationship between AC and MC

MC and AC are linked mathematically: MC pulls the average in its own direction. In the schedule above, AVC falls from $Q = 1$ to $Q = 5$ ($60.00 \rightarrow 44.00$) while MC ($60, 50, 40, 30, 40$) starts out exactly equal to AVC at $Q = 1$ (both \$60 — MC always equals AVC at the very first unit of output, since $MC(1) = TC(1) - TC(0) = VC(1)$ and $AVC(1) = VC(1) \div 1$, a general identity rather than a coincidence of this table) and then stays below AVC from $Q = 2$ to $Q = 5$ — an extra unit costing less than the current average pulls the average down. From $Q = 6$ onward, MC ($50, 60, 70$) rises above AVC , pulling AVC back up ($45.00 \rightarrow 47.14 \rightarrow 50.00$). The same logic applies to AC , which reaches its lowest value at $Q = 7$ (\$61.43) — very close to where MC (60) is last below AC before MC rises above AC at $Q = 8$ ($70 > 62.50$).

Key relationship: when $MC < AC$ (or AVC), the average is falling; when $MC > AC$ (or AVC), the average is rising; therefore the MC curve cuts the AC curve and the AVC curve exactly at their minimum points. (In a schedule of whole units the crossing may fall between two rows rather than exactly on one, because output changes in discrete steps; with continuously divisible output the crossing is exact.)

4. Economies and diseconomies of scale, and the long-run average cost curve

Economies of scale: falls in long-run average cost (LRAC) as the scale of production (all factors, including capital) increases.

Diseconomies of scale: rises in long-run average cost as the scale of production increases beyond some point.

Minimum efficient scale (MES): the lowest level of output at which LRAC first reaches its minimum; producing at or beyond the MES allows the firm to fully exploit available economies of scale.

Internal economies of scale (arising from the firm's own growth):

- **Technical economies** — large-scale, specialised machinery and production-line specialisation of labour lower cost per unit.
- **Managerial economies** — large firms can afford specialist managers (finance, marketing, production) rather than generalists, raising efficiency.
- **Financial economies** — large, well-established firms can usually borrow at lower interest rates and access a wider range of finance.
- **Purchasing (bulk-buying) economies** — buying raw materials in bulk secures discounts, lowering average variable cost.
- **Risk-bearing economies** — a large firm can spread risk across many products or markets, reducing the average cost of uncertainty.

External economies of scale (arising from growth of the whole industry in an area): a larger local pool of trained, specialised labour; better transport and communications infrastructure built to serve the industry; specialist local suppliers and support services (e.g. component makers, repair firms) clustering nearby.

Internal diseconomies of scale: as a single firm grows very large, coordination, communication and control become harder — decisions take longer, management layers multiply, and worker motivation can fall (alienation in very large workplaces) — raising average cost.

External diseconomies of scale: as a whole industry grows very large in one area, competition for local labour and premises raises wages and rents for every firm, and congestion (traffic, waiting times for inputs) raises costs industry-wide.

Stage of the LRAC curve	What happens to LRAC	Main cause
Early stage — low output	LRAC falls as output rises	Internal (and possibly external) economies of scale dominate
Minimum efficient scale	LRAC reaches its lowest point (may be a single point or a flat "dish" over a range of output)	Economies of scale are exhausted; diseconomies have not yet set in
Later stage — very high output	LRAC rises as output rises further	Internal (and possibly external) diseconomies of scale dominate

Because the long-run average cost curve is built up from many possible short-run average cost curves (one for each plant size a firm could choose), it is often described as an "envelope curve" — at each level of output, LRAC shows the lowest attainable short-run average cost for the plant size best suited to that output.

5. Revenue

Total revenue (TR): the total income a firm receives from selling its output. $TR = P \times Q$ (price $\times$

quantity sold).

Average revenue (AR): revenue per unit sold. $AR = TR \div Q$, which always equals price (P), so the AR curve is identical to the firm's demand curve.

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

5.1 Revenue under different market structures

Market structure	Shape of AR (demand) curve facing the firm	Relationship between AR and MR
Perfect competition (firm is a "price taker")	Perfectly elastic — a horizontal line at the ruling market price	$AR = MR = P$ at every level of output, because the firm can sell any quantity at the same price
Imperfect competition (e.g. monopoly, monopolistic competition, oligopoly — the firm has some price-setting power)	Downward sloping — to sell more, the firm must lower price	MR lies below AR at every output above the first unit, and falls faster than AR (for a straight-line demand curve, MR falls twice as steeply as AR)

Worked example 5 — a revenue schedule for a firm facing a downward-sloping demand curve. A firm's weekly demand curve is $P = 120 - 10Q$ (P in \$, Q in units). TR, AR and MR are calculated below.

Q	P = AR (\$)	TR (\$)	MR (\$)
1	110	110	110 (110 – 0)
2	100	200	90 (200 – 110)
3	90	270	70 (270 – 200)
4	80	320	50 (320 – 270)
5	70	350	30 (350 – 320)
6	60	360	10 (360 – 350)
7	50	350	–10 (350 – 360)
8	40	320	–30 (320 – 350)

TR rises while MR is positive, reaches its maximum where $MR = 0$ (between $Q = 6$ and $Q = 7$, here at $TR = \$360$ when MR falls to just \$10 then turns negative), and falls once MR is negative. AR falls steadily (this is just the demand curve, $P = 120 - 10Q$), while MR falls faster — confirming that MR lies below AR at every output beyond the first unit, as the table above predicts for imperfect competition.

6. Profit maximisation: $MC = MR$

Profit-maximising rule: a firm maximises profit at the output where marginal cost equals marginal revenue ($MC = MR$), provided MC is rising through MR at that point (i.e. cutting MR from

below). At any output where $MR > MC$, producing one more unit adds more to revenue than to cost, so profit rises; at any output where $MC > MR$, the last unit produced costs more than it earns, so profit would rise by producing less. Profit is therefore greatest exactly where the two are equal.

Worked example 6 — finding the profit-maximising output. Using the marginal revenue schedule from Worked example 5 (MR: 110, 90, 70, 50, 30, 10, -10, -30 for $Q = 1$ to 8) and the marginal cost schedule from Worked example 4 (MC: 60, 50, 40, 30, 40, 50, 60, 70 for $Q = 1$ to 8):

Q	MR (\$)	MC (\$)	Compare
1	110	60	$MR > MC$ — expand output
2	90	50	$MR > MC$ — expand output
3	70	40	$MR > MC$ — expand output
4	50	30	$MR > MC$ — expand output
5	30	40	$MR < MC$ — do not produce this unit

MR exceeds MC for every unit up to and including the 4th, and MC exceeds MR from the 5th unit onward — so the **profit-maximising output is $Q = 4$ units**.

Calculating profit at $Q = 4$: TR at $Q = 4 = \$320$ (from Worked example 5). TC at $Q = 4 = \$280$ (from Worked example 4, FC \$100 + VC \$180). Profit = TR – TC = $320 - 280 = \$40$.

Check using AR and AC per unit: AR at $Q = 4 = P = \$80$. AC at $Q = 4 = \$70$ (from Worked example 4). Profit per unit = AR – AC = $80 - 70 = \$10$. Total profit = $\$10 \times 4 \text{ units} = \$40 \checkmark$, matching the TR – TC calculation.

7. Normal profit, supernormal (abnormal) profit and loss-minimising in the short run

Normal profit: the minimum reward needed to keep an entrepreneur's resources (capital and enterprise) employed in their current use; it is treated as a cost of production (part of AC) because it is the opportunity cost of enterprise. Normal profit occurs where $AR = AC$ (total revenue exactly equals total cost, so economic/supernormal profit is zero).

Supernormal (abnormal) profit: profit earned above normal profit, i.e. $AR > AC$. It is the "surplus" that attracts new firms into a competitive industry.

Loss: occurs where $AR < AC$. A firm can still be minimising its loss (rather than shutting down immediately) provided $AR > AVC$ — it is covering all variable costs and making some contribution towards fixed costs.

Shutdown (close-down) rule: in the short run, a firm should stop producing immediately if AR (price) falls below AVC, because it can no longer even cover the cost of the inputs used to make each unit; producing would add losses beyond the fixed costs it must pay regardless.

Worked example 7 — three short-run outcomes, same cost structure. Using the cost schedule from Worked example 4 ($FC = \$100$), assume the firm is a price taker (perfect competition), so $P = AR = MR$, and it sets output where $P = MC$.

(a) Supernormal profit. Market price $P = \$70$. Since $MC = 70$ only at $Q = 8$, the firm produces 8 units. $TR = 70 \times 8 = \$560$. TC at $Q = 8 = \$500$. Profit = $560 - 500 = \$60$ (**supernormal profit**). Check per unit: $AR = \$70$, AC at $Q = 8 = \$62.50$, profit per unit = $\$7.50 \times 8 = \60 ✓. $AR > AC$ confirms supernormal profit.

(b) Normal profit only. If price settles at AC 's minimum, here approximately $\$61.43$ (at $Q = 7$, where AC is lowest), then $AR \approx AC$ and profit $\approx \$0$ — the firm earns normal profit only, with no economic surplus. This is the position a competitive industry tends towards in long-run equilibrium, once supernormal profits have attracted enough new entrants to push price back down to minimum AC .

(c) Loss-minimising. Market price falls to $P = \$60$. $MC = 60$ at $Q = 7$ (on the rising part of the MC schedule). $TR = 60 \times 7 = \$420$. TC at $Q = 7 = \$430$. Loss = $TR - TC = 420 - 430 = -\$10$. Should the firm keep producing? AVC at $Q = 7$ is $\$47.14$ (rounded; the exact value is $330 \div 7 = \$47.142857\dots$), and price ($\60) is above this, so the firm is covering all its variable costs and contributing $TR - VC = 420 - 330 = \$90$ towards its $\$100$ of fixed costs — reducing its loss to $\$10$ rather than losing the full $\$100$ it would lose by shutting down. (Multiplying price minus the rounded AVC by output, $60 - 47.14 = 12.86$, then $\times 7$, gives $\$90.02$; the small difference from $\$90$ is just a rounding artefact of using AVC to two decimal places — the exact figure is $\$90$.) The firm should therefore keep producing in the short run: it is loss-minimising, not shutting down.

(d) Shutdown. If price fell further, to below the minimum AVC of $\$44$ (e.g. $P = \$40$), the firm could not even cover its variable costs at any output and should stop producing immediately, limiting its loss to the $\$100$ fixed cost alone rather than losing more by continuing to produce.

Key terms

Term	Definition
Utility	The satisfaction or benefit a consumer gains from consuming a good or service.
Marginal utility (MU)	The extra utility gained from consuming one more unit of a good.
Law of diminishing marginal utility	The principle that marginal utility from successive units of a good eventually falls as consumption rises.
Consumer equilibrium	The combination of goods that maximises a consumer's total utility for a given income, where MU per dollar is equalised across goods.
Substitution effect	The change in quantity demanded due to a change in relative price alone, real income held constant.
Income effect	The change in quantity demanded due to a change in real purchasing power caused by a price change.
Normal good	A good for which quantity demanded rises as income rises.
Inferior good	A good for which quantity demanded falls as income rises.

Giffen good	A strongly inferior good whose quantity demanded falls when its price falls, because a negative income effect outweighs the substitution effect.
Short run	The period in which at least one factor of production is fixed.
Long run	The period in which all factors of production are variable.
Law of diminishing (marginal) returns	As more units of a variable factor are added to a fixed factor, marginal product eventually falls.
Total product (TP)	The total output produced by a given quantity of factors, usually labour.
Marginal product (MP)	The extra output from employing one more unit of the variable factor.
Average product (AP)	Output per unit of the variable factor, $AP = TP \div \text{factor quantity}$.
Fixed cost (FC)	A cost that does not vary with output in the short run.
Variable cost (VC)	A cost that varies directly with the level of output.
Total cost (TC)	$FC + VC$, the full cost of producing a given output.
Average cost (AC)	Total cost per unit of output, $AC = TC \div Q$.
Marginal cost (MC)	The extra cost of producing one more unit of output.
Economies of scale	Falls in long-run average cost as the scale of production increases.
Diseconomies of scale	Rises in long-run average cost as the scale of production increases beyond some point.
Minimum efficient scale (MES)	The lowest output at which long-run average cost first reaches its minimum.
Long-run average cost curve (LRAC)	A curve showing the lowest attainable average cost of production at each output when all factors are variable.
Total revenue (TR)	Total income from sales, $TR = P \times Q$.
Average revenue (AR)	Revenue per unit sold, $AR = TR \div Q = \text{price}$.
Marginal revenue (MR)	The extra revenue from selling one more unit.
Profit maximisation	Producing at the output where marginal cost equals marginal revenue ($MC = MR$).
Normal profit	The minimum profit needed to keep resources in their current use; occurs where $AR = AC$.
Supernormal (abnormal) profit	Profit above normal profit; occurs where $AR > AC$.
Loss-minimising	Continuing to produce at a loss in the short run because price still exceeds average variable cost.
Shutdown (close-down) point	The point where price equals average variable cost; below this the firm should stop producing in the short run.

Quick check

1. Define the term *marginal utility*. [2]

2. State the law of diminishing marginal returns. [2]
3. Distinguish between the short run and the long run in the theory of production. [4]
4. Explain, using the concept of diminishing marginal utility, why a typical demand curve slopes downwards from left to right. [4]
5. Distinguish between the income effect and the substitution effect of a change in the price of a good. [4]
6. A firm has fixed costs of \$50. Its variable costs are shown below.

Q	VC (\$)
1	40
2	70
3	90
4	120
5	170

Calculate the average fixed cost, average variable cost, average total cost and marginal cost of production at each level of output. [8]

7. Explain why the marginal cost curve cuts both the average variable cost curve and the average total cost curve at their lowest points. [4]
8. Analyse how internal economies of scale and internal diseconomies of scale together determine the shape of a firm's long-run average cost curve. [6]
9. Distinguish between normal profit and supernormal (abnormal) profit. [4]
10. A firm's marginal revenue is $MR = 80 - 4Q$ and its marginal cost is $MC = 20 + 2Q$, where Q is weekly output. Calculate the firm's profit-maximising output. [4]
11. A price-taking firm's average variable cost is \$30 at its current output, and the market price has fallen to \$25. Using the shutdown (close-down) rule, explain whether the firm should continue to produce in the short run. [4]
12. Discuss whether a profit-maximising firm will always choose to produce at the output where average total cost is at its lowest point. [8]

1. Marginal utility is the extra (additional) satisfaction or utility a consumer gains from consuming one more unit of a good, with consumption of all other goods held constant.

2. As successive equal units of a variable factor are added to a fixed factor of production in the short run, a point is eventually reached beyond which the marginal product of the variable factor declines.

3. The short run is the period in which at least one factor of production is fixed (cannot be altered), so output can only be changed by varying the other, variable factors. The long run is the period in which all factors of production, including capital and plant size, are variable, so the firm can change its entire scale of operation; no factor is fixed.

4. A rational consumer will only buy an extra unit of a good if the utility it provides is worth at least as much as its price. Because marginal utility diminishes as more units are consumed, the price a

consumer is willing to pay for successive units falls as quantity rises. Since the (money value of the) marginal utility schedule effectively traces out the maximum price the consumer will pay at each quantity, it behaves as the consumer's demand curve — and because MU falls, this curve slopes downward: a lower price is needed to persuade the consumer to demand a greater quantity.

5. The substitution effect is the change in quantity demanded resulting purely from the change in the good's price relative to other goods, with real income held constant — consumers switch towards the good that has become relatively cheaper (or away from one that has become relatively dearer). The income effect is the change in quantity demanded resulting from the change in the consumer's real purchasing power caused by the price change, with relative prices held constant. For a normal good both effects push quantity demanded in the same direction as the price change (in reverse), reinforcing each other; for an inferior good the income effect works in the opposite direction to the substitution effect, but is normally the smaller of the two, so quantity demanded still moves as the basic law of demand predicts.

6.

Q	TC (\$)	AFC (\$)	AVC (\$)	AC (\$)
1	90	50.00	40.00	90.00
2	120	25.00	35.00	60.00
3	140	16.67	30.00	46.67
4	170	12.50	30.00	42.50
5	220	10.00	34.00	44.00

Q	MC (\$)
1	40
2	30
3	20
4	30
5	50

Method shown for Q = 3: $TC = 50 + 90 = \$140$; $AFC = 50 \div 3 = \$16.67$; $AVC = 90 \div 3 = \$30.00$; $AC = 140 \div 3 = \$46.67$; $MC = TC(3) - TC(2) = 140 - 120 = \20 .

7. Marginal cost is the cost of producing one extra unit. Whenever MC is below the current average (AVC or AC), that cheaper extra unit pulls the average down, so the average is falling; whenever MC is above the current average, the extra unit pulls the average up, so the average is rising. The average can therefore only stop falling and start rising at the exact output where MC equals it — which is, by definition, the average's minimum point. This is why the MC curve passes through the lowest points of both the AVC curve and the AC curve.

8. As a firm's scale of production increases from a low level, internal economies of scale — technical economies from large, specialised machinery, managerial economies from employing specialist staff, financial economies from cheaper access to credit, bulk-purchasing discounts, and

risk-bearing economies from a wider product range — reduce the cost of producing each unit, so long-run average cost (LRAC) falls. This continues until the firm reaches its minimum efficient scale, where LRAC is at its lowest. Beyond this point, if the firm continues to expand, internal diseconomies of scale set in: coordination, communication and control become more difficult across a very large organisation, decision-making slows, and worker motivation can fall, all of which raise average cost. The result is a U-shaped (or saucer-shaped) LRAC curve: falling while economies of scale dominate, flat near the minimum efficient scale, and rising once diseconomies of scale dominate.

9. Normal profit is the minimum profit required to keep an entrepreneur's resources employed in their current use; because it is the opportunity cost of enterprise, it is included as a cost within average cost, and it occurs where average revenue equals average cost ($AR = AC$). Supernormal (abnormal) profit is any profit earned above normal profit, occurring where average revenue exceeds average cost ($AR > AC$); it is the surplus that can attract new firms into a competitive industry, whereas normal profit alone is enough to keep an existing firm operating.

10. Set $MR = MC$: $80 - 4Q = 20 + 2Q$. Rearranging: $80 - 20 = 2Q + 4Q$, so $60 = 6Q$, giving **$Q = 10$ units**.

11. Price (\$25) is below average variable cost (\$30), so the firm cannot cover even the variable costs of production at its current output; every unit produced adds to its losses on top of the fixed costs it must pay regardless. The shutdown rule states that the firm should therefore stop producing (produce zero output) in the short run, since this limits its loss to fixed costs alone rather than allowing losses to grow larger by continuing to produce.

12. Not necessarily. The profit-maximising rule is to produce where $MC = MR$, which is a completely separate condition from producing at the output where average total cost is lowest. For a firm facing a downward-sloping demand curve (imperfect competition, e.g. monopoly or monopolistic competition), MR lies below AR at every output, so the firm typically maximises profit at an output below the level at which AC would be minimised — it restricts output to raise price, sacrificing some productive/technical efficiency for higher profit. The two objectives coincide only in a special case: long-run equilibrium under perfect competition, where competitive pressure from new entrants and exiting firms forces price down until $AR = MC = AC$ simultaneously, so the profit-maximising output and the minimum- AC output are the same — but at that point the firm earns only normal profit, not supernormal profit. In general, therefore, a profit-maximising firm will choose the output where $MC = MR$, and this need not be, and for firms with market power usually is not, the output at which average total cost is minimised.