ALLAMA IQBAL OPEN UNIVERSITY

USMAN NOTES
www.usmannotes.com

Name:

Registration No.:

Course Code: 801

Program: MSc. Economics

Semester: Autumn 2026

ALLAMA IQBAL OPEN UNIVERSITY, ISLAMABAD

Course: Introduction to Microeconomics (801) • Level: MSc. Economics

Semester: Autumn 2026

Assignment No. 2 (Units 6–9)

Total Marks: 100 • Pass Marks. 40

SOLVED ASSIGNMENT GUIDE

(For Study & Reference)

Note: This guide is prepared for study and reference purposes. Students are advised to read the relevant units of the course book, understand the concepts, and then write the assignment in their own words.

USMAN NOTES www.usmannotes.com

Question 1: Short-Run Supply Curve of the Perfectly Competitive Firm (20 Marks)

Q. 1 Explain short run supply curve of the perfectly competitive firm.

Introduction

A supply curve shows how much a firm is willing and able to offer for sale at each possible price. Under perfect competition the firm is a price taker: the market, through the interaction of industry demand and supply, sets the price, and the individual firm, being too small to influence it, takes that price as given and decides only how much to produce. The short-run supply curve of such a firm is not drawn from a separate theory; it is read directly from its cost curves. This answer derives it step by step. It recalls the short-run cost structure, states the profit-maximising rule (produce where price equals marginal cost), introduces the shutdown condition that price must cover average variable cost, and shows that the firm's short-run supply curve is precisely the portion of its marginal cost curve that lies above the minimum point of average

variable cost. It then explains how the industry supply curve is obtained and why the result matters for market analysis.

Assumptions and the Firm's Short-Run Position

The derivation rests on the standard assumptions of perfect competition: there are very many small buyers and sellers, the product is homogeneous, firms and resources are free to enter, leave and move, and buyers and sellers have perfect knowledge of prices. In the short run the firm has a fixed plant, so its costs divide into fixed costs, which are already committed, and variable costs, which it still controls. Because the firm is a price taker, its demand curve is a horizontal line at the market price: it can sell any quantity it chooses at that price, and nothing above it, so for the competitive firm price, average revenue and marginal revenue are all equal, P = AR = MR. This equality is what gives the competitive supply analysis its special simplicity: the firm's marginal revenue does not fall as it sells more, so the question 'how much should I produce?' reduces to comparing the given price, unit by unit, with the cost of producing successive units.

The Production Rule: Price Equals Marginal Cost

A profit-maximising firm expands output so long as the revenue from one more unit exceeds the cost of producing it, and stops before the point where cost would exceed revenue. For the competitive firm, marginal revenue is simply the price, so the rule becomes: produce the output at which P = MC, choosing the rising portion of the marginal cost curve. As the market price rises, the equality P = MC is satisfied at a larger output because MC rises with output under diminishing returns; as the price falls, the equality is satisfied at a smaller output. If we performed this exercise for every possible price, we would trace out a set of price–quantity

pairs, each one optimal for the firm. Those pairs are exactly the points of the marginal cost curve itself — which is the first half of the result: the marginal cost curve is the candidate for the firm's supply curve, because it records, for each quantity, the lowest price at which the firm would willingly produce that quantity.

The Shutdown Condition: Price Must Cover Average Variable Cost

The marginal cost curve cannot be the whole story, because there is a range of low prices at which the firm should produce nothing at all. In the short run the firm's fixed costs must be paid whether it produces or not; they are, for this decision, unavoidable. What the firm can avoid by shutting down is its variable costs. It therefore pays to produce only if revenue covers at least the variable costs of production, that is, if $P \geq AVC$. If price equals or exceeds minimum average variable cost, producing contributes the whole of variable cost and something — or at worst nothing less than zero — toward fixed costs, so losses are no greater, and generally smaller, than the fixed costs that closing would leave behind. If price falls below minimum AVC, every unit sold fails even to cover its own variable cost, so each unit produced adds to the loss; the firm minimises its loss by producing zero and losing only its fixed costs. The minimum point of AVC is therefore called the shutdown point, and prices below it belong to no supply curve at all: quantity supplied at such prices is zero.

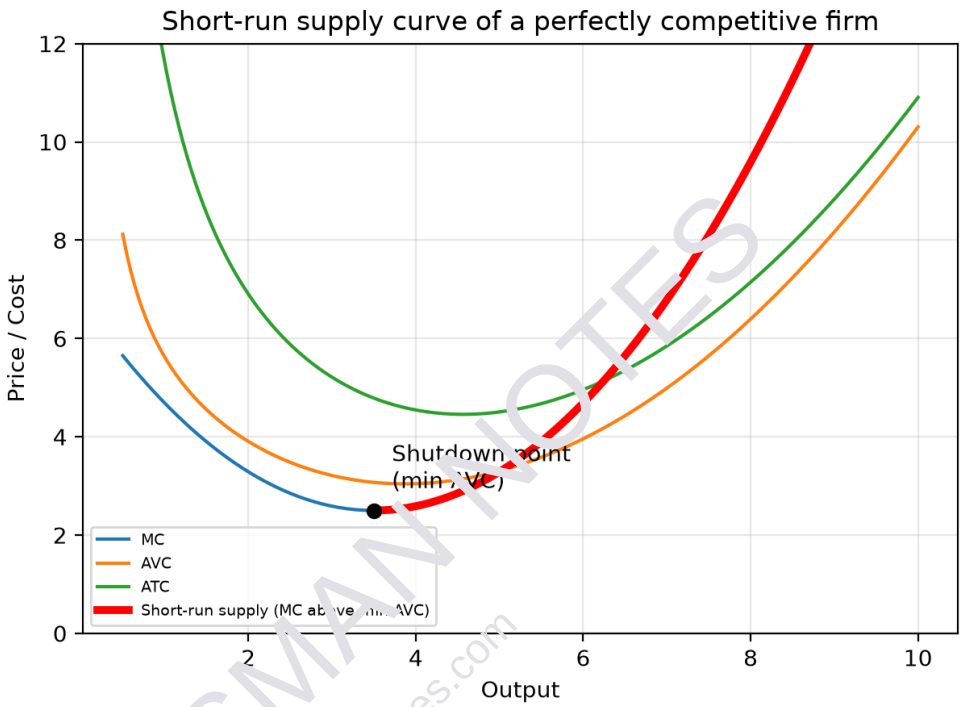


Figure 1: The short-run supply curve is the rising portion of the MC curve above the shutdown point (minimum AVC), shown in bold red.

USMAN NOTES www.usmannotes.com

The Supply Curve Stated Precisely

Combining the production rule with the shutdown condition gives the complete result. The short-run supply curve of the perfectly competitive firm is that part of its short-run marginal cost curve which lies above the minimum point of its average variable cost curve. At any price above minimum AVC, quantity supplied is read from the MC curve where P = MC on its rising branch. At the price exactly equal to minimum AVC, the firm is indifferent between producing the shutdown output and producing nothing, since its loss equals fixed cost either way. Below that price, quantity supplied is zero. Because MC slopes upward in the relevant range — a consequence of the law of diminishing returns raising the cost of successive units — the resulting supply curve slopes upward from left to right, providing the microeconomic foundation for the law of supply. Notice also what the supply curve does not depend on: fixed

costs do not appear in it at all, because neither MC nor AVC is affected by the level of fixed cost. A lump-sum tax or a rise in rent shifts ATC upward and reduces profit, but does not change the quantity supplied at any price in the short run, whereas a per-unit tax or a rise in input prices shifts MC and AVC up and reduces quantity supplied at every price.

Table 1: Price, the firm's response, and its result

Market price	Firm's decision	Result
P > minimum ATC	Produce where P = MC	Supernormal (economic) profit
P = minimum ATC	Produce where P = MC	Break-even: normal profit only
Minimum AVC < P < minimum ATC	Produce where P = MC	Loss, but smaller than fixed cost — keep producing
P = minimum AVC	Shutdown point — indifferent	Loss equals total fixed cost either way
P < minimum AVC	Shut down; supply = 0	Loss equals total fixed cost

USMAN NOTES www.usmannotes.com

From the Firm's Supply to the Industry Supply Curve

The short-run industry or market supply curve is obtained by the horizontal summation of the supply curves of all the firms in the industry: at each price, the quantities that every firm supplies along its MC curve above AVC are added together. With many firms, the industry curve is smoother and flatter than any single firm's curve. One qualification is needed for large industry-wide expansions: when all firms expand together, they bid up the prices of industry-specific inputs, which shifts every firm's cost curves upward, so the long-run industry supply may slope up even where each firm has flat long-run costs. In the strict short run, however, with plant and input prices given, the summed MC curves are the industry supply curve, and its intersection with industry demand determines the market price that each firm then takes as given — completing the circle of the competitive model.

Importance and Limitations

The result is important for three reasons. First, it anchors the supply side of the market in the real technology and costs of firms, instead of treating supply as an unexplained line: anything that changes marginal or variable costs — wages, fuel prices, technology, per-unit taxes and subsidies — shifts supply in a predictable way through these curves. Secondly, it explains observed behaviour that otherwise looks irrational, such as firms continuing to operate at a loss for months: they are covering variable costs and contributing to fixed costs, which is rational in the short run. Thirdly, it gives the benchmark against which real markets are judged in competition policy and welfare analysis. Its limitations follow from its assumptions: real firms often face downward-sloping demand, uncertain prices and multiple products, so the simple P = MC reading must be adapted; and the curve describes intended supply at each price, assuming the firm can adjust output freely, whereas capacity limits and delivery lags can make very-short-run supply almost vertical. These qualifications extend the model rather than overturn it.

Conclusion

To conclude, the short-run supply curve of a perfectly competitive firm is derived, not assumed: the firm produces where price equals marginal cost, provided price covers average variable cost, and shuts down otherwise. The supply curve is therefore the rising portion of the marginal cost curve above the minimum point of the average variable cost curve, with quantity supplied equal to zero below that shutdown price. Summed horizontally across firms, it becomes the industry supply curve whose interaction with demand sets the market price. The derivation beautifully illustrates the method of microeconomics: a market-level regularity, the upward-

sloping supply curve, is shown to follow from individual cost-minimising and profit-maximising behaviour.

Question 2: Bilateral Monopoly (20 Marks)

Q. 2 Explain the concept of Bilateral Monopoly.

Introduction

Market structures are usually classified by counting sellers and buyers separately: many sellers with one buyer is monopsony, one seller with many buyers is monopoly. Bilateral monopoly is the limiting case in which both sides of the market are reduced to a single participant — one seller, who is a monopolist, faces one buyer, who is a monopsonist. Each side therefore possesses market power and each knows that the other knows it. This situation, though rare in pure form, is far from a theoretical curiosity: it describes a single employer facing a single trade union in a company town, a sole government buyer purchasing from a sole defence contractor, or a single large retailer negotiating with a sole supplier of a patented input. This answer explains the structure of bilateral monopoly, why economic theory finds its outcome indeterminate within a range, how the upper and lower limits of that range are set, and how bargaining, negotiation and institutions decide where, within the range, the final settlement lies.

Meaning and Structure

In a bilateral monopoly, the monopolist seller wants to charge the highest profitable price and restrict quantity in the manner of an ordinary monopolist, equating marginal revenue with marginal cost. The monopsonist buyer wants the opposite: to pay the lowest possible price and

to restrict purchases in the manner of an ordinary monopsonist, buying only up to the point where the marginal value of the good to him equals its marginal expenditure, which for a sole buyer lies above the supply price because buying more bids the price up. Unlike the standard models, neither side can simply impose its preferred solution, because each needs the other: the seller has no alternative customer for his output, and the buyer has no alternative source of supply. Trade will normally still take place, because there is a surplus to be shared — the good is worth more to the buyer than it costs the seller to produce over a range of quantities — but the division of that surplus, and usually the price, cannot be read off from cost and demand curves alone.

Why the Outcome Is Indeterminate

The famous conclusion of the analysis, associated with Edgeworth and developed by later writers, is that bilateral monopoly is indeterminate: theory can fix a contract zone, also called the range of possible settlements, but cannot fix a unique price within it. Suppose the seller, if he could dictate terms, would set the monopoly price, high and with quantity restricted to where his marginal revenue equals his marginal cost. Suppose instead the buyer, if he could dictate terms, would impose the monopsony price, low, restricting purchases to where the marginal expenditure equals the good's marginal value to him. Any price between these two limits can persist, because at any such price both parties are better off trading than not trading at all, and neither has a market mechanism, such as competition from a third party, to push the price further in his favour. What determines the final resting point is bargaining strength: patience, information, the credibility of threats to walk away, the costs each side bears from delay or breakdown, legal rights, and sheer negotiating skill. Economics here reaches the

boundary of the price theory of competitive markets and hands over to bargaining theory and game theory.

Quantity, interestingly, may be less indeterminate than price. If the two parties bargain over both price and quantity together — as rational negotiators sharing a surplus should — they have a joint interest in choosing the quantity that maximises their combined surplus, because a larger total surplus leaves more for both to divide. That joint-surplus-maximising quantity is the competitive quantity, where the buyer's marginal value equals the seller's marginal cost: every unit up to that point adds more value than cost, and every unit beyond it would subtract from the total. The bargain then determines only the price, which divides this maximised surplus between them. If, however, negotiations are restricted to price alone, with quantity following from that price, the outcome is inefficient: a monopoly-level price leads the buyer to purchase too little, and a monopsony-level price leads to too little being offered. Whether bilateral monopoly is efficient therefore depends on the scope of the bargain, a result with direct practical importance for how labour and supply contracts are structured.

The Labour-Market Example: Union and Sole Employer

The classic application is a labour market with one employer and one trade union that is the sole supplier of labour. The employer, as monopsonist, would like to set the wage where marginal revenue product equals marginal cost of labour, employing relatively few workers at a relatively low wage, exploiting the fact that hiring more bids the wage up for all. The union, as monopolist of labour, would like to set a higher wage, analogously to a seller restricting supply, though unions differ in their objectives: some maximise the wage rate, some maximise total wage income, and some maximise employment for their members. Between the monopsony

wage the employer prefers and the wage the union demands lies the contract zone. The actual settlement emerges from collective bargaining: negotiations, the threat and cost of strikes and lockouts, the financial reserves of each side, public opinion and, often, government conciliation. This is precisely why industrial-relations outcomes cannot be forecast from demand and supply curves alone, and why two similar industries can settle at noticeably different wages.

Table 2: Bilateral monopoly compared with one-sided market power

Feature	Monopoly	Monopsony	Bilateral monopoly
Sellers / buyers	One / many	Many / one	One / one
Power held by	Seller	Buyer	Both sides
Price tendency	High (seller's optimum)	Low (buyer's optimum)	Indeterminate — between the two optima
Decided by	Seller's MR = MC	Buyer's marginal expenditure rule	Bargaining within the contract zone
Typical example	Patented drug supplier	Sole mill buying local cotton	Single employer facing a single union

Modern Treatment: Bargaining and Game Theory

Twentieth-century economics added formal tools for the indeterminate range. Nash's bargaining solution proposes that rational negotiators will divide the surplus so as to maximise the product of their gains over their disagreement outcomes, which in symmetric cases means splitting the gains evenly after accounting for what each side gets if talks fail. The practical lesson is that a party's share depends heavily on its outside option: a union whose members can find other work, or an employer who can relocate or automate, bargains from strength. Game theory also shows how repeated dealings change behaviour: a seller and buyer who must deal with each other every year have reasons to preserve reputation and avoid squeezing the

other to the point of exit, which pushes settlements toward the middle of the range. Legal and institutional rules — minimum wages, arbitration requirements, competition law and rules on essential facilities — deliberately narrow the contract zone where society considers an unrestricted bargain unacceptable, for example where the buyer is a public utility or the good is an essential medicine.

Advantages, Disadvantages and Welfare

Bilateral monopoly is not necessarily worse than one-sided monopoly; in one respect it can be better. Opposing market powers can partly cancel out: the buyer's power restrains the seller's price, and the seller's power restrains the buyer's attempt to depress it, so the final price can lie closer to the competitive level than under pure monopoly — an idea Galbraith called countervailing power. Against this, the indeterminacy itself is costly. Negotiations consume time and management attention, the risk of breakdown — strikes, supply interruptions, litigation — imposes real losses on both sides and on third parties, and the resulting price may reflect relative stubbornness or asymmetric information rather than costs and values, which weakens the allocative signals prices are supposed to send. There is also a risk that the two powerful parties settle comfortably at the expense of outsiders: a union and employer, for example, may agree on terms that are financed by higher prices to unorganised consumers, converting bilateral monopoly into a quiet alliance against the public.

Conclusion

In conclusion, bilateral monopoly is the market structure in which a single seller faces a single buyer, so that monopoly and monopsony power confront each other directly. Economic theory shows that the outcome is indeterminate within a contract zone bounded by the seller's

preferred monopoly price above and the buyer's preferred monopsony price below; if the parties bargain over quantity as well as price, quantity tends toward the joint-surplus-maximising, competitive level, while the price that divides the surplus is settled by bargaining strength, outside options and institutions. The concept matters because it marks the honest limit of mechanical price theory, introduces bargaining and game theory into market analysis, and fits real situations — collective bargaining, sole-supplier procurement and regulated utilities — that simpler models cannot describe.

Question 3: Excess Capacity (20 Marks)

Q. 3 Explain the concept of excess capacity in detail.

Introduction

Excess capacity is one of the most debated ideas in the theory of monopolistic competition. In the long run, firms in a monopolistically competitive industry — many sellers of differentiated products, with freedom of entry — end up producing at an output smaller than the output at which their long-run average cost is lowest. The difference between the cost-minimising, 'ideal' output and the output actually produced is called excess capacity: the plant, so the charge runs, is larger than its use justifies, resources are spread too thinly across too many firms, and society pays for product variety with productive inefficiency. This answer defines the concept rigorously, derives it from the Chamberlinian tangency equilibrium, measures it on the cost diagram, and then gives a balanced evaluation, because later economists have strongly disputed whether the unused 'capacity' is really wasted, or whether it is simply the price of the variety that consumers value.

Meaning of Excess Capacity

Excess capacity is the difference between the ideal output — the output at which long-run average cost (LAC) is at its minimum, sometimes called the optimum or full-capacity output — and the output actually produced in long-run equilibrium. It is not the same as a firm temporarily working below capacity because demand is slack in a recession; it is a permanent, structural feature of the equilibrium itself. A firm with excess capacity could lower its unit cost by producing more with the same plant and organisation, yet it does not do so, because at any larger output it could not sell the extra quantity without cutting its price so deeply that profit would fall. Excess capacity therefore expresses a tension between what is technically efficient for the firm (produce at minimum average cost) and what is privately profitable under the demand conditions it faces (stop where its downward-sloping demand curve still permits normal profit).

Derivation: The Tangency Equilibrium of Monopolistic Competition

The derivation follows Chamberlin's large-group model. Each firm sells a differentiated product, so it faces a downward-sloping demand (average revenue) curve: it has a small monopoly over its own variety, but close substitutes limit its power. In the short run the firm may earn supernormal profit. Free entry, however, means that new firms are attracted by those profits, each bringing another close substitute; the demand curve of every existing firm shifts leftward and becomes more elastic as its market is shared among more rivals. Entry continues until supernormal profit is eliminated, which occurs when each firm's demand curve is tangent to its long-run average cost curve, so that price equals LAC and only normal profit is earned. The critical geometrical fact is that a downward-sloping demand curve can be tangent to a U-

shaped LAC curve only on the falling portion of LAC, to the left of its minimum point. Tangency at the minimum itself would require the demand curve to be horizontal there, which holds only under perfect competition. Hence long-run equilibrium output necessarily falls short of the ideal output, and the gap between them is the excess capacity.

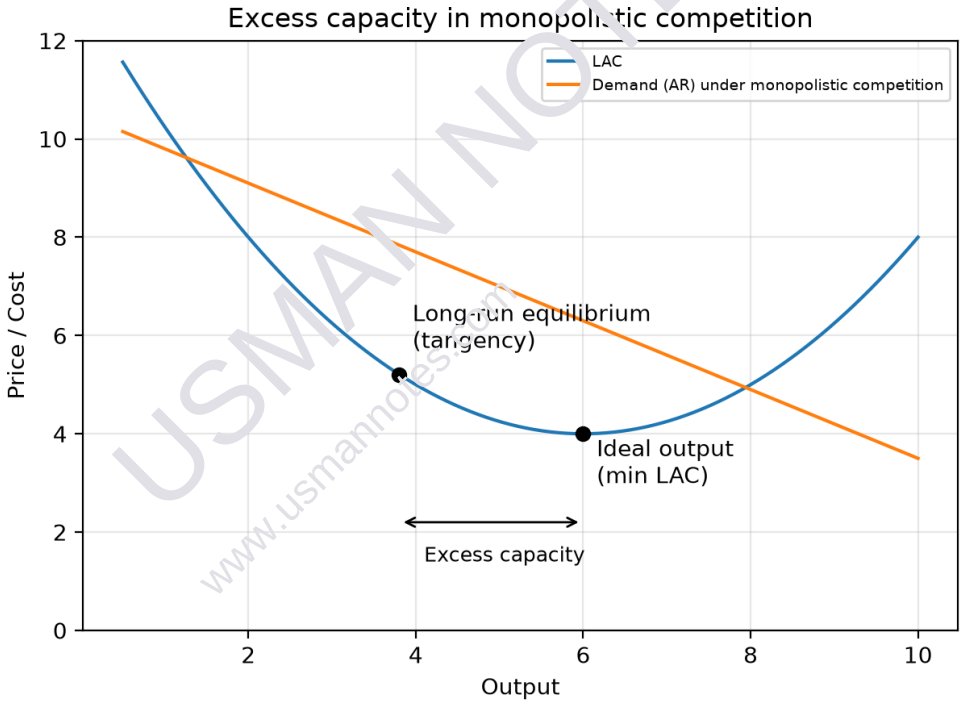


Figure 2: Long-run tangency equilibrium under monopolistic competition lies to the left of minimum LAC; the distance to the ideal output measures excess capacity.

Measurement and a Worked Illustration

Excess capacity is measured in units of output: Excess capacity = Ideal output (at minimum LAC) − Actual long-run equilibrium output. Suppose the firms in an industry have a minimum long-run average cost of Rs 50 at an output of 10,000 units per period. Under monopolistic competition, entry pushes each firm to a tangency at, say, 6,500 units, where price and LAC both equal Rs 58. Excess capacity is then 3,500 units per firm — the firm operates at 65 per cent of its ideal scale, and unit cost is Rs 8 higher than the technically achievable minimum.

Aggregated across an industry with many such firms, the same total output could, in a purely technical sense, be produced by fewer, larger firms at lower unit cost, employing less capital and fewer fixed resources in total. Chamberlin himself was careful to note that the calculation holds the product and its differentiation constant: the 'waste' exists only if the varieties among which output is spread are regarded as worthless duplication, which is exactly the point his critics attack.

Comparison with Perfect Competition

The contrast with perfect competition sharpens the concept. Under perfect competition the firm's demand curve is horizontal, entry likewise eliminates supernormal profit, but tangency — the equality of price and LAC — occurs exactly at the minimum point of LAC. There is no excess capacity in long-run competitive equilibrium: firms are of optimum size, price equals both marginal cost and minimum average cost, and productive as well as allocative efficiency is achieved. Under monopolistic competition, price exceeds marginal cost (because demand slopes down while profit maximisation sets MR = MC) and price is above minimum LAC (because of the leftward tangency). Excess capacity is thus one element of a package of differences: somewhat higher prices, somewhat smaller firms, and a great deal more variety than under perfect competition. Under monopoly, by contrast, there is no entry to force any tangency at all, so the concept applies only loosely; a monopolist may or may not produce near minimum LAC depending on the size of the market relative to the cost structure.

Table 3: Excess capacity across market structures (long run)

Feature	Perfect competition	Monopolistic competition	Monopoly

Demand curve of firm	Horizontal	Downward sloping	Downward sloping (market demand)
Price vs minimum LAC	Equal	Above minimum LAC	Not determined by entry
Output vs ideal output	Equal — no excess capacity	Below — excess capacity exists	Restricted; capacity depends on market size
Profit in long run	Normal only	Normal only	Supernormal may persist

Is Excess Capacity Really Waste? The Debate

The case for calling excess capacity wasteful is straightforward. Resources — buildings, machines, managerial effort, shop space — are embodied in firms that operate below their efficient scale, so the same output could be produced more cheaply. High streets with dozens of half-empty similar shops and restaurants, or industries with many small plants where a few large ones would do, are offered as everyday evidence. Critics add that product differentiation is often trivial or purely persuasive, created by advertising rather than by real differences, so the variety that 'justifies' the capacity may itself be manufactured and socially doubtful. On this view, monopolistic competition taxes consumers twice: through price above marginal cost and through costs above the minimum.

The defence is equally serious and is now widely accepted in qualified form. First, consumers value variety for its own sake: a world in which the same total output came from a few optimum-size firms would offer fewer styles, locations, flavours and services, and consumer surplus from matching individual tastes would fall — the supposedly wasted capacity is the supply side of a benefit the simple cost comparison omits. Secondly, some apparent spare capacity is genuinely useful: restaurants, transport and repair services must hold capacity for

peak periods, and a firm that could serve the peak with no slack at all would turn customers away for much of the day. Thirdly, differentiation is frequently real — quality, reliability, after-sales service, location and trust differ between sellers, and locating nearer to customers saves their time and transport cost even while raising the seller's unit cost. The modern verdict is therefore measured: excess capacity is a real and rigorously derived feature of the model, but labelling all of it 'waste' requires the extra judgement that the variety it supports is worth less than the cost saving, which must be argued case by case rather than assumed.

Conclusion

To conclude, excess capacity is the shortfall of long-run equilibrium output below the ideal, minimum-LAC output that arises under monopolistic competition because free entry drives each firm's downward-sloping demand curve into tangency with the falling portion of its LAC curve. It is measured simply, as ideal output minus actual output, and it distinguishes monopolistic competition sharply from perfect competition, where no such gap exists. Whether the gap represents social waste or the legitimate cost of product variety, convenience and peak-load readiness remains contested, and a good answer, like a good policymaker, should present the derivation with confidence and the welfare judgement with care.

Question 4: Collusive Oligopoly (20 Marks)

Q. 4 Explain in detail the concept of Collusive Oligopoly.

Introduction

Oligopoly is the market structure dominated by a few large firms, each large enough that its decisions visibly affect its rivals, and each knowing it. That mutual interdependence is the

source of all the distinctive problems of oligopoly, including the temptation examined in this question: instead of competing, the firms may collude — agree, formally or informally, to act jointly so as to raise price and profit above the competitive level, in effect behaving as a shared monopoly. Collusive oligopoly therefore ranges from the explicit, legally organised cartel that fixes prices and quotas, through price leadership arrangements, to the silent, tacit understanding by which firms avoid undercutting one another. This answer explains collusion in detail: the joint-profit-maximising logic of a perfect cartel, the problem of allocating output and dividing profit, the conditions that make collusion easy or fragile, the strong incentive of each member to cheat, and the legal and welfare consequences that lead most countries to prohibit cartels.

Meaning and Forms of Collusion

Collusion means co-ordinated behaviour among rival firms intended to restrict competition between them. It takes three main forms. A formal cartel is an explicit agreement, often written, in which independent firms fix a common price, allocate output quotas or market territories, and sometimes operate a central selling agency that sells all members' output and distributes the proceeds — OPEC in the oil market is the best-known international example of a quota cartel, and the pre-war European industrial cartels were classic domestic ones. Gentlemen's agreements and trade-association arrangements are softer: prices, discounts and terms are co-ordinated through meetings and information exchange without a binding contract. Tacit collusion involves no communication that can be proved: firms simply learn, through repeated experience, that price cuts are matched and therefore pointless, and a pattern of parallel pricing — often organised around a price leader whose changes others follow — settles

over the market. The economic effects of the three forms can be similar, which is why competition authorities look at outcomes and facilitating practices as well as at signed agreements, although the law usually penalises explicit collusion most severely because it is provable and deliberate.

The Perfect Cartel: Joint Profit Maximisation

The theory of a perfect cartel treats the industry as if it were a single multi-plant monopoly. The cartel first constructs the industry demand curve and the corresponding marginal revenue curve. It then obtains its marginal cost curve by horizontally summing the marginal cost curves of all member firms and plants, exactly as a monopolist would combine the costs of its own plants. Joint profit is maximised where industry marginal revenue equals industry marginal cost; that equality fixes total cartel output, and the price is read off the industry demand curve at that output. The cartel then allocates the total output among members so that the marginal cost of the last unit is equal in every plant — production is shifted toward the more efficient members until no reallocation could produce the same total output more cheaply. Profits are distributed according to the cartel agreement, which need not match the output shares: efficient firms may be paid to expand while inefficient firms are compensated for restricting output, which is why quota bargains inside cartels are often as much about side payments and market shares as about the price itself.

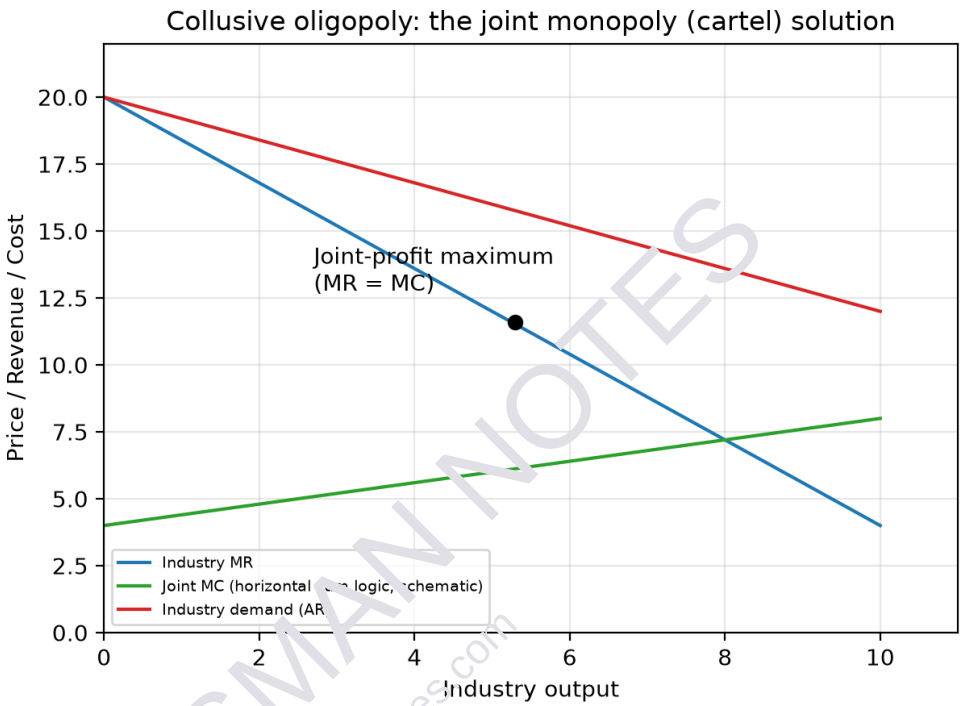


Figure 3: A perfect cartel maximises joint profit where industry MR equals the combined MC, charges the price shown by industry demand, and allocates output among members at equal marginal cost.

Output Quotas and the Problem of Agreement

Reaching the joint maximum on paper is much easier than sustaining it in a negotiating room, because members' interests diverge at every step. Efficient and inefficient firms disagree about quotas: the low-cost firm wants a large share and a price nearer its own costs, while the high-cost firm survives only under a high, protective price and a guaranteed share. Firms with excess capacity press for larger quotas; firms that have invested in new plant demand recognition of it; and every firm suspects the others of understating costs or overstating difficulties. Quota systems based on historical market shares freeze yesterday's pattern of efficiency in place, while systems based on capacity rewards firms for building capacity they do not need. Disagreement about the division of profit is in fact the commonest cause of cartel collapse, even before cheating begins, and cartels often settle for a simpler, less than fully maximising

arrangement — a common price with traditional shares — because it is negotiable and monitorable, sacrificing some joint profit for stability.

The Incentive to Cheat and the Instability of Cartels

Even a successfully negotiated cartel contains the seed of its own destruction: at the cartel price, price exceeds every member's marginal cost, so each member can increase its own profit by secretly selling a little more — shading the price through hidden discounts, better credit terms, free delivery or quality upgrades — while the others keep to their quotas. If one firm cheats, it gains at the others' expense; if all cheat, output expands, price collapses toward the competitive level, and the cartel's purpose is defeated. Game theory expresses this as a prisoners' dilemma: cheating is each firm's individually rational strategy even though all would be better off if none cheated. Cartels survive only where cheating can be detected quickly and punished credibly — through transparent prices, published sales data, homogeneous products that make secret deals visible, frequent meetings, and the threat of a price war or expulsion. This is why real cartels invest heavily in auditing members' sales and why they break down so dramatically when monitoring fails or when a desperate member, facing bankruptcy or holding unsold stocks, decides that survival today matters more than discipline tomorrow.

Conditions That Favour or Hinder Collusion

Collusion is easiest where the number of firms is small, so that negotiation and monitoring are manageable; where the product is homogeneous, so there is one price to fix and deviations are easy to spot; where firms have similar costs and sizes, so a common policy suits them all; where entry barriers are high, so the cartel's profits are not competed away by outsiders; where demand is stable and predictable, so a fall in sales can be correctly attributed to cheating rather

than to the market; and where buyers are numerous and small, with no large customer able to play firms against one another by offering a huge order for a secret discount. Collusion is hardest in the opposite conditions: many firms, differentiated products with innumerable prices and terms, volatile demand, powerful buyers, secret bidding for large contracts, and rapid technical change that continually reshuffles firms' cost positions. Antitrust enforcement works precisely by worsening these conditions for cartels — leniency programmes that reward the first firm to confess are a deliberate device for destroying the mutual trust on which collusion depends, and they have proved to be among the most effective anti-cartel instruments ever introduced.

Table 4: Conditions affecting the stability of collusion

Favours collusion	Hinders collusion
Few, similar-sized firms	Many firms of very different sizes
Homogeneous product, transparent prices	Differentiated products, secret discounts and bidding
Stable, predictable demand	Volatile demand and rapid technical change
High entry barriers	Easy entry by outsiders
Small, fragmented buyers	Large buyers who can split orders and demand concessions

Price Leadership as a Collusive Device

Where explicit agreement is illegal or risky, collusive oligopolies often settle on price leadership. In dominant-firm price leadership, the largest firm sets the price and the smaller firms follow it, taking the leader's price as given and supplying what they can sell at it; the leader, knowing this, sets the price that maximises its own profit on the demand remaining after the fringe's supply. In barometric price leadership, one firm — not necessarily the largest — acts as the

industry's barometer, announcing price changes that reflect cost and demand conditions, which others follow because experience has taught them that the barometer reads conditions correctly. Low-cost price leadership occurs when the most efficient firm sets a price the others must match. Price leadership softens price competition without any provable agreement, and prices in such industries tend to be rigid, changing only when a clear cost shock, such as an industry-wide wage settlement or raw-material increase, gives everyone a simultaneous, legitimate reason to move — the rigidity acting as a further signal that no one is breaking ranks.

Welfare Effects and the Legal Position

The welfare judgement on collusion is the clearest in the whole of oligopoly theory. A successful cartel reproduces the results of monopoly: price is raised above marginal cost, output is restricted below the competitive level, resources are misallocated, and a deadweight loss of consumer and total surplus arises, together with a transfer from consumers to producers. Collusion also dulls the incentives to cut costs and innovate, because sheltered profits can be earned without efficiency — the 'quiet life' that Hicks remarked monopolists enjoy. For these reasons, explicit cartels are illegal per se in the United States under the Sherman Act and prohibited in the European Union and in Pakistan under its Competition Act, with heavy fines and, in some jurisdictions, criminal penalties; competition commissions actively investigate bid-rigging, price-fixing and market-sharing in industries from cement and pharmaceuticals to public procurement. The main exceptions recognised in policy are arrangements defended as generating efficiencies that outweigh the harm — some research-and-development co-operation and standard-setting — and international commodity arrangements such as OPEC, which lie beyond any single country's jurisdiction and illustrate

both the power and the notorious instability of cartels, as members repeatedly exceed their quotas when prices are attractive.

Conclusion

To conclude, collusive oligopoly is the attempt by a few interdependent firms to escape the discipline of competition by co-ordinating price and output, most perfectly in a cartel that maximises joint profit where industry marginal revenue equals combined marginal cost and allocates production among members at equal marginal cost. Its analysis reveals a permanent tension: the joint gains from collusion are real, but so is each member's incentive to cheat, so collusion flourishes only under identifiable conditions of fewness, homogeneity, transparency, stability and blocked entry, and collapses when monitoring or trust fails. Because a successful cartel inflicts the same deadweight loss as monopoly, modern competition law prohibits it and deliberately engineers distrust among would-be colluders. Price leadership remains as the milder, tacit form of the same impulse, and understanding both forms is essential to understanding how real oligopolistic industries actually behave.

Question 5: Joan Robinson's Contribution to Monopolistic Competition (20 Marks)

Q. 5 Explain Joan Robinson's contribution to Monopolistic Competition.

Introduction

Joan Robinson (1903–1983), the Cambridge economist, published The Economics of Imperfect Competition in 1933, in the same year that Edward Chamberlin published The Theory of Monopolistic Competition in the United States. Between them, the two books destroyed the

old two-box world of economics, in which a market was either perfectly competitive or monopolised, and replaced it with a spectrum of intermediate market forms in which most real firms actually live. Robinson's book was the more rigorous and the more geometrical of the two, and several of its analytical tools — above all the marginal revenue curve in its modern form — became permanent equipment of every economist. This answer explains her contribution in detail: the problem she set out to solve, her construction of the demand and marginal revenue apparatus, her analysis of the firm selling in an imperfect market, her treatment of price discrimination, her celebrated discussion of the exploitation of labour, and the lasting influence and criticisms of her work.

The Setting: Breaking the Competition–Monopoly Dichotomy

Before 1933, textbook theory oscillated between two unrealistic poles. Under perfect competition the firm faced a horizontal demand curve and price was outside its control; under monopoly a single firm faced the whole market demand. Real firms — manufacturers of branded goods, local shops, professional services — fitted neither box: each had some power over its price because its product was distinguishable and its customers not perfectly mobile, yet each faced rivals whose competition limited that power. Robinson's starting insight, influenced by Piero Sraffa's famous 1926 article arguing that firms typically operate with falling costs and limited markets, was that the downward-sloping demand curve of the individual firm should be the general case of theory, with perfect competition and pure monopoly as limiting special cases. That single move — making the firm's demand curve slope downward and asking what follows — reorganised price theory, and it is the foundation on which the whole analysis

of monopolistic competition, including the tangency equilibrium and excess capacity discussed in Question 3, is built.

The Technical Apparatus: Average and Marginal Revenue

Robinson's most concrete and lasting contribution is technical. She systematically developed the relationship between the average revenue curve – the demand curve facing the firm — and the marginal revenue curve derived from it, showing that whenever the demand curve slopes downward, marginal revenue lies below average revenue at every output, because selling one more unit requires lowering the price on all units, not only on the extra one. From this she derived the connection with elasticity that every student still learns: marginal revenue equals price multiplied by (e − 1)/e, where e is the elasticity of demand, so that MR is positive when demand is elastic, zero when elasticity is unity and negative when demand is inelastic. This formula explains at a stroke why a firm with any market power never willingly operates on the inelastic portion of its demand curve. She likewise gave marginal cost its central organising role through the equilibrium condition MR = MC. It is no exaggeration to say that the diagrams in a modern textbook chapter on the firm — AR, MR, AC and MC curves in one picture — are Robinson's apparatus, inherited almost unchanged.

Equilibrium of the Firm in an Imperfect Market

With this apparatus Robinson analysed the firm in an imperfect market in a way that unified the older theories. The firm maximises profit by producing the output at which marginal revenue equals marginal cost and charging the price shown by its average revenue curve at that output. Whether it earns supernormal profit, normal profit or a loss depends on the position of its average cost curve relative to average revenue at that output — precisely the analysis still

taught today. Where Chamberlin emphasised the group of firms and the entry-driven tangency solution, Robinson worked more deeply on the individual firm's geometry, including cases of falling average cost in which competitive equilibrium in the old sense could not exist at all, since a price equal to marginal cost would leave the firm unable to cover its costs. She also analysed how a tax or a change in costs would shift the equilibrium, an early and model exercise in comparative statics. Her treatment made clear that 'imperfection' is not a minor blemish on competition but a different system with its own logic, in which price exceeds marginal cost as a normal, permanent condition.

Price Discrimination

Robinson gave the theory of price discrimination its classical formulation, building on Pigou's classification. She set out the three conditions for discrimination to be possible and profitable: the seller must have some monopoly power; markets must be separable, with no possibility of buyers reselling from the cheap market to the dear one; and the elasticity of demand must differ between markets. Her rule for a discriminating monopolist — allocate output so that marginal revenue is equal in all markets and equal to the common marginal cost, charging the higher price in the market with the less elastic demand — remains the standard solution taught today, and she illustrated it with the railway rates and professional fees that were the stock examples of the time. She also asked the welfare question that still frames the debate: discrimination can increase total output, for example by making it profitable to serve a poor market that a uniform monopoly price would exclude altogether, yet it also transfers surplus from consumers to the seller and can misallocate a given output between buyers. Modern practices — student discounts, off-peak fares, differential pricing of medicines between

countries and personalised online prices — are all analysed with Robinson's conditions and rules.

Imperfect Competition and the Exploitation of Labour

Robinson extended the analysis from product markets to factor markets in one of her most famous contributions: the economic definition of the exploitation of labour. Where an employer has monopsony power — market power as a buyer of labour — or where the firm sells in an imperfect product market, the wage a worker receives can fall below the value of his marginal revenue product. Robinson defined a worker as exploited when he is paid less than the value of his marginal product under the relevant conditions, and she showed two distinct sources. Monopolistic exploitation arises because the firm, facing a downward-sloping demand curve, hires up to the point where wages equal marginal revenue product, which is less than the value of the marginal product. Monopsonistic exploitation arises because the employer, facing an upward-sloping supply of labour, hires where marginal revenue product equals the marginal cost of labour, which lies above the wage. The distinction matters practically: a trade union or a minimum wage can, in the monopsony case, raise both wages and employment — a result directly opposed to the simple competitive story and one that continues to influence minimum-wage debates, in which Robinson's analysis is still cited.

Table 5: Two sources of exploitation in Robinson's analysis

Type	Cause	Result for the worker
Monopolistic exploitation	Firm sells in an imperfect (downward-sloping demand) market	Wage = marginal revenue product, which is below the value of the marginal product
Monopsonistic exploitation	Employer is a powerful buyer facing upward-sloping labour	Wage is below marginal revenue product, set against the

	supply	higher marginal cost of labour

Other Contributions and the Cambridge Context

Several further contributions belong in a complete account. Robinson analysed buying-side market power (monopsony and oligopsony) with the same rigour as selling power, restoring symmetry to a theory that had concentrated on sellers, the bilateral monopoly problem of Question 2 is unthinkable without this two-sided view. Her work on the classification of market imperfection — through the concept she called the degree of imperfection, related to the elasticity of the firm's demand curve — gave empirical content to the theory: the less elastic the demand facing the firm, the greater its market power and the larger the gap between price and marginal cost, an idea that survives in the modern Lerner index of monopoly power. Intellectually, the book also stands at the gateway of her later career: the young author of 1933 became the leading figure of the post-Keynesian Cambridge school, a fierce critic of the neoclassical theory she had helped to perfect, and a central combatant in the Cambridge capital controversies. That later self-criticism does not diminish the 1933 book; it underlines how completely its tools had been absorbed into the orthodoxy she later challenged.

Criticisms and Limitations

Balanced judgement requires the criticisms. Chamberlin objected that Robinson's approach still treated the firm's demand curve as if it were independent, whereas in a group of competing firms every firm's curve shifts when rivals react — his charge that her model was 'monopoly analysis with a competitive flavour', and the ensuing dispute over priority and terminology between 'imperfect' and 'monopolistic' competition, occupied journals for years; in modern

teaching the two are used together, Chamberlin for the group equilibrium and entry, Robinson for the firm's apparatus and factor markets. Her analysis is also static: it compares equilibria but says little about the process of rivalry, innovation and product development through which competition in differentiated markets actually proceeds, a gap later filled partly by Austrian and Schumpeterian criticisms and by modern game theory. Finally, parts of the book are forbiddingly geometrical, and some of its constructions, such as curves drawn on shifting assumptions about rivals' reactions, can mislead if the ceteris paribus conditions are forgotten. None of this removes the central achievement: the analytical core of imperfect competition, as now universally taught, is substantially hers.

Conclusion

To conclude, Joan Robinson's contribution to monopolistic (imperfect) competition was to make the downward-sloping firm demand curve the general case of price theory, to forge the average and marginal revenue apparatus and the MR = MC rule in their modern form, to explain equilibrium, profits and losses of the firm in imperfect markets, to give price discrimination its classical conditions and allocation rule, and to define and dissect the exploitation of labour under monopolistic and monopsonistic power. Published alongside Chamberlin's work in 1933, her Economics of Imperfect Competition moved real-world markets — branded, differentiated, imperfect — from the footnotes of theory to its centre. Every diagram used elsewhere in this assignment, from the competitive firm's cost curves to excess capacity and collusive pricing, is drawn with tools that she helped to create.