

H2 Economics Chapter 3 Notes

This document is intended as a summary of the points provided in the NJC 2024 H2 Economics Chapter 3 Seminar Notes. It is to be used as a **study guide** to provide a **structure** for studying the Chapter 3 Notes.

1 Introduction to Firms

Before we begin our study, let us clarify the focus of our study. We are studying individual firms' decisions and strategies individually. We are NOT studying the decisions and strategies of firms as a whole industry.

1.1 Key Terms and Concepts

Chapter 3 is the most difficult chapter for students to completely understand because the scope of content is extremely large. Hence, the approach to studying this chapter is to first master the big idea of the chapter, before understanding each subsection in great detail. It is highly advised that you spend a large amount of time to properly and fully understand the basic concepts covered in this section before moving on to any other sections.

A **Firm** is a business organisation that hires factors of production, combines them in a productive process to create and sell the resulting output for a profit.

A **Plant** (or factory) refers to the geographical location where the actual production is carried out. A firm may own one or several production plants.

An **Industry** refers to a group of firms producing similar goods or services. Firms within the same industry are usually competitors in the same market.

Profit (π) is the difference between total revenue (TR) and total cost (TC) of production

Short Run is defined as a period over which at least one factor of production is fixed. In the short run, output can only be adjusted by changing the quantities of variable factors.

Long Run (LR) refers to a period of time extensive enough to allow the firm to change all resources employed within the confines of a given state of technology.

Important Things to Note and Key Clarifications	
Firm	<u>What is a firm?</u> Typically, a firm is seen as a separate entity from its shareholders, i.e. those who own the firm. A firm is typically seen more as an intangible entity which is defined by (1) its profit-maximising objective and (2) its ability to hire factor inputs to produce goods and services. <u>Separation between employees and shareholders:</u> • Shareholders are the owners of a firm. Typically, they own a part of the firm by investing in shares of the firm. When the firm profits, the shareholders receive part of the profits in terms of dividends. Shareholders typically make the major decisions of the firm (electing board members, deciding on mergers and acquisitions, investment decisions, hiring and firing the CEO) • Employees typically serve the role of labour in firms. They are what keeps the firm operating, and they are rewarded by the firm in terms of wages. Firms can also hire paid managers like CEOs and CFOs which are in charge of the operations of the firm in terms of pricing and non-pricing strategies, hiring or firing workers, ordering stock, etc. These managers are the ones who provide entrepreneurial resources that the firm needs. However, entrepreneurial resources can also be provided by shareholders as they are in charge of major decisions. Traditionally, we would consider managers to also be profit maximising, as they are generally paid higher salaries if the firm's profits increase. However, since they are motivated by salaries and not profits directly, this may not always be true. • However, in smaller firms, the owner can also operate the firm. For example, in hawker centres, hawker stall owners operate the firm. These smaller firms may be known as sole proprietorships or partnerships, while the larger firms may be known as corporations or companies.

Plant	<u>What is a Plant?</u> Plants refer to factories and buildings where the production of the goods and services actually occur. It is what people would usually call ‘capital goods’. They contain labour and other machinery which are part of the production process. A firm is seen as more of the entity that makes decisions, while the plant is seen as the entity which carries out these decisions.
Industry	<u>What is an Industry?</u> An industry is typically what students have understood as the concept of ‘market supply’. An industry typically addresses one particular want or need of consumers via the selling of slightly differentiated products. For example, in the energy industry, the firms would be companies producing or distributing energy (e.g., oil, gas, renewable energy sources). <u>Competition Between Firms</u> Firms in an industry are usually competitors of each other. Competition is the rivalry between firms seeking to achieve a particular objective, typically to maximise profits, market share, or customer base. In a market context, competition arises because multiple firms or sellers offer similar or substitute products and services either at different prices or different quality. We say that goods that firms sell are price competitive or quality competitive when they are goods which consumers are more willing and/or able to buy. An increase in competition is usually associated with the concept of an increase in the number of firms in the industry. This is because with new firms, some consumers will typically switch over to these new firms because individual consumers have different tastes and preferences, and some goods and services appeal to some consumers’ specific tastes and preferences better than others.

Profit	Total Revenue (TR) is a firm's total earnings from the sale of its output for a specific period of time. Total revenue is calculated as the average price per unit (P) multiplied by the number of units sold (Q) Total Costs (TC) is a firm's total opportunity cost from the production of its output for a specific period of time. <u>Supernormal, Normal and Subnormal Profits</u> • If total revenue and total costs are equal, economic profit is zero. The firm is said to be earning normal profits. This is deemed to be the minimum amount of profits the firm must make in order to continue operations in the long-run. • A supernormal profit (i.e. positive economic profit) indicates that the total revenue accruing to the firm from sale of its goods and services exceeds the total costs of production. • A subnormal profit or loss (i.e. negative economic profit) occurs when total revenue falls short of total production costs. The firm may incur subnormal profit in the short run and choose to continue production, but it will exit the industry if it continues to make subnormal profit in the long-run. <u>Economic Profit VS Accounting Profit</u> The total costs of production to the economist are the opportunity costs of using resources in the production of a product – a measure of the sacrifice involved in using factors of production to produce that product. To produce a product, scarce resources must be diverted from other productive activities. The real costs of production to an economist are hence the sum of implicit and explicit costs. • Implicit costs are the opportunity costs of using resources that are <u>already owned</u> by the firm. The firm does not have to pay out money to use them. For example, if a firm owns some buildings, the opportunity cost of using them is the rent it could have received by letting out to another firm.
----------------------	--

	• Explicit costs are the opportunity costs of using resources that are <u>not already owned</u> by the firm. The firm has to pay money to use them. For example, if a firm buys \$1,000 worth of materials and direct payment is made to the supplier, the opportunity cost is \$1,000 because the \$1,000 could have been spent on something else. • Put simply, implicit costs are the next best use of owned <u>resources</u> that are forgone while explicit costs are the next best use of <u>funds</u> (not 'money' – too vague) forgone. • Accountants consider only the explicit costs incurred during production but economists consider both explicit as well as implicit costs hence typically, accounting profit is higher than economic profit.
Short Run	These are defined in terms of how many factors are variable, <u>not calendar time</u>. As long as one factor of production is fixed, the firm is operating in the short run. Only when <u>all</u> factors of production are variable, the firm is operating in the long run. The state of technology is assumed to be a constant in both the short and long run. • A Fixed Factor/Input is a factor of production whose quantities cannot be changed irrespective of the level of output. Typically, these would be land resources and capital resources. But this is not always true (the airline industry can just send out aeroplanes from its warehouse.)
Long Run	• A Variable Factor/Input is a factor of production whose quantities can be changed such that more of the factor can be employed in order to increase the level of output. Typically, these would include labour resources. However, if we have to factor in trade unions which bargain for shorter working hours, then this may not be true. • A fixed factor to one firm can be a variable factor to another firm. For example, a table can be a fixed factor to a restaurant, but to a furniture retailer, it is a variable factor because it can simply be retrieved from its stock of tables.

1.2 The Marginalist Principle

Marginal Revenue (MR) is the addition to total revenue from the sale of an additional unit of good/service produced.

Marginal Cost (MC) is the addition to total cost from the production of an additional unit of good/service.

If $MR > MC$, the additional revenue earned from producing an additional unit of output is higher than the additional cost of producing that unit of output. Hence, a profit maximising firm will continue to produce as long as $MR > MC$ since the firm will make a gain in profit by producing that unit of output. If $MR < MC$, the additional revenue earned from producing an additional unit of output is lower than the additional cost of producing that unit of output. Hence, a profit maximising firm will reduce production as long as $MC > MR$ since the firm's profit will make a loss in profits by producing that unit of output.

As such, the profit maximising firm will produce at the output level where marginal revenue (MR) is equal to marginal cost (MC), known as the profit-maximising level of output. At this level of output, the maximum price the firm can charge is determined by the demand curve or AR curve. Hence, the firm will set its price at P_1 . The profit of the firm is given by the difference between its total revenue and total cost. This is shown by the area C_1P_1ab in Fig 1.

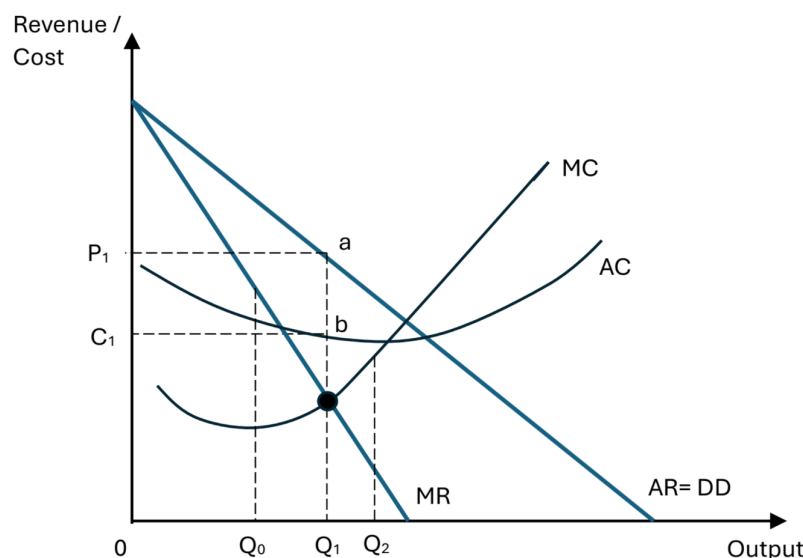


Fig 1: Profit Maximising Level of Output

2 Short Run Costs Curves

The ***Law of Diminishing Marginal Returns*** states that when increasing amounts of a variable factor are used with a given amount of a fixed factor, there will come a point when each extra unit of the variable factor will produce less extra output than the previous unit.

Consider a bakery which produces bread on a large scale. It requires a factory, as well as labour to bake the bread. In the short run, as there is only one factory, he can increase output of bread only by adding more bakers to the factory. The bakers are assumed to be identical in ability and aptitude (i.e. homogeneous). Technology is also assumed to be constant in the short run.

- It is important to understand that the law of diminishing marginal returns is a short run law. *Ceteris paribus*, returns are only diminishing because there are fixed factors of production which we cannot increase the input of together with the increase in input of variable factors of production, as I will illustrate in the following paragraphs. That means in the long run, at this point, if we increase all the factors of production by the same proportions, the marginal cost of production should theoretically not change. (This is not true. The long run marginal cost does change with output, but it is not as a result of the existence of fixed factors of production because the long run is literally defined to be a period where there are no fixed factors of production.)

Initially, with only one baker working, factory space is under-utilised. From the first to the fourth baker, each additional baker contributes a greater additional level of output compared to the previous unit of input. This is because there is greater efficiency in the use of the factory space for productive purposes.

However, as more bakers are added, the factory became congested and mobility became restricted. At some point, an additional baker will contribute a smaller additional amount of output than the previous baker due to the discomfort of working within a constrained space or having to wait to use the oven. This is when the bakery experiences diminishing returns.

- Notice that the diminishing returns only occur because the factory cannot be expanded.
- Output itself is not decreasing, but rather it is the addition to output that is decreasing.

Starting from the 5th baker, each additional baker contributes less to total output than the previous one. Eventually the 7th baker is redundant as he ceases to add to output. Total output may even fall as more variable factors are added (e.g. the 8th baker) as these extra workers may even hinder the work of the other bakers by causing arguments, miscommunications, etc.

Therefore, we observe that output can be increased by adding more and more variable factors to the fixed factor but the phenomenon of increasing output does not last forever. Eventually, over-utilisation of the fixed factor in the short-run, will result in diminishing returns from additional units of variable factor.¹

The law of diminishing returns can be used to explain the shape of the marginal cost curve, though this is not in the syllabus. Nonetheless, it is advised that you should remember this as it is useful in helping to gain a better understanding of this chapter.

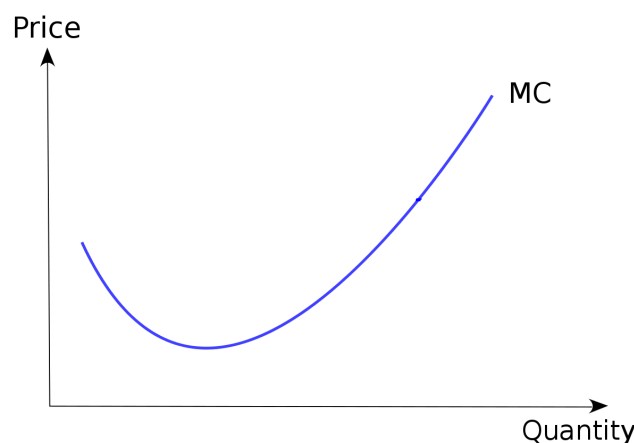


Fig 2: Marginal Cost Curve

If each additional unit of labour produces increasing amounts of output initially, assuming each unit of labour is paid the same wages, then to produce the same additional amount of output, less labour is required. Hence, the marginal cost decreases, illustrated by the downward sloping portion of the marginal cost curve. However, past a certain point, each additional unit of labour produces decreasing amounts of output, so to produce the same additional unit of output, a lot more labour is required. Hence, the marginal cost increases, illustrated by the upward sloping portion of the marginal cost curve.

¹ FYI: The ratio of the amount of variable factor to the fixed factor varies as production level changes. Hence, production in this case is said to take place under variable proportions. Production under variable proportions will eventually result in diminishing returns as more and more variable factors are added to a fixed factor. Hence, the Law of Diminishing Returns is also known as the Law of Variable (Factor) Proportions.

Total Fixed Costs (TFC) are costs that do not vary with the amount of output produced.

- **Average Fixed Cost (AFC)** is the fixed costs per unit of output.

Total Variable Costs (TVC) are costs that vary with the amount of output produced.

- **Average Variable Cost (AVC)** is the variable costs per unit of output.

Total Cost (TC) is the sum of total fixed costs and total variable costs at a specific level of output.

- **Average Cost (AC)** is the total cost per unit of output.

In the short run, some of a firm's factors of production, such as the size of the plant, will be fixed. Other factors of production will be variable. Thus, in the short run we can break the firm's costs into two categories: fixed costs and variable costs.

Fixed Costs	Variable Costs
Total Fixed Costs (TFC) are costs that do not vary with the level of output of the firm. Fixed costs are associated with fixed factors, therefore they exist only in the short run. There are no fixed costs in the long run.	Total Variable Costs (TVC) are costs that vary with the level of output of the firm. Such costs are incurred from the hiring and purchasing of variable factors. Thus, variable costs increase as output increases.
Fixed costs will be <u>incurred</u> even if the firm <u>ceases production</u> . Such costs are often referred to as overheads and are beyond management's control in the short run.	Variable costs will be <u>zero</u> if the firm <u>ceases production</u> . Management can control variable costs in the short run by changing the level of output.
Fire insurance, property taxes, rentals, Depreciation of machines due to obsolescence and ageing (not due to usage), Interest payments (e.g. for instalment contracts) on loans for the purchase of capital goods (e.g. factory and equipment)	Costs of raw materials and fuel, Wages of production workers

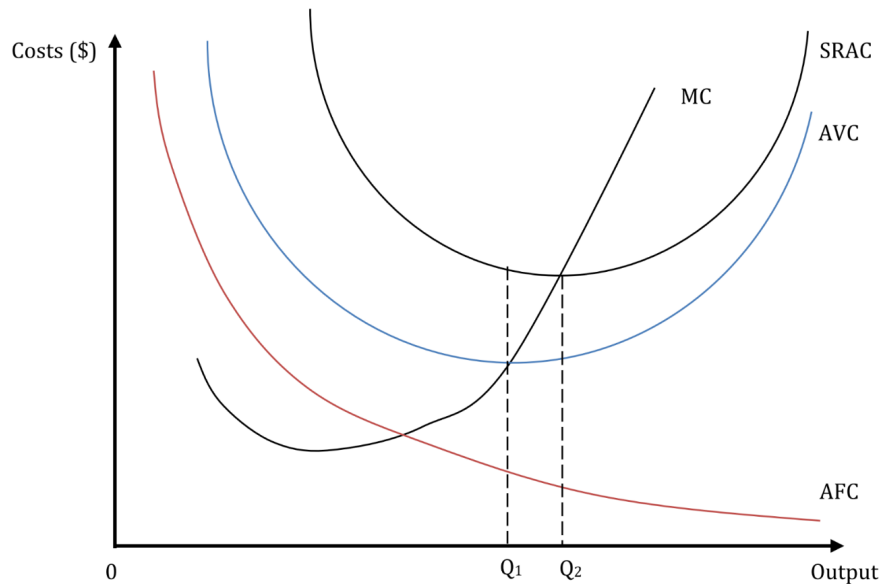


Fig 3: Short Run Cost Curves

Key Features of the Diagram

1. Average Fixed Cost (AFC) decreases at a decreasing rate and approaches zero. This is because total fixed cost is a constant and as a result, as the same cost is spread over an increasingly large level of output, the average cost becomes smaller and smaller, approaching zero. As such, MC is also the additional variable cost per additional unit of output.
2. Short-run average cost (SRAC) and average variable cost (AVC) both decrease first, then increase. By the law of diminishing marginal returns, initially, the addition to output will be greater than the addition to cost. Hence, average cost falls because the total variable costs are spread over an increasing level of output. However, after the addition to output becomes lower than the addition to cost, eventually, the total variable costs will be spread over a relatively smaller increase in output. Thus, the average costs will rise later on. It's a mathematical exercise.
3. SRAC approaches AVC. This is because $SRAC = AFC + AVC$. If $AFC \rightarrow 0$, then $SRAC \rightarrow AVC$.
4. Marginal Cost (MC) cuts SRAC and AVC at their minimum point. When the addition to costs is lower than the average cost, then the average cost will be falling. When the addition to costs is higher than the average cost, then the average cost will be rising. Only when the addition to costs is equal to average cost, the average cost will not change. This occurs at the minimum point of the AVC and SRAC.

3 Long Run Costs and Economies of Scale

Returns to Scale refers to the behaviour of production or returns when all the factors of production are increased or decreased simultaneously in the same ratio.

- Increasing returns to scale is said to occur when a given percentage increase in all inputs by the same proportion leads to a larger percentage increase in output.
- Constant returns to scale is said to occur when a given percentage increase in all inputs by the same proportion leads to the same percentage increase in output.
- Decreasing returns to scale is said to occur when a given percentage increase in all inputs by the same proportion leads to a smaller percentage increase in output.

We had previously claimed that the additional returns from each additional unit of input will decrease because of the existence of fixed factors, as there will come a point where the fixed factor is over-utilised. Does that mean that when there are no fixed factors, the returns from each additional unit of input will be constant?

Surprisingly, empirical evidence tells us the answer is no. In fact, in the long run, when all the factors of production are increased in the same ratio, the increase in output is not the same. This can no longer be explained by the law of diminishing marginal returns because there are no longer any fixed factors in the long run. Hence, we turn to the concept of returns to scale.

Why do we increase the factors of production by the same ratio?

Firms are profit-maximisers, and they will tend to produce goods and services with the least cost combination of inputs. Hence, we assume that in the long run, firms have discovered this particular ratio of inputs to be the least cost combination of inputs, and hence will not change this ratio of inputs.

Internal Economies of Scale refers to the reduction in long run average costs of production as a result of the expansion of the scale of production of a particular output/product for an individual firm.

External Economies of Scale occur when a firm benefits from lower average costs as a result of the whole industry growing in size, independent of the firm's decisions.

Be very careful when identifying internal economies of scale and external economies of scale.

Internal Economies of Scale	External Economies of Scale
An individual firm's cost of production falls because it increases its scale of production. This entails a firm expanding its plant size, hiring more employees, market share, number of shareholders, capital assets, etc.	An individual firm's cost of production falls because the industry expands. This usually is shown when the demand for goods and services of a particular type increases, or when there are more facilities and auxiliary services available for firms' usage.
This is illustrated by a point movement along a firm's LRAC.	This is illustrated by a point shift of a firm's LRAC.

Additionally, note that you should preferentially pick technical economies of scale when the firm's production is largely manufacturing in nature, while you should preferentially pick firm economies of scale when the firm's production is largely service-related in nature. This does not mean that you should solely pick technical economies of scale or firm economies of scale. If you are required to write both, then you should do a 2 + 1 approach (e.g. if the firm's production is largely manufacturing related, 2 technical economies of scale + 1 firm economies of scale). Ultimately, the economies of scale chosen must be contextualised to the firm's production.

3.1 Internal Economies of Scale

Due to internal economies of scale, the increase in the firm's cost of production due to its larger scale of production is thus spread across an even larger level of output. Noting that average cost is equal to total cost divided by the level of output, the firm long run average cost will be lowered from C_1 to C_2 illustrated by a movement from point A to point B along its long run average cost curve (LRAC) when the firm output increases from Q_1 to Q_2 as illustrated in Fig 4.

- This paragraph should be written as set up if the focus of the essay is largely on economies of scale. This will save you tremendous amounts of time as you do not have to rewrite the analysis for each IEOS paragraph.
- **Minimum Efficient Scale (MES)** refers to the lowest level of output at which the long run average cost is at its minimum.

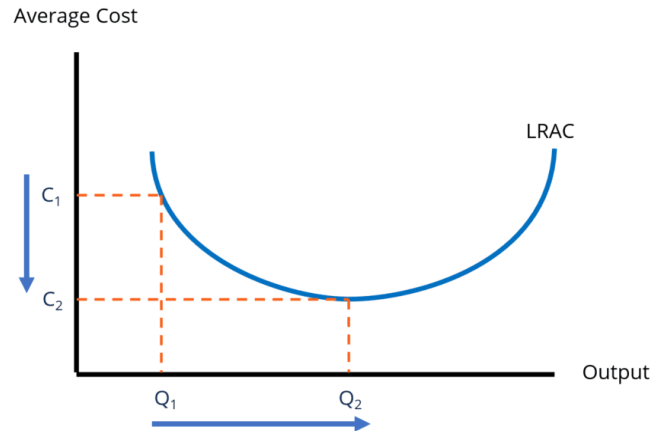


Fig 4: Internal Economies of scale

Specialisation and Division of Labour (Technical IEOS):

A complex manufacturing production process can be broken down into simpler specialised multi-stages. Workers can therefore be allocated to their most efficient specialised stages. Specialisation and division of labour raises productivity because the workers can produce with greater speed and dexterity, and because lesser time is lost from task switching. Training cost is also reduced as with specialisation because workers only need to be trained in their specialised jobs. A lot more output can therefore be produced given a smaller but proportionate increase of inputs.

Higher Utilisation of Indivisible, Highly Efficient Machinery (Technical IEOS):

Certain types of machines are only available in a fixed productive capacity. Firms with larger scale of production can better utilise these machineries via, for example, working on 24 hours shift as compared to firms with smaller scale of production where the machines may be under-utilised. The percentage increase in the cost of operating the machines such as labour and overhead costs can be insignificant compared to the level of output that can be produced due to the high productivity of the machines when inputs are increased by the same proportion.

Marketing Economies of Scale (Firm IEOS):

The firm can purchase more factor inputs thus giving the firm a greater bargaining power to negotiate for higher bulk purchase discount. Moreover, the firm can also enjoy lower per unit distribution costs because, for example, it can use a slightly more expensive but more cost efficient large truck to deliver its larger amount of goods to its distributors instead of a less cost efficient smaller van. Also, with the larger scale of production, high advertising cost is shared across the larger amount of output produced by increasing inputs by the same proportion.

Organisational Economies of Scale (Firm IEOS):

Rationalisation can take place when the use of a centralised administration makes possible the reduction of duplicated administrative tasks such as the reduction in the combined size of the finance department. For example, despite the relatively larger manpower, procurement and output of the newly merged firm, only one human resource department handling wage matters or one accounting department is required instead of two.

Financial Economies of Scale (Firm IEOS):

The firm may accumulate and thus will have more assets which can be offered as collateral to banks for cheaper loans. Banks are more willing to provide lower interest rates to larger firms because of their lower risk as their high collateral value would reduce the probability of them defaulting the repayments of the loan. A small reduction in the interest rate would result in huge cost savings for loans of large amounts. In addition, the larger firm can also raise funds by issuing shares which increases its cost of production insignificantly as compared to getting a bank loan with high interest payment.

Managerial Economies of Scale (Firm IEOS):

A complex administrative process such as finance process can be broken down into smaller processes such as budgeting and actual payroll matters which are of very different nature and assigned to specialised administrative departments headed by specialised finance trained managers. The specialised department will therefore be better managed, for example, via better budget planning and can better provide specialised in-house training for their specialised administrative personnel thus would improve productivity and communication. As a result, a lot more output can be processed using the same amount or even lesser time and manpower

Risk-Bearing Economies of Scale (Firm IEOS):

The firm can diversify its risk especially when it operates in several markets. For example, in a market where generally consumers are trend-setters and in a market where consumers are generally trend-followers. When there is a change in tastes and preferences away from its product in the trend-setting market resulting in a fall in demand of the product, the large firm can transfer its excess stock of the outdated goods in the trend-setting market to sell in the trend-following market as the goods can still be considered trendy in the trend-following market. Large firms are therefore able to minimise wastage cost and save on storage cost as compared to a small firm which cannot readily divert its stock due to its limited distribution networks.

OR

The firm can diversify its risk especially when it operates in several markets. For example, the firm can produce products of different income elasticity for countries facing different economic conditions to lower its cost of uncertainty. It can produce a relatively more luxurious product to cater to countries having economic boom and it can produce a relatively more inferior product to cater to countries having economic recession. Doing so allows the firm to offset its fall in revenue in some countries by the increase in revenues in other countries during times of global economic boom or recession. The fall in revenue and hence profits in one market can be offset by those earned in the other markets resulting in the lowering of its cost of uncertainty, for example, via cheaper firm insurance risk premium and cheaper bank loan.

3.2 Internal Diseconomies of Scale

Loss of Control

In the real world, there is imperfect information thus the firm may not know whether it is producing below or above its minimum efficient scale. If it is already producing near or at its minimum efficient scale, it may soon suffer internal diseconomies of scale, thus higher average cost of production as it expands its scale of production further. In an oversized firm, there might be loss of control as it employs more and more workers. It may become difficult to monitor the work performance of the large number of workers and ensure they are fairly treated. Motivation to work will be affected. Besides, problems such as lack of motivation and worker absenteeism may lead to lower productivity and thus increasing the average cost of production.

Loss of Communication

In an oversized firm, there might also be coordination and communication problems across decentralised sections of the firm especially when bureaucratic structure takes shape and coordination becomes more complex. This might result in possible spillover effects of disruptions from one section to later sections especially if the firm involves complex interdependencies. Possible production delays can be costly to the firm especially when the delays involve liquidation damage. Delays may also result in the firm having to operate beyond the usual operating hours in order to deliver the goods, thus increasing the average costs of production due to expensive payment of overtime wages and utilities.

3.3 External Economies of Scale

The lowering of long run average cost can be illustrated by a shift of the long run average curve from $LRAC_1$ to $LRAC_2$ as shown in Fig 5. Assuming a firm's output remains at Q_0 , its long run average cost of production falls from C_1 to C_2 .

- This paragraph should be written as set up if the focus of the essay is largely on economies of scale. This will save you tremendous amounts of time as you do not have to rewrite the analysis for each IEOS paragraph.

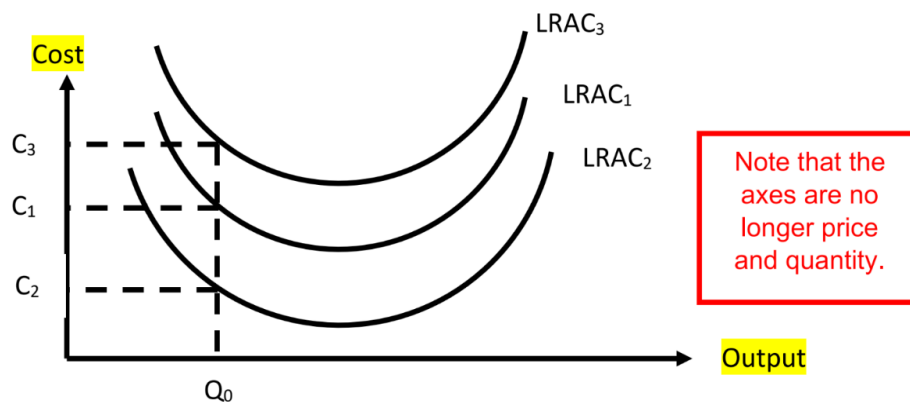


Fig 5: External Economies of scale

Economies of Disintegration

With the expansion of the industry, a network of auxiliary firms providing support to the production process of the industry, for example production of intermediate products (e.g. spare parts of medical equipment), might develop. This creates the possibility of many subcontracting relationships between the firms and auxiliary firms. As the specialised components are mass produced by respective specialised auxiliary firms, these auxiliary firms could enjoy internal economies of scale such as plant internal economies of scale via specialisation and division of labour; thus lowering their average costs of production. (Note: Do not be confused, The focus of this whole point is on EEOS not IEOS thus you do NOT need to explain the plant internal economies of scale in depth. We are just using IEOS as part of the explanation to explain the cost-savings that auxiliary firms can enjoy. This would then enable us to explain why LRAC of the firms in question can be lowered.) On one hand, firms in question can outsource the manufacturing of certain components of their final product to the auxiliary firms so that the firms in question can concentrate on their core production. On the other hand, the auxiliary firms can pass the lower costs in terms of lower prices to firms that outsource to them. Overall, firms in the industry can lower their production cost via outsourcing as compared to producing all the components themselves; thus might lower their long run average costs.

Economies of Information

With the expansion of the industry, supporting firms offering research services or common information for the industry, for example data analysis of less sensitive information (e.g. efficacy of medical treatments), might develop. The firms can therefore obtain up-to-date information on recent and most efficient techniques (e.g. medical treatments) or best practices at a lower cost through cost-sharing instead of spending on costly independent research. Overall, firms in the industry can lower their costs of obtaining information as compared to collecting their own information; thus might lower their long run average cost.

Economies of Concentration

With the expansion of the industry (e.g. healthcare), the demand for trained workers (e.g. nurses) increases. Courses might be set up by educational institutions to provide the necessary training leading to an increase in the availability of the required trained labour (e.g. nurses). This provides the firms ready access to a larger pool of trained labour which will minimise their searching costs such as costs arising from intensive local and overseas advertisement and recruitment. Also, some in-house training can be reduced as the trained labour would already have the skill prerequisites (e.g. new medical processes) required to take up the job. Therefore, training costs can be reduced. In addition, due to the concentration of rival firms selling substitutes in consumption locating in close proximity as the industry expands, the firms can share the maintenance and operating costs of costly infrastructures (e.g. drug testing lab) that was set up by a third party or the government for the expanding industry. Overall, firms in the industry can lower their searching, training and infrastructure costs as compared to having their own search, training and infrastructures arrangement thus lowering their long run average costs.

3.4 External Diseconomies of Scale

When an industry becomes too large, the competition for common but scarce factors of production would be intense. The huge increase in demand would result in a significant increase in the price of the factor of production especially when the supply of the factor of production is price inelastic perhaps because the required skills of the required labour would take a long time to master. In order to obtain the factor of production to produce their goods and services, firms in the industry would bid up the prices resulting in significant increase in their cost of production as illustrated by a shift of the firm's long run average cost curve from LRAC1 to LRAC 3 as shown in Diagram 1; thus given output is at Q_0 , the cost would increase from C_1 to C_3 .

Moreover, as more and more firms share the common facilities and infrastructures, wear and tear would be rapid and soon the facilities and infrastructures would be stretched to the limit giving rise to overcrowding and other problems e.g. frequent breakdown of facilities, causing lower productivity and hence long run average cost increases as well.

4 Firms' Revenue

Average Revenue (AR) is total revenue (TR) divided by output (Q).

Marginal Revenue (MR) is the addition to total revenue from the sale of an additional unit of good/service produced.

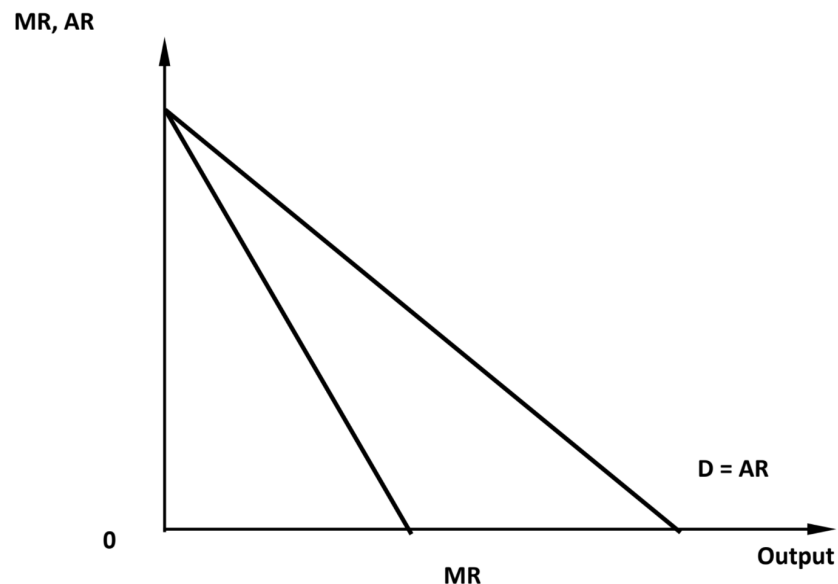


Fig 6: Average Revenue and Marginal Revenue Curves

Key Features of the Diagram

1. Downward sloping average revenue (AR) curve. The AR curve is effectively the demand curve for a single firm. It represents the highest price consumers are willing to pay at each output level. Hence, the downward slope is explained by the law of diminishing marginal utility.
2. The marginal revenue (MR) curve is twice as steep as the AR curve. This is very hard to derive but just know that it can be mathematically derived.

Market power refers to the ability of a firm to influence or control the output and price of a good or service that they sell.

While in chapter 2, we learnt that the market price in the free market is determined by the intersection of market demand and supply, this is not true when a firm has significant market power. In the free market, the assumption of perfect competition holds. That is, no single buyer or seller is influential enough to control a market and exploit other sellers or buyers. However, in reality, firms can influence the market if they are large enough! We will explore this in greater detail in later sections, but right now, it is important to understand that in general, firms can make pricing and output decisions. Hence, a firm can artificially (not due to market forces) create a shortage by producing less output i.e. restricting its output so that quantity supplied is lower than quantity demanded at the original price that the firm sells its goods at. Consumers who are more willing and able to buy the goods will bid the prices up, and as prices are bidded up, those who are less willing and able to buy the goods stop buying from the firm. Hence, the firm's lower output is allocated to those who are most willing and able to buy its goods. In short, we say when a firm exploits its market power, it is "restricting output to set higher prices".

Since the AR curve is taken to be the firm's demand curve, the same determinants (TEN PIGS) that affect the market demand curve tend to affect the individual firm's demand curve as well. For example, if there is an increase in income, there will be a rightward shift of the firm's AR curve. The MR curve also has a rightward shift because of mathematics. Likewise, PED determinants affect the AR and MR curve. For example, if there are less substitutes of the firm's products, then its AR curve becomes more price elastic because when it increases its price by restricting output, its consumers have less options to switch to and hence would be willing to pay a higher price for its products. Hence, a restriction of output will result in a more than proportionate rise in price.

The one exception to this is the level of competition. In fact, when firms face higher competition in the form of other firms entering the market, the market supply curve rises, but the firm's individual AR curve falls. This is because other firms may sell products which appeal to some consumers' tastes and preferences better. Hence, some consumers' switch away from incumbent firms towards newer firms which sell goods which better appeal to their tastes and preferences, causing a fall in incumbent firms' AR and MR curves.

5 Shut Down Condition

Set up Context

The firm's decision to close depends on the level of profit/loss it is making. As firms are profit maximisers, they will aim to maximise revenue and minimise costs. As long as $AR \geq AC$, the firm is making supernormal or normal profits and will continue production. If the firm makes subnormal profits on the other hand, where $AR < AC$, variable and fixed costs of production need to be considered to determine whether the firm will shut down in the short run.

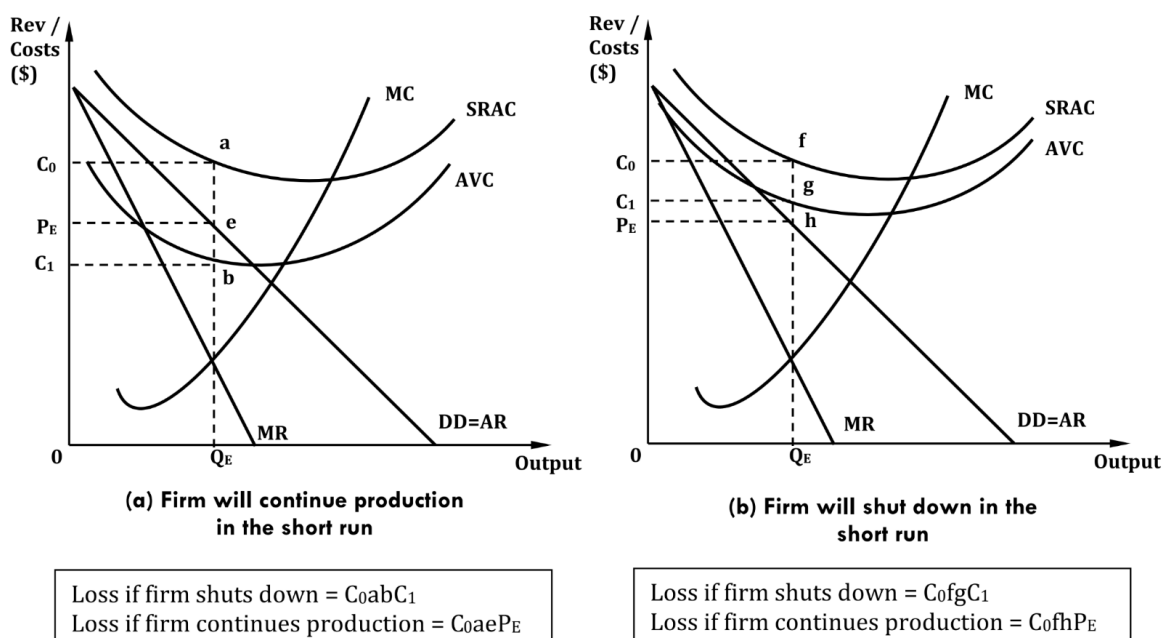


Fig 7: Shut Down Condition in the Short Run

Suppose at the new profit-maximising output of Q_1 where $MR_1 = MC$, the firm charges a price of P_1 which is now lower than AC and the firm is making subnormal profit. Its decision whether to shut down in the short-run will depend on whether its revenue is able to cover at least its variable costs. If P_1 is still higher than AVC , then the firm will not shut down as $AR > AVC$. This is because in the short-run, fixed costs must be paid, even if the firm is shut down, the firm is committed to paying rent. Therefore, if the firm stops production, it will make a loss equivalent to its total fixed costs (area). However if it continues production, even though it starts to incur the variable costs, the revenue earned can pay for the variable costs. And if the revenue can cover more than the variable costs, the firm can cover part of the fixed costs. This means that the loss incurred by continuing production in this case (area) is smaller than by shutting down. The firm will not shut down so as to minimise costs.

Suppose at the new profit-maximising output of Q_2 where $MR_2 = MC$, the firm charges a price of P_2 which is lower than AC and AVC and the firm makes subnormal profits. Since the $AR < AC$ and the $AR < AVC$, the firm will shut down because the revenue the firm is making cannot cover its variable costs, much less its fixed costs. If it continues production, its loss is (area), which is equal to both its fixed costs and part of the variable costs not covered by its revenue. In shutting down, the firm only incurs the fixed costs of production illustrated by (area) and not the variable costs, thus shutting down would result in a smaller cost compared to continuing production.

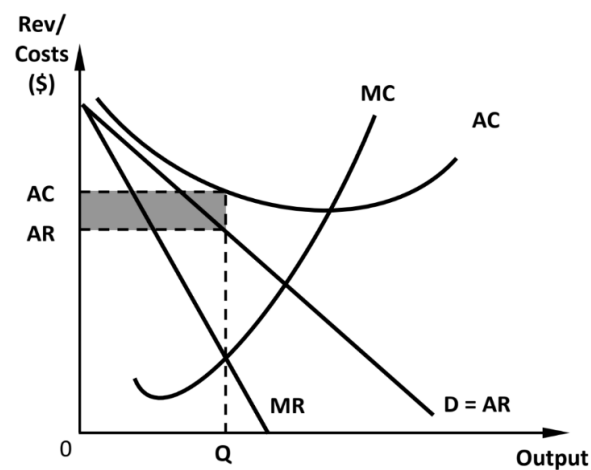


Fig 8: Shut Down Condition in the Long Run

In the long run, when there are no more fixed costs, the firm will shut down if the AR falls below AC . In the long run, if the firm continues to incur losses, it is not making the minimum profit required for it to remain in the business, and hence the firm will decide to shut down.

- This part is frequently left out by many students, but it is a crucial part of the analysis for why a firm decides to shut down.

6 Alternative Objectives of Firms

Profit-Satisficing Behaviour

In reality, firms do not have perfect information about the demand for their products nor their implicit costs of production. Given the constraint of imperfect information, they may aim to achieve a satisfactory level of profit rather than maximise their profit. For example, during uncertain economic conditions where there are large fluctuations in crude oil prices and changes in consumers' taste and preferences for the distillates of crude oil, ExxonMobil may aim to profit-satisfice because the firm is not able to accurately estimate the marginal costs of production and marginal revenue to achieve profit maximisation. Given that ExxonMobil is a multinational company with several production locations and multiple product offerings thus difficult to achieve accurate cost and revenue estimation, to accurately achieve profit maximisation is challenging thus making profit-satisficing the more achievable objective. Also, as a large multinational company which is managed by employed executives, there is a separation between the ownership of ExxonMobil (by shareholders) and the control of the company by paid managers. This gives rise to the principal-agent problem, where the salaried managers' may only aim to achieve a profit level that will satisfy the shareholders. As its shareholders are far removed from the operations of the firm to be fully aware of the decisions that need to be made in order to maximise profit, the firm's managers have greater autonomy in deciding the acceptable level of profit that will satisfy the shareholders. Hence, the firm's objective may be to achieve a satisfactory level of profit rather than the maximum profit possible.

Market Share Dominance

Firms may aim to increase their market share in order to gain greater dominance in the market, thus giving them more market power. Globally, ExxonMobil is operating in a competitive industry and the firm may be willing to forgo short-run profits through the use of price and non-price measures to increase its market share. For example, ExxonMobil might practise predatory pricing, engage in costly advertising or provide free gifts to induce demand for its products. During the screening of Rio Olympics in 2016, ExxonMobil engaged in a high-profile advertising campaign. Since the cost of running the worldwide advertising campaign could have been very significant while it may not have immediate impact on its demand, ExxonMobil would not have been maximising its profit in the short term. Nevertheless, in the longer term, the image of the company would improve as a result of the advertising campaign, which would lead to an increase in the demand of its products and hence its revenue. Moreover, with an increase

in market share, the firm faces a more price inelastic demand curve. This allows it to set higher prices for its products, which increases total revenue and profits in the long run, allowing it to maximise profit in the long run.

Managerial Utility Maximisation

As a large multinational company which is managed by employed executives, there is a separation between the ownership of ExxonMobil (by shareholders) and the control of the company by paid managers e.g. pricing and output decisions are left up to the managers. This gives rise to the principal-agent problem where the hired managers control the company on behalf of the shareholders who own the company. The managers do not necessarily aim to maximise profit. The managers may seek to maximise their own utility in areas like staff salary and bonuses as well as incentives such as entertainment allowances and lavishly furnished offices. These additional expenses will lead the firm to experience higher marginal and average costs. This means that the firm may not be minimising its costs, thus resulting in lower profit.

Growth Maximisation

ExxonMobil's internal expansion may have reflected its growth maximisation objective. For the oil refining industry, high start-up costs are required for the building of infrastructure such as the processing plants to refine oil. So, the MES is only attained at a relatively high level of output. By maximising growth through expansion of output, the firm is then able to enjoy significant cost advantages. It can optimally utilise its indivisible processing plants to spread the cost over larger output levels, thereby lowering average costs, compared to smaller production scale. However, in the process of expanding, profits may not be maximised due to excessive expansion of production capacity, leading to inability to fully utilise it in the short run. If the expansion of the production is not matched by an increase in demand, the firm may be producing at an output where its marginal revenue is less than its marginal cost, thus failing to maximise its profit. Nevertheless, ExxonMobil might have expanded to be recognised as a rapidly growing 'dynamic' organisation, especially with the opportunity to capture the market around Asia. As ExxonMobil expands in size, it may be able to increase its market share, thereby increasing its market power and revenue in the longer term. Even though it may not be maximising profit currently, growth maximisation could be an objective that helps ExxonMobil to profit-maximise in the long run since it brings about significant cost and revenue advantages in the long term.

Revenue Maximisation

ExxonMobil may choose to maximise revenue rather than profits especially since it is a large company where decisions are made by managers rather than the shareholders. This is because sales revenue is often used to judge the success of the managers. Thus, output may be set where marginal revenue (MR) is zero, at Q_3 so that total revenue is at a maximum, as shown in Figure 1. For every additional unit beyond Q_1 , which is the profit maximising level of output, the marginal cost (MC) is greater than the marginal revenue. The production of the additional units of output add more to cost than revenue, lowering profits. The higher output level leads to a lower price and lower profit as compared to the profit maximising level. At output Q_3 , the firm will set its price at P_2 which is determined by the AR curve. The profit is now C_2P_2ef . However, at Q_3 , the MC is greater than MR, which means that total cost has risen by more than total revenue, and thus the profit would have fallen. Hence the supernormal profit is lower than the supernormal profit at Q_1 where $MC=MR$.

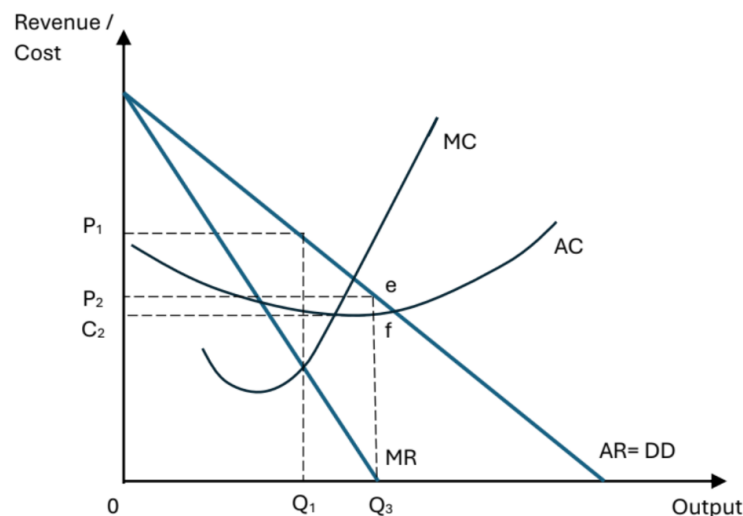


Fig 9: Revenue Maximising Firm

7 Market Characteristics

There are primarily four types of markets that we will be covering in the H2 syllabus — perfect competition, monopoly, monopolistic competition and oligopoly. The **characteristics** of each of these markets are different, which then results in different **decisions** made by firms. These decisions made in the context of characteristics then result in different **market outcomes** in each firm. In order to understand the decisions made and expected market outcome of firms, we will first study the characteristics of firms.

7.1 Number of Sellers and Nature of Product

The **Number of Sellers** refers to how many producers, vendors, or firms are offering a particular good or service in a market. It is directly related to the level of competition in an industry.

The **Nature of Products** refers to the extent to which the goods or services sold by a firm are different from the same goods or services sold by a different firm in the same industry.

These characteristics are essentially the same as the closeness and availability of substitutes, a concept taught in chapter 2. The number of sellers refers to the availability of goods that other consumers can switch to, while the nature of products refers to the closeness of goods that other firms sell to the goods that one particular firm sells.

The implication of this is that the number of sellers and nature of products effectively determine the price elasticity of a firm's demand curve i.e. its AR curve, which then determines its market power/ monopoly power/ price setting power. For example, suppose there are relatively few sellers in the market and the goods they sell are different from the goods that a particular firm sells. When this particular firm decides to raise its prices by restricting its output, most of the customers of this firm will likely not switch to patronising another firm in the market, so the quantity demanded falls less than proportionately. We have shown in chapter 2 that this increases the firm's total revenue, and holding costs constant, its profits as well. Hence, the firm, being a profit maximiser, will likely raise its prices. Likewise, if there are many sellers and products sold are similar, the firm will likely lower its prices to maximise profits.

7.2 Level of Barriers to Entry

Barriers to entry refer to any impediments that prevent new firms from competing on an equal basis with existing firms in an industry. Barriers to entry can be natural or artificial.

- It is important to state this in any question about barriers to entry.

The main implication of the level of barriers to entry is the level of profits firms have in the long run. Let us first start with our understanding of the standard perfectly competitive market.

Since all firms are profit motivated, they will try to maximise profits and minimise losses. We have learnt that the price mechanism serves an incentive function, as when price increases, firms divert more resources to producing the goods and tend to enter the market, and when prices fall, firms divert resources away from producing the goods and tend to leave the market. With the absence of barriers to entry, new firms incentivised by the supernormal profits will enter the market and attract consumers to their businesses until the supernormal profit of original incumbent firms is completely eroded away. However, firms may not always be able to enter and exit the market freely. For example, if the market price has been set to be quite low, smaller firms may not be able to make at least normal profits due to a lack of economies of scale and as a result, they will inevitably shut down and leave the market in the long run. In such markets, firms which can survive would earn supernormal profit even in the long run. due to high barriers to entry in the market which prevent new firms from entering the market easily to share any existing supernormal profit in the market.

The level of barriers to entry also affects the number of firms in the market. Naturally, with higher barriers to entry, there will be less firms in the market. This may or may not be explicitly stated in your answer since

Natural or Structural Barriers

These are barriers that exist by nature of the industry or product or due to existing technological or structural factors relating to the incumbent monopolist that prevent other firms from entering.

1. Control of Essential Resources/Control of Access to Markets

A monopolist may enjoy the exclusive ownership or control of raw materials which are absolutely essential in making a product and thus preventing the entry of rival firms. These higher profits could create a natural barrier to entry to prevent potential competitors from entering the market due to their lower demand of its potential rival and not being profitable.

2. Technological Superiority

A firm which possesses technology which is far ahead of that of its rivals can be in a monopoly position. Technological superiority creates a situation where potential rival firms are unable to be quality competitive since the products they produce are inferior.

- It is important to note that this is not done intentionally by the firm to prevent the entry of other firms, but rather just as a nature of an industry that has advanced significantly.

3. Economies of Scale

In some industries, the Minimum Efficient Scale (MES) of production occurs at a high level of output. This means that the firm which enters the industry first will be able to operate on a vast scale and produce at a low average cost and hence pass on lower costs in terms of lower prices. This deters new entrants, who do not have a large consumer base to warrant a high level of production. New entrants would hence have to produce at higher average costs above market price which might be unprofitable.

Artificial or Strategic Barriers

These are man-made barriers that are erected by firms or institutions to restrict the entry of new firms into the industry. They could be created by the deliberate action of incumbent firms as part of their strategy to reduce competition or prevent new firms from entering the industry.

1. Contrived Barriers/ Deliberately Erected Barriers

An established firm or group of firms may attempt to preserve its dominant position by making entry for others difficult. This usually involves high spending such that any potential entrant is forced to match up to that expenditure should it wish to be in the market. This may also create customer loyalty, making it difficult for new firms to attract customers. E.g. advertising, branding, R&D.

2. Limit Pricing

Limit pricing is a pricing strategy adopted by a monopolist to deter potential entrants by charging a price low enough to make it unprofitable for the new firm entering the industry. Limit pricing is possible when incumbent firms have certain cost advantages that are not available to new firms, thus resulting in new firms having higher cost structures (e.g. due to internal economies of scale enjoyed by the incumbent monopolist as it is supplying the whole market). This allows the monopolist to charge a lower price that would be below the average cost of the new firms.

3. Legal or Statutory Barriers

In many cases, monopolies arise because the government has given one person or firm the exclusive right to sell some goods or services. The firm's monopoly position may be directly protected by:

- Various forms of government licensing (allowing, say, only one firm to operate in a particular area – example, postal and bus services).
- Tariffs (i.e. customs duties) & other trade restrictions.
- Patents. A patent is a privilege granted to an inventor that prohibits anyone else from producing or using that invention without the permission of the patent holder, for a specified period of time. Often, this includes discovering new uses of a product, or a new way of producing a product.
- Exclusive rights given to certain firms are another form of statutory barriers which result in firms having monopoly power.

4. Network Effects

Network effects occur when the value of a good or service increases as more and more people use the good or service. They refer to the value individuals derive from the use of the good by other users. They are also known as network externalities or demand-side economies of scale. Network effects can create a bandwagon effect, as the increase in the use of a good or service by more people in turn attracts more users, resulting in a positive feedback loop.

Existing firms tend to have an extensive network with many users. This makes it more valuable to current and new customers, thereby creating consumer loyalty which makes it difficult for new entrants to draw customers away from incumbent firms.

When one type of technology dominates, it is very hard for a new firm to enter the market with different technology, especially if it is not compatible with the existing technology. E.g. Most people use Microsoft Office software such as Word, Excel and Powerpoint in their work. It is very difficult for a new entrant to introduce a different product as it would not be compatible with the documents used by most people, and thus would not gain a critical mass of users.

7.3 Level of Knowledge

Level of Knowledge refers to the amount of information that buyers and sellers possess about the products, competitors, and market conditions.

This is arguably the most useless of the markets' characteristics, because this primarily serves as an assumption, rather than a direct cause of firms' actions. Typically, we will not cite this as a reason for firms' actions and it is mentioned more for completeness. It may be brought up when discussing firms' decisions and strategies if you really do not have anything else to write. Otherwise, this can generally be omitted as the level of competition is assumed to be imperfect unless it is a perfectly competitive market, in which case you would have stated it in its name. We will dive into more detail of the implications of perfect and imperfect information in markets as that is primarily the focus of our study of perfectly competitive markets.

7.4 Characteristics of Different Market Structures

	Perfect Competition	Monopolistic Competition	Competitive Oligopoly	Monopoly
Number of Sellers	Very Many Sellers	Many Sellers	A Few Dominant Sellers	Sole Seller
Nature of Product	Homogeneous	Slightly Differentiated ²	Differentiated or Homogeneous	Unique
Barriers to Entry	None	Low to None	High	High
Level of Knowledge	Perfect	Imperfect	Imperfect	Imperfect

² Differentiation in products does not make them 'different' per se. They still satisfy the same want or need of consumers, but they may differ slightly in their features. In other words, they are not the same product. Only when a product is unique, then we can consider it to be completely different.

8 Market Long Run Equilibrium

Now that we have a good understanding of market characteristics, let us study the implications of the different characteristics in different markets. We will first be studying the two extreme ends — a perfect competition and a monopoly — to build a solid understanding of the idea of competition in section 8.1.

Thereafter, we will study in depth the nuances in real markets which are usually monopolistic competitions or oligopolies in section 8.2 and section 8.3 to understand that ultimately, the theories presented are merely arguments which are only true to a certain extent in real markets.

8.1 Perfect Competition VS Monopoly

In this section, we will compare firms in perfect competition and a monopoly. To do so, we will first understand what the combined characteristics of each market entail in each market. In general, if a question requires comparison of competition, we will try to use the two extremes, unless the requirements of the question do not allow us to do so (e.g. price rigidity).

Perfect Competition

A perfectly competitive (PC) firm faces a perfectly price elastic demand (i.e. $|PED| = \infty$) for its good as there is a **large number of small firms** in the market selling **homogenous goods** so the PC firm has **negligible market share** so it has no **market power**. This is because if a particular PC firm increases its price, all its consumers, having **perfect information**, would switch to other relatively cheaper **perfect substitutes** in the market and thus, quantity demanded for this particular PC firm would decrease to zero, causing a loss in revenue for the firm. If a PC firm can sell all its output at the current higher price, the profit-maximising firm would not want to sell its goods at a lower price. These suggest that a PC firm is a **price taker**, taking market equilibrium price as given.

- The market equilibrium price is determined by the intersection of the market demand curve illustrated by D_0 and the market supply curve illustrated by S_0 in Fig 1. The perfectly price elastic demand curve of a PC firm is thus illustrated by $AR_0 = MR_0$. A profit-maximising PC firm would produce an output level Q_0 where marginal revenue (MR) is equal to marginal cost (MC) in Fig 1 and charge the market price P_0 . (this part is provided for understanding. It is not required but it would be good to include in tutorial answers to increase your depth of understanding.)

An increase in market demand from D_0 to D_1 causing market prices to rise from P_0 to P_1 would cause an increase in the demand for a PC firm's products, illustrated by an increase in their average revenue (AR) from AR_0 to AR_1 . Hence, at the new profit-maximising output Q_1 where $MR_1 = MC$, a PC firm earns supernormal profit in the short run.

With the **absence of barriers to entry**, new firms incentivised by the supernormal profits will enter the market until the supernormal profit is completely eroded away. The increase in the number of firms would shift the supply curve from S_0 to S_1 , decreasing prices from P_0 to P_1 . With a larger number of firms in the market, each PC firm would face a smaller market share, illustrated by a fall in demand of a PC firm from AR_1 back to AR_0 .

Therefore, the PC firm will decrease its output from Q_1 to Q_0 . The price the PC firm charges will also decrease from P_1 to P_0 , where $AR = AC$. Hence, in the long run, PC firms earn **normal profit**.

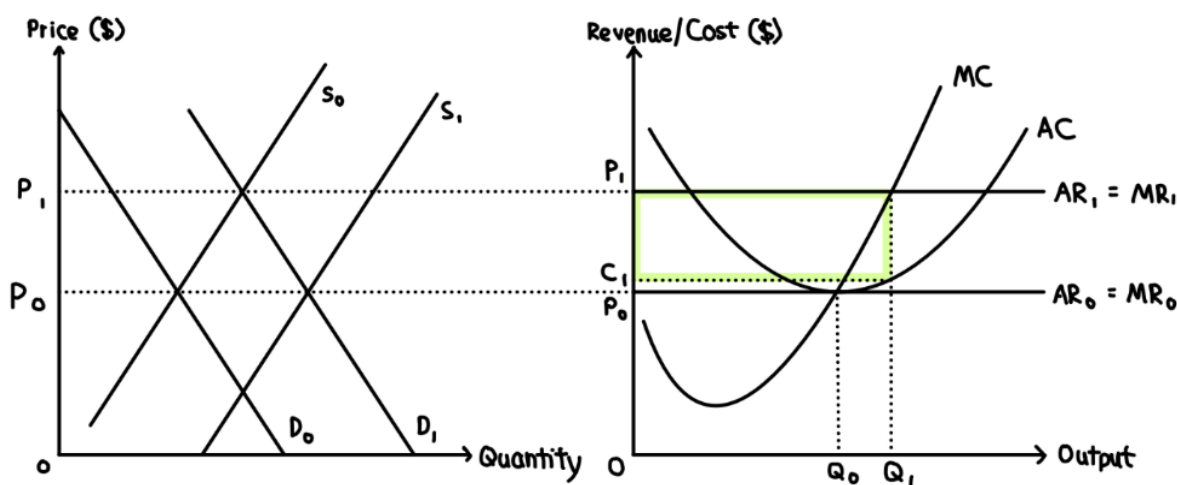


Fig 10: Increase in Demand in a Perfectly Competitive Industry

Conversely, when demand falls causing price to fall from P_0 to P_2 , there is a decrease in the AR of the perfectly competitive firm from AR_0 to AR_1 and the profit maximising level of output shifts from Q_0 to Q_2 . In the short run, at Q_2 , $AC > AR = P_2$ and all firms make subnormal profits. Firms which do not have sufficient accumulated profits will eventually shut down in the long run as they cannot sustain losses for a sufficient period of time. Hence, this causes the number of sellers to fall, decreasing supply for the good until prices rise back to P_0 .

- This is provided for understanding. Typically, you should only pick one of the two changes in demand instead of both.

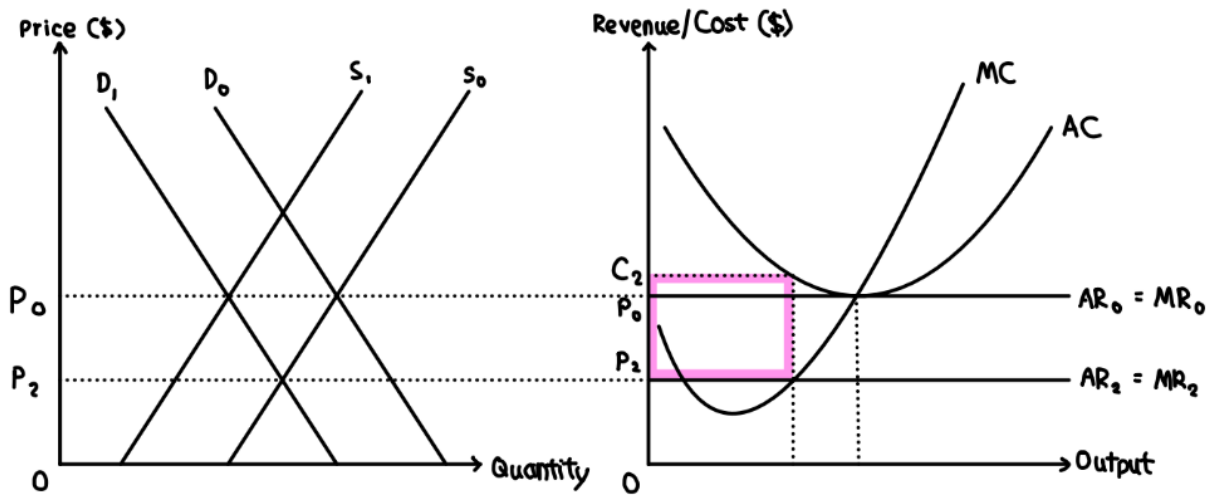


Fig 10b: Decrease in Demand in a Perfectly Competitive Industry

When there is a small increase in marginal cost of production, the profit maximising level of output where $MR = MC$ also falls, so profit maximising firms will decrease output, causing market supply to decrease and the resulting shortage would lead to an increase in price, vice versa.

Monopoly/ Collusive Oligopoly

A firm with market dominance such as a monopoly faces a downward sloping price inelastic demand curve (i.e. $|PED| < 1$) illustrated by AR_0 in Fig 4. Given that the monopolist is the **only dominant firm** in the market and that the monopolist sells a **unique good**, it has a high **market share** and hence a high level of **market power**. When the monopolist increases its price, consumers would have **no close substitutes** to turn to and thus, quantity demanded for the monopolist's goods would decrease less than proportionately. Hence, it is a **price setter**. The profit maximising monopolist would produce an output level where marginal revenue (MR) is equal to marginal cost (MC) as at Q_0 in Fig 4. The highest price the monopolist can charge at a particular output level is determined by its price inelastic demand curve which is also its average revenue curve (AR). The monopolist will thus **charge the price P_0 by³ restricting its output to Q_0 .**

due to **high barriers to entry** in the market which prevent new firms from entering the market easily to share any existing supernormal profit in the market, such as (Explain one barrier to entry), holding average cost at AC_0 in Fig 4, the monopolist would earn **supernormal profit** even into the long run.

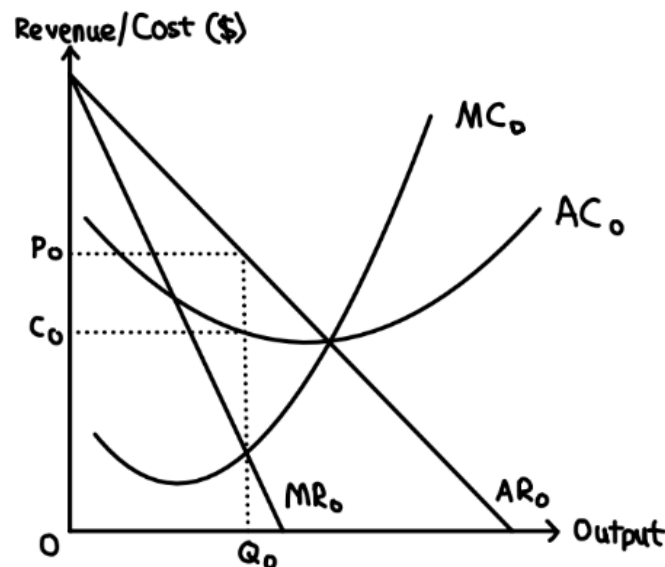


Fig 11: Long Run Equilibrium of a Monopoly

³ The word 'by' is very important here. It indicates the cause and effect relationship between a monopoly setting prices and output. A monopoly can choose to set high prices with the trade off of lower quantity demanded, or it can choose to set high output with the trade off of lowering prices from creating an artificial surplus. It cannot set BOTH price and quantity. This is why we say the highest price that a firm can set is determined by the firm's demand curve.

Natural Monopoly

A **Natural Monopoly** occurs when a single firm can supply a good or service to an entire market at a lower long run average cost than two or more firms can.

A natural monopoly typically occurs if a monopoly experiences substantial economies of scale, serving as a high barrier to entry to other firms' entry because a monopolist already experiencing economies of scale can charge a price below the cost of the new entrant and drive it out of business. This is illustrated by an AC curve which is falling over the entire range of consumer demand. Natural monopolies occur because of the high set-up cost of the technology or infrastructure required for an industry. The high capital outlay results in significant plant economies of scale as it is spread over a large output. (Note: a natural monopoly can still erect artificial barriers to entry, but it **MUST** have natural barriers to entry in terms of IEOS.)

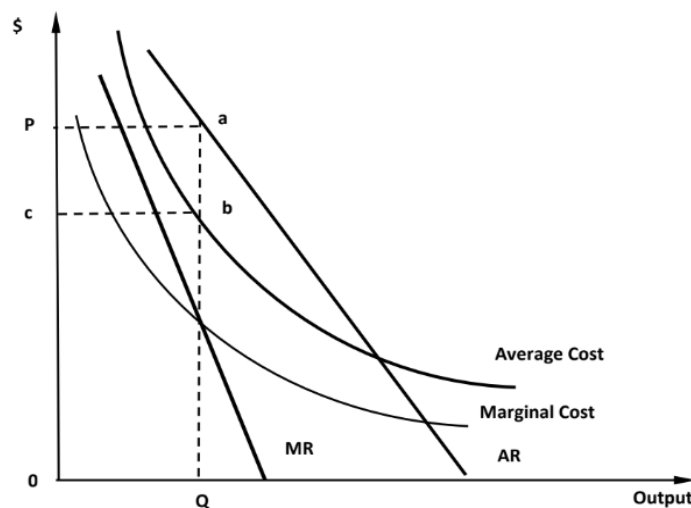


Fig 12: Natural Monopoly

In Fig 12, the profit maximising equilibrium of the monopolist is at Q , where the monopolist makes a **supernormal profit of $Pabc$** . However, if another firm were to enter the industry and assuming that the market is divided equally between the two firms, each firm will now face a demand curve that is half of the original demand curve of the monopolist. The demand curve facing each of the two firms will thus be equivalent to the MR curve of the original monopolist. Since MR lies below the AC throughout its range of output, this means that the demand curve or AR of each of the two firms in the industry falls below the AC. Hence, whilst it may be possible for a single large firm to earn supernormal profit, if two large firms were to supply to the market, they might both make losses.

When we are comparing a monopoly and a perfectly competitive firm, there are two diagrams to be considered — the diagram to show the market outcome of a monopoly is better than a perfectly competitive firm and the diagram to show the market outcome of a perfectly competitive firm is better than a monopoly.

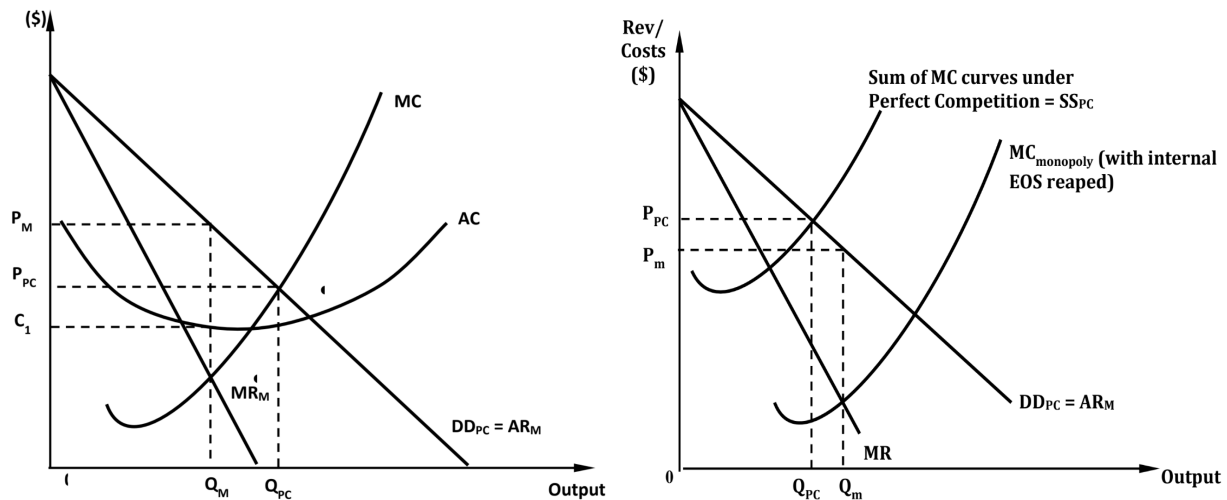


Fig 13: Comparing a Monopoly and Perfect Competition

With its market power, a profit maximising dominant firm such as a monopoly is very likely to restrict output to Q_M to charge a high price P_M at where marginal revenue (MR) is equal to marginal cost (MC) as illustrated in the left diagram. The price P_M is higher than the price level in a perfectly competitive market illustrated by P_{PC} and higher than the marginal cost of the dominant firm as illustrated by C_M . Output level is below competitive level Q_{PC} .

- This has implications on consumer welfare and allocative efficiency as we will learn later. This also shows that the profit level of the monopoly is higher than that of the perfectly competitive firm, as we have previously explained.
- The profits of the firm also have an implication on the ability to carry out innovation and other costly strategies, thus having an effect on dynamic efficiency. Generally, firms which face high levels of competition face normal profits in the long run and do not have the ability to undergo innovation, generally only carrying out less pricey methods of product differentiation like using banners or posters. In a perfectly competitive industry however, any differences will quickly be matched by other firms due to the assumption of perfect information, making their goods homogeneous again, so it is a pointless endeavour.
- The price level has an implication on equity.

When a firm dominates the market, it suggests that the firm holds significant market share and thus is producing at a high level of output and having a large scale of production. As a result, the dominant firm is able to enjoy internal economies of scale thus lowering its long run average cost. If the dominant firm passes the lower cost to consumers in terms of lower price P_M compared to the higher price P_{PC} in a perfectly competitive market as shown in Fig 13, consumer surplus would increase. For example, the dominant firm might be able to enjoy (exemplify).

- This has implications on consumer welfare and firms' profits, as well as equity, ability to carry out innovation and dynamic efficiency.
- The diagram is not drawn usually because it is usually very time consuming to draw this diagram due to how specific it has to be.
- Alternatively, you can draw the natural monopoly diagram to illustrate this point if the context is set in a natural monopoly. (Fig 11)

Important Thing to Note:

1. You will notice that these arguments are seemingly contradictory. How can a monopoly have a higher price and a lower price than a perfect competition? Indeed, it is not possible for both events to occur at the same time. Therefore, we should view these as arguments for whether a monopoly will result in a higher price or lower price than a perfect competition instead. They are sort of like 'forces' which influence whether the price will be higher or lower in a monopoly compared to a perfect competition.
 - a. If the argument based on market power wins out (which it usually does), then the monopoly will have a higher price. If the argument based on IEOS wins out (in the case of a natural monopoly, where there are substantial IEOS to be enjoyed, this is more likely. But it is not exclusively in a natural monopoly where this argument wins out), then the monopoly will have a lower price.

8.2 Monopolistic Competition

Perfect Competition VS Monopolistic Competition

In this section, we will study what the operation of a firm in the real world typically looks like when there is relatively high competition — a monopolistic competition. So, what makes a monopolistic competition different from a perfect competition?

The primary differences between a monopolistic competition and a perfect competition are the nature of products and the level of information. Essentially, the **imperfect information** allows firms in monopolistic competition to carry out product differentiation like using banners or posters. This is because unlike in a perfectly competitive industry, differences in products will quickly be matched by other firms. This is because of the large number of firms in a monopolistic competition, which makes trying to keep track of differences in competitors' products incredibly difficult. Under imperfect information, there is no realistic way to do this.

Hence, in a monopolistic competition, firms can carry out product differentiation. However, firms which face high levels of competition also face normal profits in the long run due to low barriers to entry and do not have the ability to undergo costly product innovation, so they generally only carry out less pricey methods of product differentiation like using banners or posters. That is why in a monopolistic competition, goods are only **slightly differentiated**.

As a result of a **large number of sellers** selling **slightly differentiated** goods, In a monopolistic competition, the demand curve is **price elastic** ($|PED| > 1$). The quantity demanded of a firm's product would fall more than proportionately if it raises its prices since many would switch to cheaper close alternatives which satisfy a very similar purpose. On the other hand, a firm would decrease its price since the quantity demanded would increase more than proportionately to the fall in price, causing a net increase in revenue (using Chapter 2 TR analysis), allowing the firm to maximise profits, holding costs constant. As such, the large number of close substitutes means firms tend to have **low market power** and **sell at high output to lower prices**.

Similar to a perfectly competitive firm, a monopolistically competitive firm can make normal, supernormal or subnormal profits in the short run, illustrated in Fig 14 below. This will typically last for a longer period due to imperfect information, so firms do not immediately know there are supernormal profits to be shared and do not immediately start entering the market.

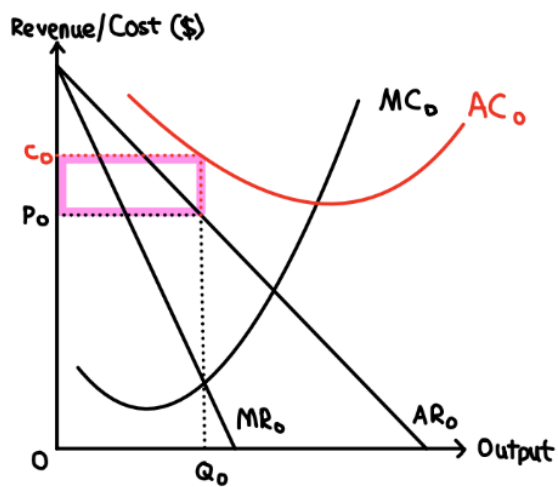
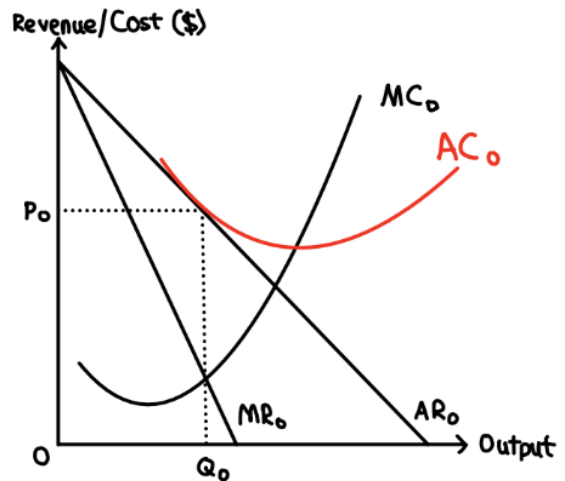
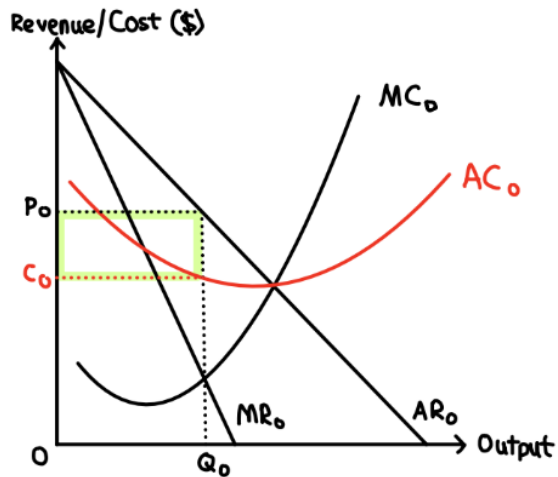


Fig 14: Equilibrium Position in Monopolistic Competition

In the long run, a firm in monopolistic competition makes normal profits where $P_0 = AR_0 = AC$. The analysis here is effectively the same as that in a perfectly competitive firm, with the cause being low barriers to entry.

8.3 Competitive Oligopoly

Monopoly/Collusive Oligopoly VS Competitive Oligopoly

In this section, we will study what the operation of a firm in the real world typically looks like when there is relatively low competition – a competitive oligopoly. Arguably, this is the most confusing market structure for students to understand because the concepts taught are quite difficult to understand. While there are some similarities between a competitive oligopoly and a monopoly, the differences are usually more pronounced.

Let us first understand the similarities between a monopoly and a competitive oligopoly – high barriers to entry and imperfect information. By now, you should understand that the two implications of high barriers to entry are that supernormal profits are not eroded and that there will be relatively little firms in the market. Indeed, dominant⁴ firms do earn supernormal profits even into the long run. While at the extreme end, there are monopolies with only one dominant seller, realistically, there are a few dominant sellers (3 to 7) which dominate a market when there are high barriers to entry. There really isn't any important implication for imperfect information. This is because with dominant firms having such a large influence on the market price and quantity, any decisions to change their price or (not and) output will have a huge and quite noticeable impact.

- The degree to which a market is dominated by a few firms is measured by the concentration ratio. The market concentration ratio is the combined market share of the specified top number of firms in the industry i.e. the combined sales revenue of the top specified number of firms as a percentage of the total sales revenue of all firms in the market. For example, the 4-firm concentration ratio is the combined market share of the top 4 firms in the industry.

Now, we will proceed to explore the differences. The main difference is that a competitive oligopoly has a few dominant sellers instead of just one seller. This results in an economic phenomenon known as mutual interdependence. This has two DIFFERENT effects: (1) there is price rigidity in the market OR (2) there is an incentive (in terms of the greater ease of significantly increasing market share) to collude or engage in mergers and acquisitions.

⁴ Non-dominant firms in the oligopoly will have a relatively small share of the market. Hence, they may make normal profits or at the very most, slight supernormal profits. This may be due to imperfect information which means the dominant firms are unaware that it has some competitors to eliminate. I.e. The dominant firms engage in profit satisficing behaviour.

Mutual Interdependence occurs where a firm is affected by its rivals' pricing and output decisions and its own pricing and output decisions affect its rivals' decisions.

Price Rigidity refers to a resistance of firms in a competitive oligopoly to change their prices, even when marginal costs of production change.

In the case of a competitive oligopolistic market where there are relatively **high barriers to entry** and thus there are only a **few dominant firms** with each firm having **large market shares**, there is **stronger but less competition**. Price in a competitive oligopolistic market is more stable due to price rigidity caused by mutual interdependence between firms in the market. Mutual interdependence occurs where a firm is affected by its rivals' decisions and its own decisions affect its rivals' decisions.

For example, when one of the firms decreases its price, consumers of its rivals would switch to buying from the former and the latter would lose market shares if the latter do not lower their prices. Therefore, all other firms in the competitive oligopolistic market are likely to match the lowered price rather than risk losing their market shares. If so, the firm that lowers its price will experience a less than proportionate increase in quantity demanded i.e. it faces a **price inelastic demand above P_0** in Fig 15 below. On the other hand, when one firm raises its price, rival firms would not follow suit because the rival firms can gain a larger market share if they just keep their prices constant. If so, the firm that increases its price would face a more than proportionate decrease in quantity demanded i.e. it faces a **price elastic demand below P_0** in Fig 15 below. Furthermore, due to the mutual interdependence between firms in a competitive oligopolistic market, there is little incentive for the firms to increase or decrease prices even if there is an increase in cost of production because of the vertical break in the MR curve caused by the kinked demand curve. Shifts in the MC curve within the vertical break will have no effect on the profit maximising output and price at P_0 and Q_0 .

- Note: Graphical illustration is NOT required. You will omit the graphical illustration if time is a constraint.

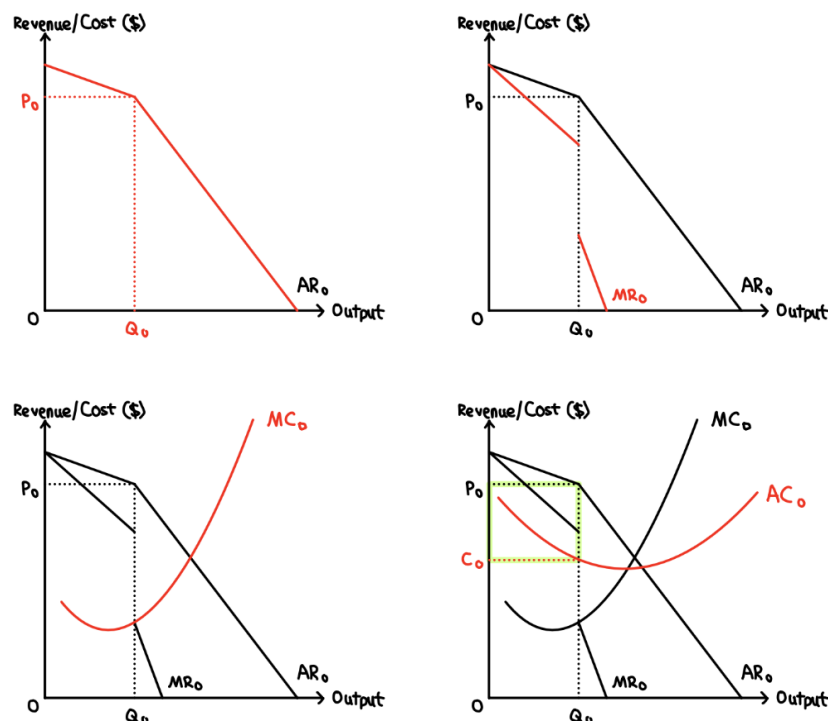


Fig 15: How to draw the Kinked Demand Curve

Collusion occurs when oligopolistic firms agree to limit competition amongst themselves.

A **cartel** is a group of firms which agree to cooperate and act jointly in setting their output as a group or fixing the price of their product.

Tacit collusion can occur in the form of price leadership when oligopolistic firms follow the pricing behaviour of a dominant firm without any formal agreement or informal price fixing.

Collusive oligopolies (non-competitive) oligopolies will face a downward sloping demand curve that is similar to a monopoly's demand curve. This is because effectively, the firms operate as one large firm to dominate the market. While strictly speaking, there are still multiple firms in the market, they can be treated as one sole seller. All the firms may also agree to sell the same product for purposes of standardisation to make pricing and output decisions simpler. However, this is not a necessary condition as firms may be informally colluding, in which case they are unlikely to sell the same product in order to prevent being accused of collusion, which is actually illegal in most countries. Firms in a collusive oligopoly may agree on prices, market share, advertising expenditure etc.

Collusion is most likely to occur and continue based on the ease of coordination, which is affected by three main factors:

- (1) The market is more **concentrated** – when a market is dominated by a few firms, it is easier to collude compared to one in which there are a larger number of firms.
- (2) There is a **clear market leader** – when the market is largely dominated by a particular firm, other smaller firms will tend to follow the pricing and strategies of the market leader (tacit collusion).
- (3) The **nature of product** – firms in an oligopoly are more likely to collude if they sell a homogeneous product, as setting different prices from competitors can lead to loss of market share. Where the products are differentiated, there is less need for collusion as they can still set different prices from their competitors without losing market share.

Formal Collusion and Cartel	Tacit Collusion and Price Leadership
Product is usually homogeneous in nature, so that they can act as a single firm in restricting their combined output or fixing a single price for their product.	Product tends not to be homogeneous in nature, in order to avoid suspicions of tacit collusion which is illegal in most countries.
Firms formally decide to coordinate their pricing and output decisions in order to act effectively as a monopoly. This may be achieved in the form of production quotas for each member country. In order to maximise the profit of the cartel as a whole, the cartel decides on the output level where their combined MR is equal to their combined MC. OR A cartel may decide to divide the market among themselves by assigning different geographical territories or market segments to each cartel member and agreeing not to compete with one another in their respective territories or market segments.	If one firm controls 50% to 95% of the market with the remaining firms accounting for the remaining market share, dominant price leadership may arise. It is likely that the large firm's immense size and efficiency will result in the smaller firms accepting its leadership. Price competition is not likely, as the substantial IEOS experienced by the large firm makes it unlikely that the smaller firms will choose to undercut the prices set by the price leader. The smaller firms are likely to experience higher average costs compared to the large firm, therefore any price increases set by the leader will be followed without hesitation.

9 Size of Firms

Limited Market Size (Demand Factor)

The size of a business can be determined by the size of its demand i.e. demand factor. Even if a business can enjoy huge average cost savings from having a large scale of production via enjoying internal economies of scale (IEOS) as illustrated by a high marginal efficient scale (MES) that occurs at Q_{MES} in Fig 9, having a large output beyond Q_1 would result in huge wastage cost from disposal of unsold outdated consumers electronics that would decrease the business' profit if it faces a much smaller demand for its goods as illustrated by D_1 , perhaps due to the personalised nature of its products arising from consumers' preferences for such products (must explain if not explained in later part like in these set of notes). The size of the business would thus be limited to Q_1 . However, if a business faces a much larger demand as illustrated by D_2 , the business can then grow big while enjoying internal economies of scale. Normally, such products are standardised products that lack personalisation are mass produced via large scale of production by the big businesses.

- **Subcontracting Relationships:** Small firms may also exist as they serve as auxiliary firms to support large firms via subcontracting. Small firms manufacture parts and components that are used by large firms in the making of the final product for sale. For example, the transportation machinery industry, which includes car making, has a large number of firms making car parts for larger firms like Toyota so that the larger firms can concentrate on their core production. Since the demand for parts and components generally only comes from firms, the target consumer groups for such markets are usually quite small. This implies that the subcontracting firms that exist in these industries do not need to be big to capture the market share.

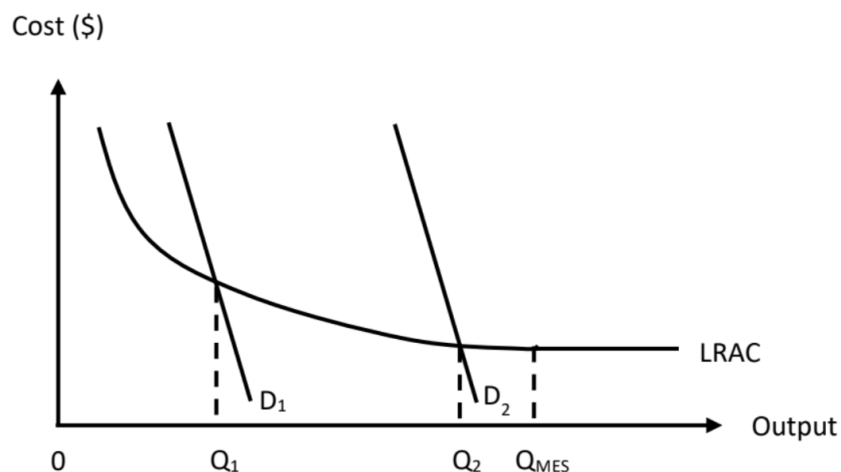


Fig 9: Demand Factors

Consumer Preference for Specialised Products or Services (Demand Factor)

Small businesses exist (alongside with big businesses) to cater to consumers' taste for varieties and personalised services. Consumers demand a wide variety of styles, patterns, and designs. For example, in shirt making businesses, small bespoke tailors exist to cater to a small niche group of consumers who want custom-fit shirts to look better. To tailor-make a bespoke shirt, close contact with the customers and one-to-one service by tailors themselves are necessary due to the unique physique, special design, cutting and/or the request for different materials by different individuals. Big shirt making businesses are unlikely to be able to provide such personal touches. On the other hand, the existence of a much larger size of consumers who value lower prices over personalisation explains the existence of big shirt making businesses alongside the small bespoke tailors because the former can mass produce standardised shirts at cheaper cost due to their enjoyment of internal economies of scale. The coexistence of big and small firms can thus be explained by the size of demand for the products which is a demand factor.

Such personalised services enable small firms to gain greater customer loyalty, allowing them to enjoy a relatively more price-inelastic demand so they can increase the price of their product without losing significant market share, thus increasing its revenue and profits. Hence, firms providing such services are likely to be small in size as they need to provide individual attention to their clients, while larger firms are less likely to be able to maintain that personal touch.

Niche Markets & Personalised Services (Demand Factor)

Small firms are able to cater to more specialised niche markets, which typically have a smaller consumer base. Small firms are also better able to cater to individual needs and preferences and thus provide more personalised services to their customers. Such personalised services enable small firms to gain greater customer loyalty, which the large firms may be unable to achieve. Hence, smaller firms cater to consumers with more specific needs or preferences and thus have relatively more price-inelastic demand. This allows the firm to increase its price without losing significant market share, thus increasing its revenue and profits. Examples can be boutique hotels with special themes, travel agents specialising in eco-tourism and minimarts selling only organic products.

Technological Disruptions in the Industry (Supply Factor)

Changes in demand for businesses due to significant advancements in technology may also allow for the existence of smaller firms. Disruptive technology is an innovation that significantly alters the way that consumers and firms operate. It may sweep away the systems or habits it replaces because it has attributes that are recognisably superior. Recent disruptive technology examples include e-commerce, online news sites, ride-sharing apps, and GPS systems. Firms that are able to ride on and exploit such technological changes may expand and eventually dominate the market, while firms that are slow to respond to technological changes will gradually lose market share and may eventually exit and disappear even though they could have once been a dominant force in that market. Thus, traditional providers of these services would see a fall in demand.

Nature of Product (Supply Factor)

The size of a business can also be determined by the nature of the product which affects its production process and in turn affects the ability of the business to enjoy IEOS.

Standardised shirts of the standard cuttings and materials etc can be mass produced by big shirt making businesses via the use of machines. This would involve huge capital outlay thus afford huge enjoyment of internal economies of scale with MES occurring at high level of output Q_{MES} as illustrated in Fig 10 i.e. significant long run average cost savings to be reaped by having large scale of production.

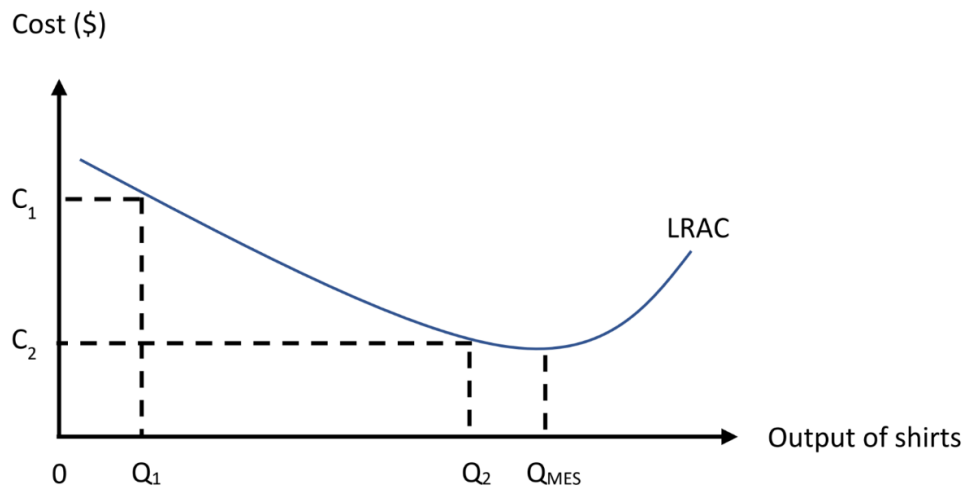


Fig 10: MES at a high level of output

For example, the big businesses can enjoy plant internal economies of scale via the use of indivisible but highly efficient shirt making machines with their large scale of production. The big businesses can better utilise the machineries via for example working on 24 hours shift for non-stop production as compared to bespoke tailors with significantly smaller scale of production where the use of even one of these machines may not be rational due to their definite under-employment i.e. not in operation most of the time. The percentage increase in the cost of operating the machines for the big businesses (e.g. labour and overhead costs) can be much lesser compared to the level of output that can be produced by leveraging on the high productivity of these machines. As a result, their higher initial capital outlay is spread across an even larger level of output thus lowering average cost as illustrated from C_1 to C_2 as output increases from Q_1 to Q_2 as shown in Fig 10.

On the other hand, small bespoke tailors exist alongside the big shirt making businesses because the bespoke shirts, due to the personalisation, can only be hand-made a piece at a time using a combination of hand-held tools. Hence, their MES occurs at low level of output Q_{Low} and rises rapidly past that point of output as illustrated in Fig 11. Larger levels of output would cause the tailors to suffer internal diseconomies of scale (IDOS), for example, resulting from miscommunication with clients and thus resulting in shirts being thrown away due to inaccurate fitting, wrong design or materials used. As a result of the wastage, average cost increases as illustrated from C_1 to C_2 as output increases from Q_{Low} to Q_2 as shown in Fig 11.

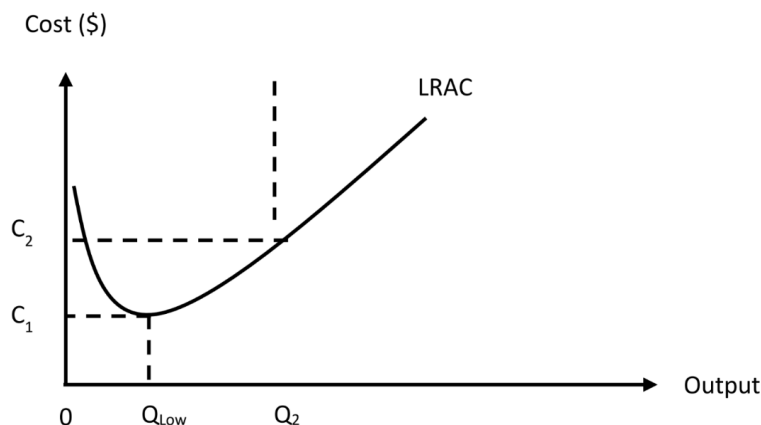


Fig 11: MES at a low level of output

There are also other industries where minimum efficient scale is reached at a low level of output and average cost tends to remain constant or drops only slightly such that big firms do not enjoy any significant average cost advantage over small firms. With reference to Fig 12, small firms could therefore produce at Q_1 and large firms could produce at Q_2 level of outputs and both big and small firms can co-exist in the industry.

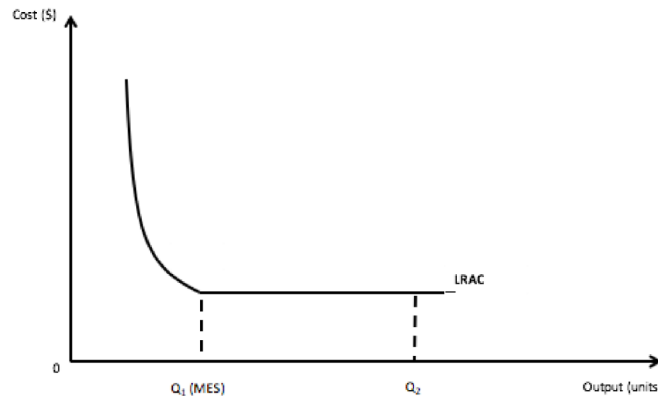


Fig 12: Co-existence of small and large firms

Better Communication/ Adaptability and Flexibility (Supply Factor)

Small firms are more flexible, thus making them more adaptable to changes in market conditions. Being small means that decision making can be quicker as fewer people are involved and there is less bureaucracy in the organisation. The firm is thus able to make changes to its strategies or adjustments to its production faster in response to external changes. For example, in times of recession, the firm may be able to quickly switch its products to more inferior goods to cater to the falling income of consumers. When faced with falling demand, small firms may be better able to adopt cost cutting measures to reduce their cost due to their flexibility e.g. using cheaper materials or offering less services. Moreover, small firms will be better able to react to technological disruptions by adapting their products to align with the new technologies.

Better Control (Supply Factor)

Small firms can more easily get its workers to extend their working hours when faced with a surge in demand, whereas a large firm may have to go through several procedures such as negotiating with trade unions before making changes to working hours or other working conditions. Small firms may be able to maintain better working relations among the workers as there could be stronger bonding and sense of identity among the workers due to their smaller numbers, making them feel like they are part of a family. This increases the motivation and hence productivity of the workers in the firm, lowering average costs. In contrast, workers in

large firms may feel more disengaged and more distant from the management of the company, and thus less motivated in their work. Hence their costs can be lower compared to large firms, even though they are operating on a smaller scale.

Financial Support from the Government (Supply Factor)

Small firms may also have access to various assistance schemes and financial support provided by the government. For example, Enterprise Singapore, a government statutory board, offers various financial assistance schemes such as Productivity Solutions Grant and Enterprise Development Grant to help Small and Medium-Sized Enterprises (SMEs) in Singapore to develop their capability or venture into overseas markets. Strategic industries tend to receive government support, which should be brought in to contextualise your answer to the question. In Singapore, this has come in the form of lab grown meat subsidies.

Lack of Capital (Supply Factor)

Small firms may also be unable to expand due to lack of capital. Banks may be unwilling to lend to small firms as they are often considered less credit-worthy. Small firms do not have substantial assets to offer as collateral; hence they are usually financed from personal savings and loans from family and friends. As a result, opportunities for expansion of small firms can be limited. (covered under internal economies of scale)

Barriers to Entry (Supply Factor)

(Covered under barriers to entry)

Nature of Product

Each time a consumer buys a good or service, it is an additional unit of quantity demanded. Quantity demanded is NOT the number of consumers. It is RELATED to the number of consumers, but it is the unit of goods and services bought.

- Firms which sell dispensibles tend to have a larger market size since people will continue to buy from the firm over time so there will be a consistent firm. However, firms which sell consumer durables tend to have a smaller market size.