RAFFLES INSTITUTION

YEAR 5 H2 ECONOMICS 2022

FIRMS AND DECISIONS

ECONOMICS

CONTENTS:

This series of lectures explores the characteristics, behaviour and performance of different types of firms in the spectrum of market competition. The behaviour and performance of firms depends on the objective of firms as well as the environment which they operate in. The lectures will apply revenue-cost framework to analyse each market structure and compare between the different market structures, with focus on efficiency, innovation, equity and choices outcomes.

RAFFLES INSTITUTION
YEAR 5 H2 ECONOMICS 2022

FIRMS AND DECISIONS (PART 1)

Lecture Outline

Appendix:

References:

Sloman, J., Economics, 6th Edition, Hertfordshire: Prentice Hall
Lipsey & Courant, Economics, 11th Edition, Harper Collins
Beardshaw, J., Economics: A Student's Guide, 5th Edition, Pearson Education
Maunder, P. & Miller, R., Economics Explained, 3rd Edition, Collins Educational
Begg, D., Economics, 7th Edition, McGraw Hill.
Parkin, Economics, 7th Edition, Addison Wesley

At the end of the lecture series, students should be able to:

- Explain the significance of the marginalist principle in determining profit maximising equilibrium
- Evaluate the profit maximisation goal of firms assumed by traditional economic theory
- Identify and explain alternative theories of firm behaviour: revenue maximization, profit satisficing and market share dominance
- Distinguish between the short and the long run
- Define various short and long run cost concepts
- Sketch the various short run and long run curves and illustrate how changes on costs affect the various curves
- Explain the factors that account for the shape of the long run average cost curve
- Explain factors that affect the size of firms
- Explain the survival of small firms

1 OBJECTIVES OF FIRMS

1.1 Profit Maximising Objective

In a competitive environment, firms will seek to maximise profit in order to have the best chance of surviving in a competitive environment. Entrepreneurs receive income from firms in the form of dividends, and a higher profit level translates to higher incomes for entrepreneurs. Larger firms are also able to reinvest profits into research and development activities.

The traditional theory of supply, or ***theory of the firm***, assumes that firms aim to **maximise profit**. Profit is made by firms earning more from the sale of goods than they cost to produce. A firm's total profit (π) is thus the difference between its total revenue (TR) and its total costs of production (TC).

$$\pi = TR - TC$$

Faced with scarce resources, firms have to decide **what and how much goods and services to produce** in order to yield the highest level of profits. Firms also have to decide on **how to produce** these goods and services (i.e. the method of production).

➢ **Economic Profit**

Positive economic profit is known as supernormal profit	TR > TC or AR > AC
'0' economic profit is known as normal profit.	TR = TC or AR = AC
Negative economic profit (i.e. a loss) is known as subnormal profit	TR < TC or AR < AC

1.1.1 Marginalist Principle in Determining Profit Maximising Output and Price

Recall from Topic 1: 'Central Problem of Economics', we argue that rational decision-making involves weighing up **marginal costs** and **marginal benefits** of any activity. An individual should take an action if and only if **MB ≥ MC**. Similarly, producers make price and output decisions based on the marginalist principle, maximising profits at the output where marginal revenue equals marginal cost. This is illustrated using the table below.

Table 1: Profit Maximising Equilibrium

Output (Q)	Price (=AR)	TR	MR	TC	TVC	ATC	MC	Π=TR-TC	MR vs. MC
0	14	0	-	2	0	-	-	-2	-
1	12	12	12	6	4	6	4	6	MR>MC
2	10	20	8	8	6	4	2	12	MR>MC
3	**8**	**24**	**4**	**12**	**10**	**4**	**4**	**12**	**MR=MC**
4	6	24	0	20	18	5	8	4	MR<MC

Note: Table is used for illustration purposes, knowledge of the numbers is not required.

The first two columns of the table represent a given demand schedule for a price- setting firm. According to the law of demand, as price falls, the quantity demanded for a firm's output increases. Using this information, we can derive the values for total revenue (TR) and marginal revenue (MR), where TR = P x Q and MR = $\frac{\Delta TR}{\Delta Q}$ respectively.

Marginal Revenue (MR) is the additional revenue that a firm makes from selling one more unit of output produced.

The fifth column reports the total cost of producing 0 to 4 units of output. Using this information, we can derive the values for average total cost (ATC) and marginal cost (MC), where ATC = $\frac{TC}{Q}$ and MC = $\frac{\Delta TC}{\Delta Q}$ respectively.

Marginal Cost (MC) is the additional cost that a firm incurs from increasing output produced by one unit.

Finally, profit for the firm can be calculated from the difference between TR and TC.

It is best to think about profit maximisation in terms of whether an extra unit of output should be produced or not.

As seen from the table above, if **MR > MC**, then the additional revenue obtained from the selling the additional unit of output produced is greater than the additional cost of that output produced. Thus, there is a still *higher profit* to be gained by *increasing output.*

Key concept: Condition for profit maximization is MR=MC.

Similarly, if **MC > MR**, the additional revenue received for one additional unit of output is less than the additional cost it would incur, so it should cut back on its output produced to increase profits.

The firm will thus choose to produce 3 units of output (where MR=MC) and gain $12 in profit.

[The diagrammatic representation of the profit maximising equilibrium will be discussed in greater detail under Firms & Decisions (Part 2)].

1.1.2 Limitations of the Traditional Theory of Profit Maximisation

The traditional profit-maximising theories of the firm have been criticised for being unrealistic. This is mostly due to the fact that (i) firms wish to maximise profits but are unable to do so and/or (ii) firms have other aims other than profit maximisation (will be elaborated on in the next section).

In the real world, firms may **lack sufficient or accurate information** about demand and cost conditions that exist. Hence they may be unable to accurately estimate marginal revenue and marginal cost due to unavailable or imperfect information As such, they may not be able to use concepts of marginal cost and marginal revenue when making price and output decisions. For example, marginal revenue may be difficult to compute, given the difficulty in estimating the quantity demanded by consumers at different prices. Therefore, pricing decisions are often made based on firms' evaluation of the *estimated demand* for the good or service, rather than based on precise information on the actual demand for the good or service. Firms also tend to work on the basis of explicit costs instead of economic cost (see Section 2.1.4). In such situations, firm may resort to simpler cost-plus pricing, where firms put a mark-up on top of their average cost of these goods and adjust the mark-up based on how well the good is selling. Hence, true profit can only be *maximised by chance.*

The **cost of obtaining sufficient information** to make profit maximising decisions may also be significantly high in some cases, such as in firms that have several production locations (possibly internationally) and multiple product offerings.

Moreover, there is the problem of deciding the **time period** to maximise profits. Firms operate in a dynamic environment where both demand and cost conditions change. These may arise from both the actions of a firm as well as many other factors outside the firm's control (i.e. non-price determinants of demand and the actions of other firms). The firm is not, therefore, faced with static cost and revenue curves from which it can read off its profit-maximising output and price.

Nonetheless, the assumption that firms maximise profits makes it possible to arrive at

useful predictions regarding their behaviour. The long term survival of firms usually hinges on the firm's ability to generate profits. Hence, profit maximisation assumption of firms is a reasonable one in our analysis of firms' behaviour.

1.2 Alternative Objectives of Firms

Key Concept: Identify and explain alternative theories of firm behaviour: revenue maximization, profit satisficing and market share dominance

Besides the difficulties that firms may encounter in making decisions that allow them to maximise profits, it is important to understand that some firms may choose to focus on achieving other objectives instead. This is particularly so in firms where there are different groups of decision-makers, with each group having a different objective in mind when making decisions. This section seeks to explain the non-profit maximising behaviour of real world firms by looking at non-traditional objectives of firms and other factors.

1.2.1 Revenue Maximisation

One of the features of a complex modern economy is that entrepreneurs/employers (principals) have to employ others (agents) to carry out specialised tasks. Firms frequently employ consultants to give them advice or engage the services of specialist firms such as an advertising agency. It is the same with the employees of the company. They are often seen as 'agents' of their employer. As a firm grows in size, the organisation of these companies is such that there is a separation of ownership from control. Ownership is in the hands of shareholders while the control of day-to-day operations is largely in the directors and managers. This results in the **principal-agent problem**, where the objective of the shareholders (principal) which is profit maximisation may be different from that of the directors and managers (agents). Shareholders tend to want strong returns in the form of dividend payments and a rising share price. Managers, on the other hand, may have objectives such as power, bonuses, large expense accounts, prestige and status. In their self-interest, managers would thus maximise what is known as 'managerial utility'.

Perhaps the most famous of all alternative theories of the firm is the **theory of sales revenue/total revenue maximisation**. Often, the income of sales managers and commission-based employees are largely dependent on the firm's total revenue. The higher the total revenue generated, the greater the commission. Thus, firms with a dominant sales department may choose to maximise revenue rather than profits. Hence, a manager may aim to maximize revenue than profits as his job performance may be measured by the former.

Total revenue (TR) is maximized at the output where no additional revenue can be reaped from producing and selling an extra unit of output i.e where MR = 0

1.2.2 Profit Satisficing

Profit satisficing behavior arises due to the separation of ownership and management. Maximisers behave in a traditional economic way and always try to make the **best possible choice** from the **available alternatives**. Satisficers on the other hand examine only a **limited set of alternatives**, and **choose the best** between them. This is because aiming for the optimal solution may necessitate needless expenditure of time, energy and resources. Satisficing behaviour occurs when firms aim for *minimum acceptable levels of profit.*

As explained above, the existence of a divorce of ownership from control and the principal-agent problem often leads to the assumption that large firms are profit satisficers, where managers in the firm aim for a **target level of profit** rather than the maximum level. This is possibly due to the fact that the shareholders of the firm (who are primarily interested in maximising profits) are removed from the operations of the firm to be fully aware of the optimal decisions that need to be made in order to maximise profits. The decision-makers, such as regional managers, could decide against making profit-maximising decisions in cases where they do not stand to benefit. Instead, the regional managers could decide to **achieve a given level of profits that are deemed to be acceptable by the owners** even though it falls below the profit-maximising level. These

actions are congruent with the argument that provided satisfactory levels of profits are achieved, managers often have the discretion to pursue their own interests and managerial utility is maximised.

In addition, some firms may consider the impact of their goods and production processes on society and the environment and hence use less harmful inputs or production methods, even though this may come at the expense of higher cost of production and lower profits. For example, food manufacturers may use healthier ingredients like unsaturated oils to cook their products or bio-degradable packaging like paper rather than non-biodegradable ones made from plastics to pack and store their products. While critics may consider such actions to be merely responses that cater to the demand of increasingly health and environmentally conscious consumers, it is possible that the managers or owners of these firms could be undertaking such decisions at the expense of profits as they have altruistic intentions or alternative aims.

There could also be businesses i.e. social enterprises that have primarily social objectives. Profits are reinvested for the good of the community rather than in the pursuit of higher profits to satisfy investors. These firms look to achieve social aims over the long run.

1.2.3 Market Share Dominance

Market share is commonly measured by the proportion of the firm's total sales (revenue) vis-à-vis the market. Firms might aim for market share dominance because larger firms are able to attract better talent. This is because employees may prefer to work for bigger companies as it leads to greater prestige and higher salaries. Managers and executives may also aim for market share dominance because gains and losses in market share may correlate with stock performance.

To increase their market share, firms will try to influence the demand/revenue by:

- reducing prices (subject to making normal profits in the long run).
- engaging in strategies to shift their demand curve outwards and making demand relatively less price elastic (increase their market power). This will enable the firm to raise prices and increase total revenue and profit (in the long run), ceteris paribus.

Note:
Strategies of firms will be covered in greater detail under 'Firms and Decisions (Part 2)'.

Specifically, the firm may engage in one or more strategies below:

- **Entry deterrence**: If firms in an industry are earning supernormal profits, there are likely to be potential entrants. Successful entry will lower market share and possibly profits for incumbent firms. To avoid losing market share to these new entrants, incumbent firms could decide to focus their price and non-price decisions on deterring the entry of new firms. An example of how this could be done is to engage in product differentiation through brand proliferation (i.e. investment in developing new products and spending on marketing and advertising to reinforce consumer / brand loyalty). A new entrant may reconsider entry as it would also have to match the significant amount spent on advertising and promotion. The firm could also price the good low enough that makes it unprofitable for new firms to enter the market i.e. limit pricing.

- **Predatory pricing**: The incumbent firm is said to engage in predatory pricing when it prices its goods below the profit-maximizing price to deter new entrants. Sometimes, the price set could be lower than the average variable cost of production (OECD). While the incumbent firm engaging in predatory pricing is likely to earn subnormal profits in the short term, it is able to do so because it has sufficient past profits to cope with the losses incurred in the process. Over time, rival firms might not be able to cope with the losses incurred as a result of matching the low prices set by this firm and choose to exit the market. This allows the remaining firms to increase their market share as well as market power.

1.3 Conclusion

Large businesses are complex organizations made up of many different stakeholders. Each group of stakeholders will have different objectives or goals that may or may not be conflicting. It is unlikely that a single theory based on restrictive assumptions should be able to take account of so many forces, all of which, to some degree or another will have a say in the formulation of a firm's price and output decision. Often, the **dominant group** at any moment will focus on their own objectives – for example price and output decisions may be taken at a local level by managers – with shareholders taking only a distant view of the company's performance and strategy.

Moreover, many firms do not make profit in the **early stages of their existence**. This is particularly true of technology start-ups. One example is Skype, a firm created in 2003. When Skype filed for an IPO in 2010, it was still operating at a $7 million loss. Similarly, Twitter, which was valued at $18bn at IPO in Nov 2013, had a loss of around $80 million in 2012. **Thus, 'profit maximization' is a relevant long run objective, but may not be possible in the short run.** In the short run it may be preferable to run at a loss, and aim for revenue maximisation or establishing a brand presence and aim for growing market share. Similarly, in a recession, firms may not deem it to be the best time to pursue profit maximisation.

Refer to Appendix 3 for Netflix's business plans.

In some cases, firms may **sacrifice profits in the short term to increase profits in the long run.** One could argue that all strategies are ultimately short run sub-sets of long run profit maximisation. For example, by investing heavily in R&D, firms may make less supernormal profits or even subnormal profits in the short run due to higher costs. However, successful R&D may enable higher profits in the future. This is also evident in strategies that aim to deter entry and increase market share.

Finally, firms that **do not face competition** may have little or no incentive to maximise profits compared to those who face high levels of competition. The luxury of being able to wait for deferred long run profits may only be the preserve of established multinationals. The significance of the concept of profits in determining real-world competitive behaviour will be covered in *Firms & Decisions (Part 2)*.

Sectional Summary:

- ✓ Profit is the difference between total revenue (TR) and total cost (TC). It is maximised when marginal revenue (MR) is equal to marginal cost (MC).
- ✓ Firms use the marginalist principle when making decisions, weighing marginal cost and marginal revenue.
- ✓ Most firms aim to profit maximise, though it may be difficult to do so in reality.
- ✓ Besides profit maximisation, firms do have various other objectives including entry deterrence, revenue maximisation, profit satisficing and market share dominance.
- ✓ The ultimate aim of the firm will depend largely on the dominant stakeholder, their stage of development and the prevailing conditions in the market.

2 SHORT RUN PRODUCTION AND COSTS

2.1 Introduction

2.1.1 Definitions

- **Market:** It exists whenever producers and consumers come together to transact with each other. i.e. demand and supply.

- **Firm:** Organization or enterprise formed by entrepreneurs who bring together factors of production - land, labour and capital - to produce goods or services for sale.

- **Industry:** Comprises a group of firms that produce a single good or service, or a group of related goods or services.

Industry	Firms
Smartphone	Samsung, Apple, Huawei, Sony, Xiaomi, BlackBerry
Fast food	McDonald's, KFC, Burger King
Telecommunications	M1, Starhub, Singtel, Circles.life

2.1.2 The Production Function

- **Production:** process by which factors of production are used to create goods and services.
- **Production function:** shows the mathematical relationship between the output of a good/service and the factors of production used in producing them, for a given level of technology within a specific period of time. It shows how output/product will be affected by changes in the quantity of one or more factors of production.

Total Product = f (land, labour, capital, entrepreneurship)

2.1.3 Short and Long Run Production

Key concept: Differentiate/ Distinguish between short run and long run in production and costs.

To increase production, firms need to acquire greater quantities of certain factors of production. However, this may not always be possible immediately.

If the firm wants to increase output in the short run, it will only be able to increase the quantity of certain inputs. E.g. raw materials, fuel, tools and possibly labour. These are known as **variable factors.** But it will have to make do with its existing buildings and most of its machinery. These are known as **fixed factors**.

The **short run** is a time period during which at least ONE factor of production is fixed. In the short run, output can increase ONLY by using more **variable factors**. On the other hand, the **long run** is a time period long enough for **all inputs to be varied**, except the level of technology.

For example, to meet rising demand, an airline has several options to increase the number of flights (Q): hire more flight crew and/or use more fuel. These are variable factors that may be altered in the short run. In the long run, the airline can increase its fixed factors such as purchasing more airplanes.

Note: The actual length of the short run will differ from firm to firm and from industry to industry. It is NOT a fixed period of time.

2.1.4 Accounting and Economic Cost

Key concept: Differentiate between accounting and economic costs.

When measuring costs, economists always use the concept of opportunity cost. Economic cost is a measure of the total opportunity cost of using scarce resources to produce one particular commodity in terms of the next best alternative foregone. It consists of explicit and implicit costs of production.

Total Economic Cost = Explicit Cost + Implicit Cost

- **Factors not owned by the firm → Explicit costs:** Payments made to *outside suppliers of inputs*. These include salaries and wages paid to employees, prices of raw materials, overhead costs (rent, annual maintenance, fire insurance)

- **Factors already owned by the firm → Implicit costs:** Costs which do not involve a direct payment of money to a third party, but which nevertheless involve a *sacrifice of some alternative*. For example, this includes the salary that an entrepreneur could have earned from working his next best alternative job or the interest foregone on funds supplied by business owners. These tend to be difficult to estimate / calculate in reality.

2.2 Short Run Production

Production in the short run is subject to *diminishing returns*.

The **law of diminishing marginal returns (LDMR)**[1] states that as more units of a variable factor (e.g. labour) are applied to a given quantity of a fixed factor (e.g. machinery), there comes a point beyond which the additional output from additional units of the variable factor employed will eventually diminish.

Recall: You have already been introduced to the concept of LDMR under 'Central Problem of Economics'. This section provides a more in-depth explanation of the concept.

Production in the short run goes through **3 stages**:

Assuming a bakery uses only two factors of production to make bread (output):
1) 1 Oven (Fixed factor that cannot be varied in the short run) and
2) Workers (Variable factor)

Stage	Number of Workers	Total output (Loaves of bread)		Marginal product/output (Additional loaves that each new worker adds)	
1	1	4	Increases	4	Increases
	2	10	Increases	6	Increases
2	3	14	Increases	4	Decreases
	4	17	Increases	3	Decreases
	5	19	Increases	2	Decreases
	6	20	Increases	1	Decreases
3	7	18	Decreases	-2	Negative

STAGE	REASON	EXAMPLE
STAGE 1 **Total output rises** (Total Product Rises) Increasing marginal output (**Marginal Product of labour rises**) i.e. Total output rises at an increasing rate **Note: Marginal Product is the extra output produced by employing one more unit of a variable factor.*	More efficient labour-capital combination: fixed capital used more effectively by adding more workers Division of labour and specialisation of tasks lead to **greater efficiency**	A bakery has a fixed quantity of machines that cannot be varied in the short run. In the short run, decisions may only be made on the number of workers to employ. As the first two workers are employed, the number of loaves of bread initially rises more rapidly. The assumption behind this is that with only one worker, efficiency is low since that worker has to manage all the tasks. With one more worker employed, each worker can specialize in specific tasks, and thus can use the available resources more efficiently. *(Efficiency rises with each additional worker employed)*

[1] Not to be confused with law of diminishing marginal utility

STAGE 2 **Total output rises** Decreasing marginal output **(Marginal Product of labour falls)** i.e. Total output still rises, but at a decreasing rate	Inefficient labour-capital combination. Overcrowding arises and the fixed factor is being over-utilised. **This is where the law of diminishing marginal returns sets in where the MP of labour starts to fall.** Each additional unit of labour eventually adds less to total output than the previous unit of labour. This means that **marginal output is falling.**	When a third worker is employed, total output still rises. However, marginal product falls due to insufficient space or limited machines at the bakery. The third worker is able to contribute to the production process but by a smaller extent compared to the previous workers. The same applies for the fourth, fifth, and sixth worker.
STAGE 3 **Total Output falls** **Marginal Product of labour is negative.**	The additional output attributable to the next unit of the variable factor is negative.	Eventually, as more and more workers are employed, marginal product will become negative, because they will get in one another's way. This is the case for the seventh worker and beyond. **How can we then expand production?** Build an additional factory or increase the size of the existing factor or increase the number of machines etc. That is, increase the quantity of the fixed factor. However, this can only happen in the long-run.

Note: An in-depth knowledge of the short-run production function is not required.

2.2.1 Short Run Costs

In the short run, a firm cannot vary its fixed factors. It varies its output by varying the amount of *variable factors.*

Key concepts: Define various short and long run cost concepts and distinguish between total, average and marginal cost.

Fixed and Variable Costs

- **Fixed cost** is cost that *does not vary with output level.* It is a cost that must be paid even when production does not take place. E.g. Rent, Fire insurance, paint for the factory walls
- **Variable cost** is cost that *varies with output level.* It is not incurred when production does not take place. E.g. Cost of raw materials, Hourly wages

2.2.2 Total Costs, Average Costs and Marginal Costs

Term	Definition	Mathematical Representation
Total Cost (TC)	sum of the costs of all the factors of production a firm uses in production.	$TC = TFC + TVC$

Total Fixed Cost (TFC)	cost incurred in the utilisation of fixed factors of production	
Total Variable Cost (TVC)	cost incurred in the utilisation of variable factors of production	
Average (Total) Cost (AC)	cost per unit of output	$AC = \frac{TC}{Q}$ OR $AC = AFC + AVC$
Average Fixed Cost (AFC)	total fixed cost per unit of output	$AFC = \frac{TFC}{Q}$
Average Variable Cost (AVC)	total variable cost per unit of output	$AVC = \frac{TVC}{Q}$
Marginal Cost (MC)	change in total costs from increasing output by one additional unit.	$MC_n = TC_n - TC_{n-1}$ OR $MC = \frac{\Delta TC}{\Delta Q}$

Numerical Illustration (will not be tested)

Consider the same bakery - firm – RafflesBake, which produces bread.
It utilises machines and workers in the production of bread. In the short run, assume that machines are fixed factors of production. The firm has in its possession **1 unit of machine** which costs **$20**. Workers are variable factors of production. It can hire **workers** by paying **$25 for each** worker.

The table below shows the production and costs figures of RafflesBake:

Table 2

Machines	Workers	Output (Q)	MP	TFC	TVC	TC	AFC	AVC	ATC	MC
1	0	0	0	20	0	20	-	-	-	-
1	1	4	4	20	25	45	5.0	6.3	11.3	6.3
1	2	10	6	20	50	70	2.0	5.0	7.0	4.2
1	3	14	4	20	75	95	1.4	5.4	6.8	6.3
1	4	17	3	20	100	120	1.2	5.9	7.1	8.3
1	5	19	2	20	125	145	1.1	6.6	7.6	12.5
1	6	20	1	20	150	170	1.0	7.5	8.5	25

2.2.3 Short Run Cost Curves

Note: You are not expected to illustrate these curves with numerical figures, a sketch of them will do. Derivation of the cost curves are not required.

Observations regarding Table 2:

- Although TFC is constant regardless of the level of output, AFC falls as total output increase, hence **AFC curve is continuously downward sloping**. This is because as output rises, TFC is spread over larger quantities of output, hence, AFC decreases.

- TVC increases as output increases because more variable inputs are used to produce more output, holding the price of variable inputs constant. TVC initially

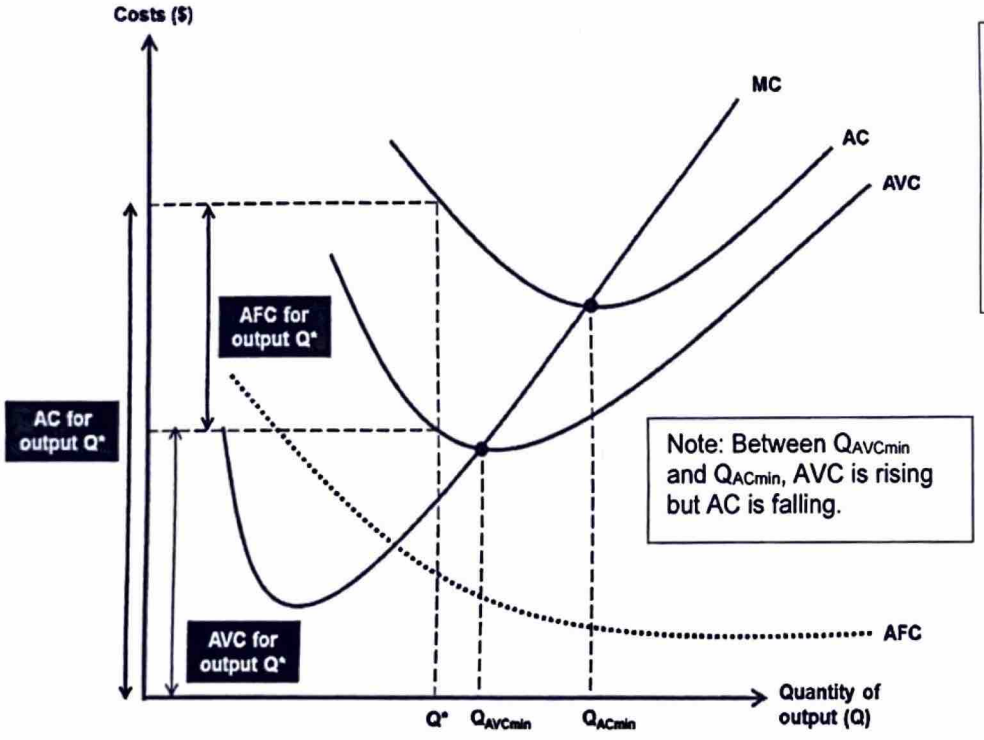
Study tip:
Revising the production patterns of firms will help in the understanding of the derivation of costs curves.

increases at a decreasing rate but eventually increases at an increasing rate due to the law of diminishing returns.

- The MC is the change in TVC between one unit and the previous unit. MC initially falls then rises as output increases. Hence the **MC Curve is U-shaped.** MC initially decreases as production experiences ***increasing marginal returns*** (or rising MP) from the hiring of the first two workers because of more efficient labour-capital combination or better utilisation of fixed factors. Assuming that every worker costs the same to hire, this will result in an initial **fall in MC** for the range of output produced by the first 2 workers.

- MC subsequently increases as production experiences ***decreasing marginal returns*** (or falling MP) due to diminishing returns setting in from the hiring of the third worker because the fixed factor becomes over-utilised. Assuming that every worker still costs the same to hire, but the third, fourth and fifth, sixth worker adds fewer and fewer units of output to production, MC will rise for the range of output produced by these workers. *Note: the law of diminishing marginal returns is a short run concept and diminishing returns sets in when MC starts to rise.

- AVC initially falls then increases as output increases. Hence, the **AVC curve is U-shaped.** The rise in MC (due to the law of diminishing marginal returns) will cause the AVC to eventually rise. Note that AVC starts to rise only after MC has risen above the AVC due to the mathematical relationship between marginal and average values.



Key Concept: it is important to draw this diagram accurately to analyse the decision of firms which will be discussed further in Part 2 of this lecture series.

Figure 1: Short Run Cost curves

Relationship between AFC, AVC and ATC Curves

- As ATC = AFC + AVC, the shape of the ATC curve is determined by the vertical summation of the AFC and AVC curves.
  - *Note: We commonly denote the ATC curve as simply 'AC'.*
- As AFC and AVC are initially decreasing as Q increases, thus ATC also decreases.
- AVC subsequently increases while AFC continues to decrease. However, ATC continues to fall as decreasing effects of AFC outweighs increasing effects of AVC.

- ATC eventually increases as increasing effects of AVC outweighs decreasing effects of AFC.
- As AFC continues to fall with increases in Q, the vertical distance between ATC and AVC decreases as Q increases (the vertical distance between ATC and AVC represents AFC).

Relationship between MC and AVC/ATC Curves

- To produce additional units of output in the short run, more workers must be employed which means that **TVC increases**. TC rises as well (as TC=TFC+TVC). MC also rises.
- There exists a mathematical relationship between MC and AVC and AC: (ATC is the same as AC):
  - When MC < AC, AC must be falling (as additional units of output cost less than the average hence their production pulls AC down)
  - When MC > AC, AC must be rising (as new units of output cost more than the average)
  - Applying the above conditions to the plotting of MC, AVC and AC curves, we observe that the **MC curve must intersect the AVC and AC curves.** The points of intersections are **at the minimum points** of the AVC and AC curves.

Sectional Summary:

- ✓ It is important to distinguish between fixed and variable costs in **the short run.** Changes in fixed and variable costs have differing effects on MC, AC and AVC.
- ✓ We tend to **sketch these cost curves** rather than plot them. Thus, there is no need to memorize exact numerical values.
- ✓ Often, we use the MC, AC and AVC curves in an economic analysis regarding the short run. AFC is usually derived from subtracting AVC from AC.

2.3 Decision Making in the Short Run

In the short run, at least one factor is fixed in supply. Fixed costs have to be paid even if the firm is producing nothing at all. Thus, fixed costs do not have any bearing on a firm's short run decision making. Variable costs, on the other hand, are significant in determining short-run pricing and output decisions. Provided that the firm is able to cover its variable costs, it will continue to produce. Otherwise, it will shut down production.

Note:
An in-depth explanation of the shut-down condition will be discussed in 'Firms and Decisions (Part 2)'.

In order to increase its profits, one strategy the profit-maximising firm can adopt is to reduce its costs of production. If rental costs are increasing at its original location, the firm can look for a new location with relatively cheaper rent so that in the new short run time period, it can lower its fixed costs. If only fixed costs change, marginal costs will be unaffected (as the additional cost of producing an extra unit of a good remains unchanged), hence price and output remains unchanged. However, average total costs will fall. Faced with lower costs, the firm's total profit will increase, ceteris paribus.

The firm can also source for cheaper variable inputs so that in the new short run time period, the firm can lower its variable costs of production. However, when variable costs change (e.g. wages rise), both marginal costs as well as average total costs will rise. When marginal costs change, the firm can pass on some of the cost changes to the consumer, hence the profit-maximising price will increase and output will fall. Average total cost increases and total profits will fall, ceteris paribus.

2.4 Long Run Production and Costs

2.4.1 Long Run Production

Long run is defined as the time period where **all** factors of production can be **varied**, except for level of technology. Firms may hence consider the **scale of production**. Assuming that the costs of labour and capital and the state of technology are constant, different scales of production have implications for long-run production.

If the firm were to double all its inputs, would it double its output? Or will it more than double or less than double? Three possible situations can be distinguished as follows:

*Note that the phrase "to scale" mean that **all** inputs increase by the same proportion.*

Note:
Though the Law of Returns to Scale is not stated in the syllabus, an awareness of the concept is required.

- **Constant Returns to Scale**
  - Output increases proportionately to the increase in inputs.
  - E.g. When all inputs are increased by 8%, output increases by the same percentage i.e. 8%.

- **Decreasing Returns to Scale**
  - Output increases less than proportionately to the increase in inputs. Internal diseconomies of scale are present. Decreasing returns to scale can be approximately represented by the rising section of LRAC.
  - E.g. When all inputs are increased by 8%, output only increases by 5%.

- **Increasing Returns to Scale**
  - Output increases more than proportionately to the increase in inputs. This implies that average cost falls as output increases (the firm is able to reap internal economies of scale).
  - E.g. When all inputs are increased by 8%, output increases by more than proportionately to 20%.

2.4.2 Long Run Costs

In the long run, all factor inputs can be varied. There are thus no long-run fixed costs. All costs, then, in the long run are variable costs.

Long Run Average Cost curve (LRAC curve)

The long run average cost (LRAC) curve shows how **average cost** varies with output.

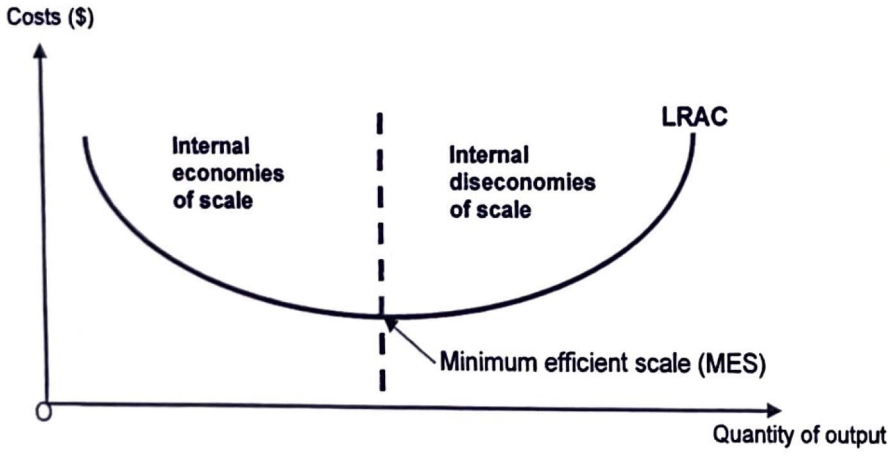


Figure 2 LRAC curve

Key Concept:
Explain the factors that account for the shape of the long run average cost curve

➢ The LRAC curve is **typically U-shaped**

- As output rises, **average costs fall initially** as the firm benefits from ***internal economies of scale*** (Refer to Section 2.4.3).
- When the long run average cost is **at its minimum**, the firm has reached its ***minimum efficient scale***. It is the point where the LRAC stops falling. It is the optimal output level beyond which no significant or additional internal economies of scale can be achieved (Refer to Section 2.4.5 for more information on MES).
- The rising part of the LRAC curve shows **average costs increasing**, **reflecting *internal diseconomies of scale*** (Refer to Section 2.4.3)
- All points on the LRAC represent least-cost factor combinations for that given level of output; points above the LRAC curve are attainable but unwise, while points below are unattainable given present factors of production and level of technology.

Refer to Appendix 5 for more information on the relationship between SRAC curve and LRAC curve.

Besides the conventional U-shaped LRAC curve, some firms (i.e. natural monopolies e.g utilities companies) do experience **continuously falling LRAC curves** as economies of scale are enjoyed over an extremely large output. This is especially for firms with **significantly high startup costs**.

2.4.3 Internal Economies & Diseconomies of Scale

Internal economies and diseconomies of scale are measured in terms of a firm's costs of production. They are concerned with how the average cost (AC) of production changes as the scale of production increases.

i. Internal Economies of Scale[2]

- These are cost savings that occur as a result of the **firm**'s expansion / increasing the scale of production, and have been created by the **firm's own policies and actions.** Diagrammatically, it is represented by the falling part of the LRAC curve. There are various sources of internal economies of scale.

- Internal Economies of scale (IEOS) arise from the **spreading out of fixed costs, productivity improvements and greater buying power over inputs, when a firm expands it production.**

Note:
Important to know and explain a few types of internal and external EOS and DOS well with the use of examples.

The following are some examples of IEOS:

a. Technical Economies of Scale

These refer to gains in productivity/efficiency from scaling up production.

- **Factor Indivisibility**

Some inputs are of a minimum size and cannot be divided into smaller units i.e. they are indivisible. Purchasing such factors of production may not be feasible if a firm's output is small as it means operating below the factor's maximum capacity. For example, a larger supermarket might invest in database technology that improves stock control and reduces transportation and distribution costs. This might not be economical for a smaller grocer to do. In industries like mining and agriculture, heavy equipment such as drills, excavators and harvesters can significantly improve output and reduce average costs of production, but are only bought by large firms that can spread the equipment cost over a larger output.

[2] See Appendix 3 for more examples of internal economies of scale.

- **Law of Increased Dimensions (the "container principle")**

Any capital equipment that is used to contain materials, for example, a blast furnace or an oil tanker, will tend to cost less per unit of output the larger its size. This is linked to the cubic law where doubling the height and width of a tanker or building leads to a more than proportionate increase in the cubic capacity. For shipping companies, a greater increase in capacity means that the average cost of containing and transporting goods will fall as scale increases. It is also important in warehousing and distribution, where a larger plot of land permits a much bigger warehouse than one built on a smaller plot of land. In these industries, large firms can thus enjoy lower unit cost of production. For instance, Amazon has invested in several huge warehouses at its central distribution points all capable of storing hundreds of thousands of items.

- **Specialisation & Division of Labour**

In large-scale plants, workers can be assigned to do *specific and more repetitive jobs*. With this specialisation and division of workers, less training is needed. Workers also become more efficient in their particular job, and less time is lost in workers switching from one operation to another. Workers equipped with specific skills can be employed in those specific areas. The manufacture of cars is often used as an example of the advantages of division of labour, but the principle applies to any mass production technique. In a car assembly line, each worker (or team of workers) is only involved in a particular phase of car assembly, such as installing the car's transmission, the electrical wiring system or doing the paintwork. Workers better at performing each of these tasks can thus focus solely on them, resulting in higher output per worker. This gain in productivity in turn helps to lower unit cost of production for the firm.

Refer to Appendix 2 for observations from labour specialisation.

In another example, workers in a soda factory are also in charge of only a narrow part of the production process. Some may specialize in processing raw materials (e.g. water sugar), some in carbonation while others in bottling and canning.

b. Financial Economies of Scale

Larger firms are often given lower interest rates and larger loans because of better credit ratings and greater collateral. In contrast, smaller firms, especially new startups or small and medium enterprises (SMEs), often face higher rates of interest on loans because banks think it is riskier to lend to smaller companies.

Large firms also tend to be public limited companies. Such firms can raise capital more easily through the issuance of bonds (debt securities issued by borrowers such as firms seeking to raise funds from the financial markets) to the public. They are consequently better able to take advantage of financial economies of scale.

c. Marketing Economies of Scale

Large scale buying and selling give the firm important savings in cost. Large firms have bargaining advantage and are accorded a preferential treatment by their suppliers because they buy raw materials and components in bulk. This allows them to dictate their requirements with regard to price, quality and delivery more effectively. Amazon has huge buying power in the publishing industry. It has a 30 per cent share of the physical book market in the US and more than 60 per cent of eBooks, and uses this power to reduce the prices it pays publishers for the books sold on the Amazon web site.

There are also economies in selling. This can result from bulk advertising and large-scale promotion, as the cost is spread over a larger volume of sales.

ii. Internal Diseconomies of Scale

- Internal diseconomies of scale (IDOS) are increases in costs (particularly average

costs) that occur to a firm as a result of the expansion of the firm, which is the result of the **firm's own policies and actions**. This occurs beyond the MES along the upward sloping portion of the LRAC curve.

- **IDOS arise from the fall in productivity due to greater complexity and lower labour motivation, as well as the employment of more resources to manage such problems, when a firm expands its production.**

The following are some examples of IDOS:

a. High Cost of Monitoring and Management

Monitoring the productivity and the quality of output from thousands of employees in big, complex corporations is imperfect and expensive – this links to the concept of the *principal-agent problem* i.e. the difficulties of shareholders monitoring the performance of managers.

Management problems of co-ordination may increase as the firm becomes larger and more complex. Larger firms often suffer poor communication because they find it difficult to maintain an effective flow of information between departments, divisions or between the head office and its subsidiaries. This can be true for franchise companies like Starbucks and McDonald's. Monitoring costs will start to increase as the firm needs to monitor the quality of food and ensure standardization of menus across all outlets. Franchise owners will also need to get approval from layers of management to make changes to restaurant layouts and to express concerns, which increase average costs. Transportation costs for staff who need to travel between the subsidiaries and the headquarters will also tend to be higher, thus increasing average costs.

Time lags in the flow of information can also create problems in terms of the speed of response to changing market conditions. For example, a large supermarket chain may be less responsive to changing tastes and fashions than a much smaller, independent retailer, and thus, may incur more costs to make the change faster.

b. Low Morale of Employees

When firms become too large, relationships tend to become impersonal. Workers in large firms may develop a *sense of alienation* and loss of morale. The relationship between management and employees is important for maintaining productivity and efficiency. A long chain of authority in large firms may lead workers to feel that they are 'only a small cog in a very large machine', and therefore do not consider themselves to be an integral part of the business. Their productivity may fall leading to wastage of factor inputs and higher costs

Refer to
Appendix 3 for more examples of internal economies of scale.

2.4.4 External Economies & Diseconomies of Scale

i. External Economies of Scale

- External economies of scale (EEOS) can be defined as the savings in costs that occur to all firms as a result of the **expansion of the industry** or the concentration of firms in a certain location.

- Diagrammatically, it is represented by the **downward shift of the LRAC** curve as average costs fall for each level of output. There are various types of external economies of scale.

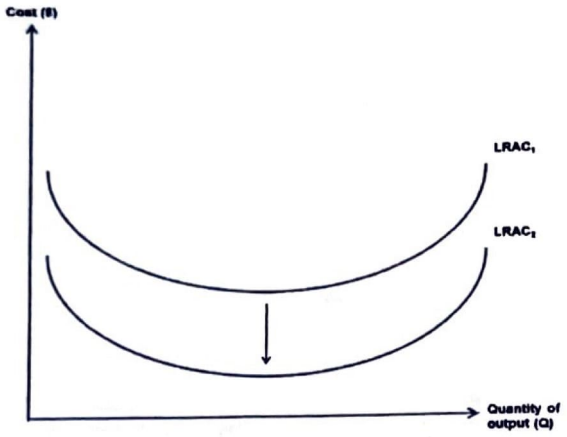


Figure 4: External economies of scale

- EEOS arise from the sharing of common resources between firms and the outsourcing of production processes to supporting firms, due to industry expansion.

The following are some examples of EEOS:

a. Economies of Information

Firms in an industry may share common R&D knowledge or facilities. Increased information flows about R&D processes have allowed for firms to share information on cost-saving technologies. This may also improve the productivity of individual firms and reduce average costs. Firms can also tap on the research of universities nearby that are likely to tailor their research toward the industry. Many scientific and trade journals that are published in the region will likely provide information to all the firms which relate to new markets, sources of raw materials and latest techniques of production.

b. Economies of Concentration / Agglomeration Economies

Agglomeration economies may also result from the clustering of businesses in a distinct geographical location e.g. software in Silicon Valley in the State of California or Singapore's twin hubs of biomedical and engineering research – Biopolis and Fusionopolis.

➢ **Availability of Skilled Labour**

- When the demand for a particular type of skill is large enough, special educational institutions can be set up to train people in such skills. Firms can also join together to develop training facilities for their workers. This will help firms in the industry to reduce costs of training for their workers.
- Well known geographical regions such as Silicon Valley also attract talent for firms, which reduce labour search costs.
- Recently, Amazon chose New York City as the location for its second headquarters mainly because of the availability of highly-skilled tech workers in the city. The firm will incur lower labour search costs.

➢ **Well-Developed Infrastructure**

When many firms are concentrated in one area, governments may be encouraged to invest in infrastructure to cater to that industry. These include better connected roads, railway lines and airports for transport of finished goods or raw materials, or public utilities and other commercial facilities to support a larger work force. Such well-developed infrastructure can help improve productivity, which increases output per unit input and reduce average costs for individual firms.

ii. External Diseconomies of Scale

- External diseconomies of scale (EDOS) are increases in costs that occur to all firms in an industry as a result of the **expansion of the industry** or the concentration of firms in a certain location. Diagrammatically, it is represented by the **upward shift of the LRAC** curve as average costs rise for each level of output. There are various types of external diseconomies of scale.

- EDOS arise from infrastructural bottlenecks and increased competition for common inputs, due to industry expansion.

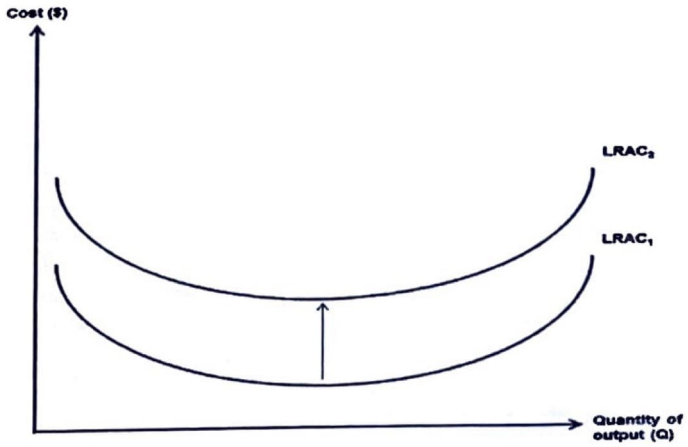


Figure 5: External diseconomies of scale

The following are some examples of EDOS:

a. Increased Strain on Infrastructure

With localisation i.e. concentration of firms in a geographical region and expansion of productive activities, infrastructure will be taxed to its limits. For example, traffic congestion results in increased average costs of production due to loss of time and increased fuel consumption.

b. Shortage of Industry-specific Resources

As an industry grows larger, this may create a growing shortage of specific raw materials or skilled labour. Competition for such resources will push up prices and increase the firms' costs. For example, the increase in demand for skilled engineers may result in a rise in wage costs. Firms may also poach talented workers from one another, resulting in higher turnover and labour search costs.

Sectional Summary:
- ✓ In the real world, various economies and diseconomies of scale exist. When asked about the effects of scaling up, be sure to explain the benefits and costs of larger output to **long run average costs** (instead of just 'costs').
- ✓ It is important to identify **explicitly** and **elaborate** on the most relevant source(s) of economies / diseconomies of scale in the given context.
- ✓ It is better to elaborate on multiple unique sources of economies or diseconomies of scale rather than just one (e.g. technical iEOS).

2.4.5 Minimum Efficient Scale

The extent of economies of scale can be measured by looking at a firm's *minimum efficient scale* (MES). It is the scale of production where the internal economies of scale have been **fully exploited** and corresponds to the lowest point on the long run average cost curve. It is also known as an output range over which a firm achieves *productive efficiency*.

Besides measuring the extent of economies of scale, MES in relation to the size of the market tells us the **maximum number of firms** of minimum efficient scale that the market can accommodate. When limited opportunities for economies of scale exist, MES is usually achieved at low levels of output. Thus, MES is usually a small percentage of total market demand and it is likely that the market will be competitive with many firms able to achieve the MES. The converse is also true.

Minimum efficient scale is critical to strategic decision-making with economies of scale because it identifies the level of scale required to enjoy optimal cost savings. Firms with scale less than MES have a cost disadvantage. Often, firms with scale greater than MES also have a cost disadvantage. In some cases, the lowest average cost is attainable over a large range of output. In these cases, minimum efficient scale is the smallest level of output that can deliver the lowest cost.

For example, for a firm in a given industry to be producing at MES, it needs to produce 2,000 units of product ($Q_{MES} = 2000$). If total market demand is 100,000 units, the industry can accommodate 50 of such firms. Therefore, it makes economic sense in this case to have several firms/plants of MES to produce for the market rather than just one large firm (as it will incur very high costs due to diseconomies of scale that may arise).

The minimum efficient scale and the contestability of the market also depend on the **nature of costs** of production in a specific industry. In industries where startup costs are high, there is plenty of scope for reducing unit cost by **increasing the scale of output** (i.e. MES is only realised at high levels of output). The market will become less contestable and more highly concentrated and that will result in a concentrated market structure (e.g. an oligopoly, a duopoly or a monopoly) – economies of scale may even act as a barrier to entry because existing firms have achieved cost advantages.

2.4.6 Decision Making in the Long Run

In the long run, all factors are variable. The firm may experience constant, increasing or decreasing returns to scale. But although all factors can be increased or decreased, technology is assumed to be constant. Hence, costs affect firms' decision in the long run, such as whether to increase the scale of production. For example, firms with significantly high startup costs (such as those in the telecommunications industry) may choose to increase the scale of production in order to reap internal economies of scale and lower the long-run average cost of production which will be reflected by a movement along the LRAC curve.

Although all factors can be increased or decreased in the long run, technology is assumed to be constant. If technology can be improved such that more output can be produced for a given amount of factors of production, LRAC will fall and this will be reflected by a downward shift of the LRAC curve.

In the long run, if the firm's average revenue is unable to cover its long-run average costs and at least earn normal profits, it will shut down by leaving the market.

3 FACTORS AFFECTING THE SIZE OF FIRMS

Note:
You are required to explain factors that affect the size of firms.

There are some indicators we use to gauge the size of firms, these include the quantity of output sold, profits generated, total sales revenue, number of employees, market share, number of outlets etc.

The market structure of an industry is affected by both the extent of economies of scale available to individual suppliers as well as the total size of market demand. Even though the concept of economies of scale seems to put large firms at a cost advantage, in many industries it is possible for smaller firms to make a profit because of various demand and supply side reasons. There are many reasons why firms may decide to remain small or why small firms can coexist with large firms, survive or even thrive.

3.1 Supply Side / Cost Factors

➢ **Firms reach MES at very low levels of output relative to market output**

- Firms tend to be small if diseconomies occur at low levels of output. Such industries usually provide services that require personal attention e.g. tailoring and private clinics. Average cost rises sharply as output increases. If such specific detail has to be mass produced, diseconomies of scale will set in, and any advantages to large-scale production are more than offset by the disadvantages. The optimum size of firms in such industries is small.

- **Vertical disintegration**: Small firms tend to emerge when an entire production process is broken up into a series of separate processes. Diseconomies of scale quickly emerge in each of these processes. Thus, different small firms each performing a small part of the whole task would incur a lower unit cost. Smaller firms can complement larger firms in the same industry if they specialize in a single process or make components for the larger organisations e.g. air conditioning units for the automobile industry are usually made by smaller companies.

➢ **Saucer-shaped LRACs**

- In some industries, small and large firms may **co-exist** because of cost reasons. This is likely in instances where economies of scale are quickly exhausted, followed by constant unit / average costs over a wide range of output e.g. in breweries. In this case, the LRAC curve is saucer-shaped and has a horizontal portion. It would be possible for small and large firms to be equally cost efficient and hence coexist in the same industry.
- In the figure below, a small firm producing output Q_1 will be as cost competitive as a large firm producing Q_2. It is therefore possible for small and big firms to coexist.

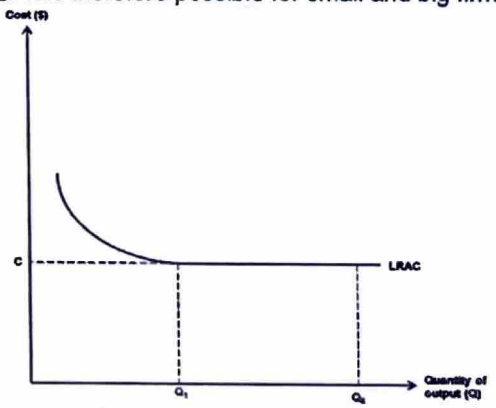


Figure 6: Saucer-shaped LRAC

➢ **Banding and Joint Ventures**

- Independent businesses engage in joint ventures or band together to gain the advantages of bulk buying, advertising and promotion while still retaining their independence. This is not the same as mergers, as the respective owners retain ownership of their businesses. This strategy involves cooperation among smaller firms to protect their interests.
- They can also set up jointly owned enterprises to source for raw materials, allowing them to obtain many of the economies of scale enjoyed by the larger firms.

3.2 Demand Side / Revenue Factors

On the demand side, some firms remain small because the market demand is limited and may not be able to support a large firm. In addition, small firms in some markets that favour large firms can implement strategies that enable them to survive or even thrive alongside larger firms. This could be because of:

➢ **Nature of the product**

- **Personalised goods and services services** require direct individual attention, hence, it is impossible to have mass production. This include legal services provided by lawyers and specialized medical procedures provided by surgeons.

- Firms providing repair services and hairdressing services are also usually small. These firms are small because personal attention to the individual is required. This means the demand for the firm's services is limited because the good or service is very specific.

- For other goods, consumers might have a preference for variety and choice rather than standardisation and mass production. This includes ladies' fashion apparel and accessories.

➢ **Market segmentation and specialization**

- If an industry caters to a diversified range of products and customers, it can be segmented into smaller markets. Large firms can focus on mass produced items while small firms concentrate on customized items. This allows small firms to cater to ***niche markets*** which may be smaller due to high and restrictive prices e.g. expensive sports cars, luxury yachts and high quality jewellery. For example, Honda and Toyota can produce cars for the mass market while Porsche and Ferrari can produce more up-market sports cars. The market for specialised products also tends to be small. For example, highly specialised machines and tools or products of religious significance may have limited markets.

- The demand for goods in niche markets tend to be relatively more price inelastic than in mass markets. Even though, the firm producing goods for niche markets may not be able to exploit fully available internal economies of scale, it can set a relatively higher price to cover its higher unit costs of production.

➢ **Geographical factors**

- If the product has great bulk in relation to value, transport costs will be high relative to total production costs. In such cases, the market for the good is likely to be local rather than national. Fresh produce such as fish and vegetables also tend to have small and localised markets as they are perishable, and can only be transported over a limited geographical region. Finally, small provision shops can distinguish themselves by being placed in convenient areas within an inaccessible location that contains a smaller market. Despite having higher average costs due to a smaller scale, these provision shops can compete with huge supermarket chains.

➢ **Profit-cycles**

- New products appear continually while others disappear. At the early stage of a product cycle, total demand for the product tends to be low. The firm tends to be small as it takes time to grow, outpace rival firms, merge or force others out of business.

3.3 Business Risk & Uncertainty

➢ **Unwillingness to take greater risks**

- Expansion and large-scale production require funds and starting a larger firm involves larger capital outlay. If firms are unable to raise the money, they will have to borrow from banks. This means the risk of investment is greater as the losses can be potentially greater if the business does not work out. Therefore, some firms remain small because they are unwilling to bear the significant risks involved.

- Firms may also fear of a future fall in the price of the final product. A sudden expansion of output may lead to a large enough increase in market supply. The huge surplus will cause the price of the product to decrease. Prospects of lower prices and lower profits may prevent firms from expanding. Thus, if the prospects of a larger market are poor, firms may choose to remain small.

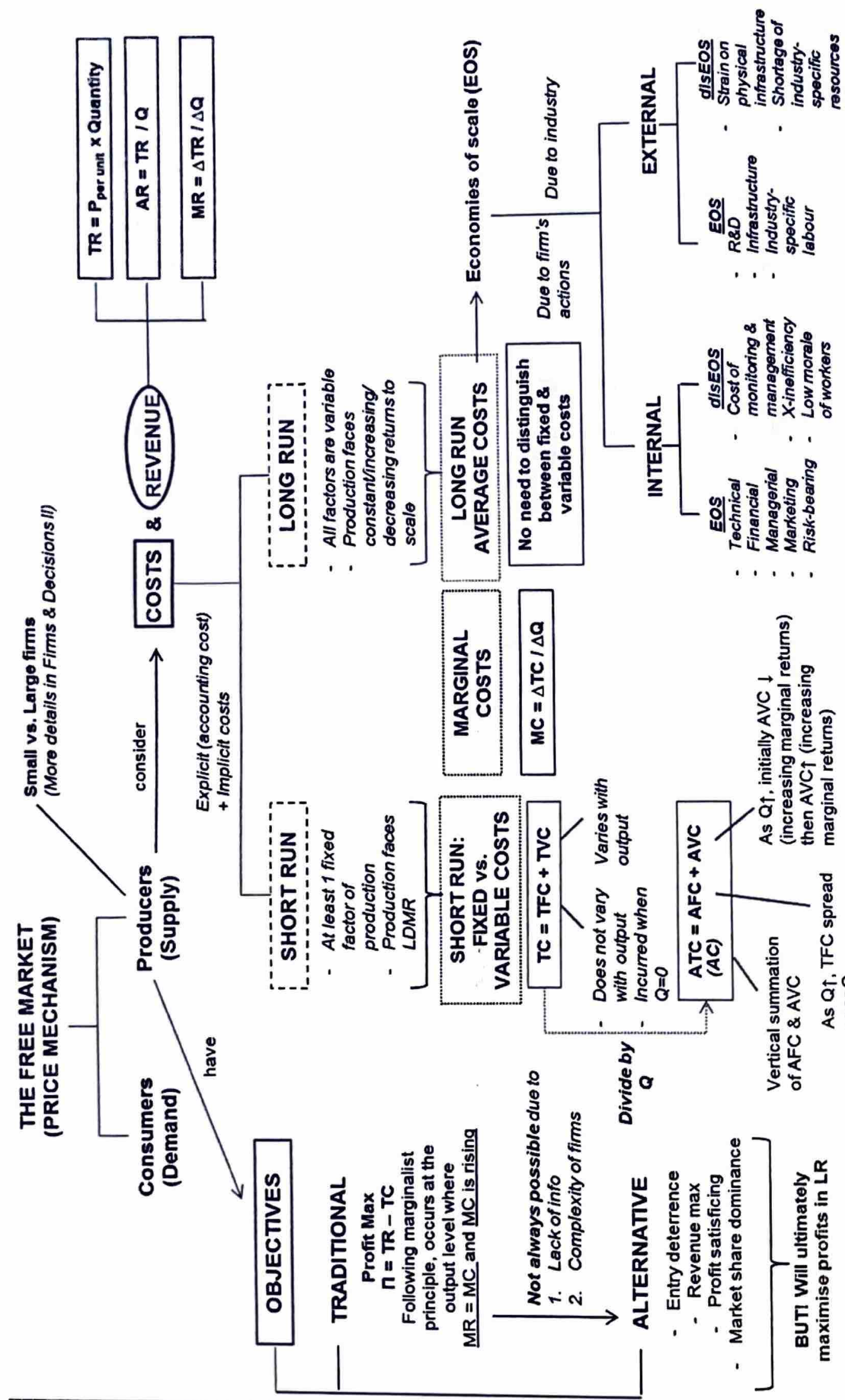
3.4 Alternative Objectives of Firms

➢ **Profit satisficing**

- Some firms remain small because the owners prefer to keep them small for reasons not related to the profit motive. The entrepreneur could be contented with a reasonable income from the domestic or localised market and is unwilling to take on increased risks, stress or perceived challenges from growth. Instead, entrepreneurs could decide to achieve a given level of profits (with a lower output) that are deemed to be acceptable even though it falls below the profit-maximising level, and enjoy other benefits such as shorter operating hours and lower levels of stress.

- The factors stated above suggest that <u>in the real world</u>, large and small firms might **co-exist** in the same industry. A good example is the retail furniture market. The industry has major players such as the Swedish giant IKEA. However, much of the market is also taken by smaller-scale suppliers with consumers willing to pay higher prices for bespoke furniture owing to the low price elasticity of demand for high-quality, hand crafted furniture products. Because these firms can charge a higher price, they can survive in the industry despite competition from a lower cost producer.

Appendix 1: Topic Summary



Appendix 2: Lessons from a pin factory

'Jack of all trades, master of none'. This well-known adage helps explain why firms sometimes experience economies of scale. A person who tries to do everything usually ends up doing nothing very well. If a firm wants its workers to be as productive as they can be, it is often best to give each a limited task that he or she can master. But this is possible only if a firm employs many workers and produces a large quantity of output.

In his celebrated book An Inquiry into the Nature and Causes of the Wealth of Nations, Adam Smith described a visit he made to a pin factory. Smith was impressed by the specialisation among the workers and the resulting economies of scale. He wrote:

One man draws out the wire, another straightens it, a third cuts it, a forth points it, a fifth grinds it at the top for receiving the head; to make the head requires two or three distinct operations; to put it on is a peculiar business; to whiten it is another; it is even a trade by itself to put them into paper.

Smith reported that because of specialisation, the pin factory produced thousands of pins per worker every day. He conjectured that if the workers had chosen to work separately, rather than as a team of specialists, "they certainly could not each of them make twenty, perhaps not one pin a day". In other words, because of specialisation, a large pin factory could achieve higher output per worker and lower average cost per pin than a small pin factory.

The specialisation that Smith observed in the pin factory is prevalent in the modern economy. If you want to build a house, for instance, you could try to do all the work yourself. But most people turn to a builder, who in turn hires carpenters, plumbers, electricians, painters, and many other types of workers. These workers specialise in particular jobs, and this allows them to become better at their jobs than if they were generalists. Indeed, the use of specialisation to achieve economies of scale is one reason modern societies are as prosperous as they are.

Appendix 3: Other examples of Internal Economies of Scale (IEOS)

i. Managerial Economies of Scale

Managerial economies of scale are also related to specialisation and division of labour. The difference is that specialisation occurs on a *supervisory level.* Just as workforce specialisation increases productivity by training employees on a specific step in the production process, managerial scale economies increase productivity by employing specialists to supervise production systems. Better management and increased investment in training managers and the use of specialist equipment, such as networked computers can improve communication, raise productivity and thereby reduce unit costs.

Managerial scale economies also extend to human resources management. HR specialists raise productivity and reduce unit costs by creating efficient and cost effective hiring and labour search processes.

ii. Risk-Bearing Economies of Scale

Risk bearing economies of scale refer to the ability of firms to spread the uncertainty in cost of production over a large level of output and thereby reduce unit costs.

Spreading the risk of research and development costs is one benefit for large firms. The returns from research and development (R&D) are highly variable and uncertain. It is easier for large firms to carry the overheads of sophisticated research and development. For example, in the pharmaceutical industry, R&D is crucial. Yet the cost of discovering the next blockbuster drug is enormous and increasing. Several of the mergers between pharmaceutical companies in recent years have been driven by the companies' desire to spread their R&D expenditure across a greater volume of sales. Thus, average cost is

reduced.

Producing a wide variety of products, and operating in many geographic locations, are also ways to spread risk, and hence keep average cost of production at competitive levels. For example, a sharp hike in wage cost in a factory situated in Europe can be mitigated by relatively cheaper costs of production in other factories in other countries. As another example, large firms can obtain materials from different sources to guard against events such as crop failures and strikes which will significantly cause costs of raw materials to spike.

iii. Network economies

Many networks have huge potential for economies of scale. That is, as they are more widely used (or adopted), they become more valuable to the business that provides them. Good examples to use include online auction sites such as eBay, social networking sites, wireless service providers, and businesses such as Amazon. In most cases, the marginal cost of adding one more user or customer to a network is close to zero, but the resulting financial benefits may be huge because each new user to the network can then interact, trade with all of the existing members or parts of the network. Given the high fixed costs of establishing a network, the more users there are the lower are the fixed costs per unit. Thus, as the network expands, not only are there potential gains from extra revenues, but the long run cost per user diminishes - an internal economy of scale.

Linking network effects to IEOS:
The presence of network effects could lead to lower costs of production per user for firms (i.e. lower AC). Network economies of scale can be applied in the following ways:

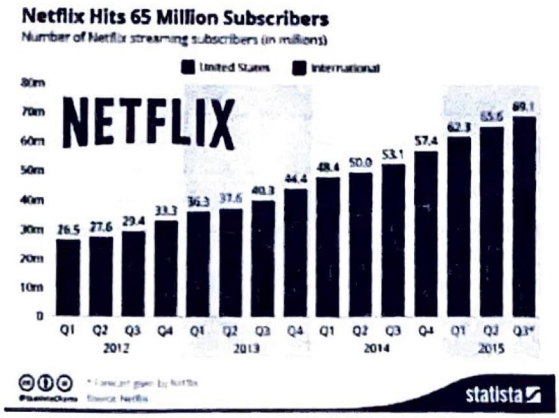
1. Spreading of high start-up costs over a large number of users
There are usually high start-up costs involved in establishing a network such as setting up telephone lines or servers for processing interactions between users. Hence, the more users there are, the lower will be the average start-up costs due to the spreading of overheads. Furthermore, the MC of serving one more user in a network is close to zero. As the network expands, the long run cost per user diminishes.

2. Reducing customer acquisition costs
Beyond a certain critical mass, the network effect helps to draw customers and service providers to use a company's platform, reducing the costs the company needs to spend to attract and acquire new customers (e.g. advertising costs). For instance, as the number of both drivers and riders increase on Grab, new passengers are also drawn to use Grab because of the ease of finding a driver. The same analysis also applies to e-commerce sites like Amazon and Qoo10 where customers are more easily able to find a merchant catering to their needs in a larger network. With the expansion of the network, customer acquisition costs fall and AC falls as output increases.

Appendix 4: Netflix's growth maximisation plan

http://www.tutor2u.net/economics/blog/65-million-subscribers-for-netflix

In the increasingly contestable market for online streaming of movies, TV series and other forms of entertainment, the drive to grow your installed user base is right up at the top of the objective of senior executives. Netflix has announced that now have more than 65 million subscribers with nearly 3 million overseas customers added in the second quarter of 2015 alone.



Netflix is an interesting example of how disruptive technology businesses have **different objectives** from the standard theory of the firm. The Financial Times reported that Netflix is launching an **aggressive international expansion** with the **aim of reaching 200 markets** by 2016. Netflix will feature in Japan, Spain, Portugal and Italy this year and is exploring expansion into China next year.

The business remains some distance from achieving profitability. In a letter to shareholders, Netflix CEO Reed Hastings explained that the company plans to operate around **break-even** through 2016 and to rake in substantial profits starting in 2017.

Appendix 5: Relationship between SRAC curve and LRAC curve

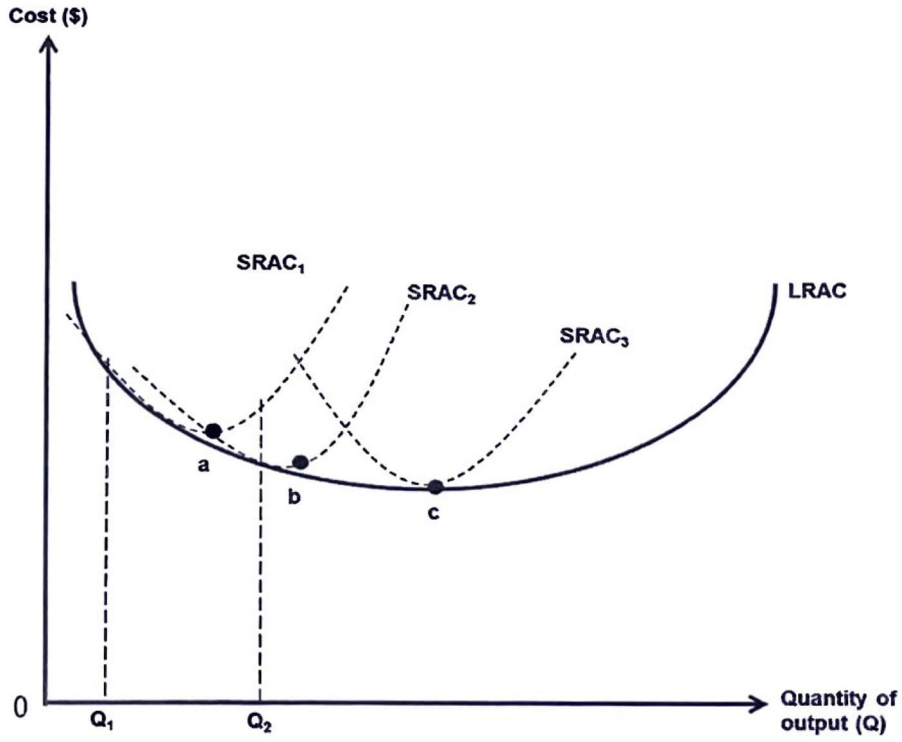


Figure 3 The relationship between SRAC and LRAC

- The LRAC curve is commonly known as a firm's **planning curve** because on the basis of this curve, the firm **decides the size and scale of its plant** *(including land,*

buildings, factories and heavy machinery) in order to produce a given level of output at the lowest possible cost.

Diagrammatic illustration

Let's assume that a farmer only produces with two inputs: land and labour. The farmer would like to expand his production and considers his long-run options.

- To simplify our analysis, let's assume that there are only 3 possible farm (or plant) sizes.
- Each farm size, with its own particular level of the fixed input of land, is represented by the curves $SRAC_1$, $SRAC_2$ and $SRAC_3$ in Figure 3.
- Assuming that the farmer is initially producing at Q_1 on $SRAC_1$, but would like to increase production to Q_2. The farmer has 2 options:
    - Retain his current farm size by remaining on $SRAC_1$ in the short run
    - Increase the size of his farm and produce on $SRAC_2$, producing Q_2 at a lower average cost (due to cost benefits arising from economies of scale)
- As the farmer increases output in the short run, it will be possible only to produce at the lowest possible cost on $SRAC_1$ up to **point a** (minimum point of $SRAC_1$). At this point, the farmer should consider increasing the size of his farm (going into the long run) and move onto $SRAC_2$.
- If the decision is made to increase the farm size and move onto $SRAC_2$, output can continue to increase to approach the lowest possible cost until **point b** is reached, where the farmer once against should consider increasing farm size (going into the long run again) and switching to $SRAC_3$.
- From this succession of an infinite number of SRAC curves we can construct a LRAC curve which *envelopes* the SRAC curves. At any output level, the LRAC curve is *tangent* to each of the SRAC curves (not necessarily the minimum point of the SRAC curve). It represents the lowest possible average cost for each level of output, given that all resources are variable in the long run.
- The points of tangency between LRAC and SRAC curves do not occur at the minimum points of the SRAC curves except at the point where the minimum efficient scale (MES) is achieved.