CHAPTER 4A (H2 only)

Firms and Decisions: Objectives of Firms and Costs of Production

OVERVIEW OF CHAPTER

To economists, production is the transformation of inputs into outputs that yield satisfaction directly to the consumers. The transforming agent is the firm as it produces and supplies goods and services to consumers.

As the firm has limited resources and cannot produce everything, it has to choose the type of goods and services to produce. The firm must also decide how to produce these goods and services. It has to decide the factor combinations (e.g. more labour and less capital, or less labour and more capital) to be used in the production process. For a self-interested (assuming profit motivated) firm, it will obviously want to use the least cost combination of factors to produce a given level of output i.e. to achieve productive efficiency.

When a firm wishes to produce more, it has to increase the quantity of factor inputs. However, by purchasing more factor inputs, the total costs of production of the firm will rise. In addition, when the firm offers more for sale, the price would also have to fall, thus affecting its revenue. All these will affect the firm's profit and thus its production decision.

As a result, a sound knowledge of a firm's production behaviour and cost of production in Chapter 4A sets the groundwork for the study of Spectrum of Market Structures (Chapter 4B) – which considers how a firm's price and output decisions are affected by the characteristics of the market structure and market conditions within which it operates, and the strategies firms (the firm and its competitors) can adopt.

KEY DECISIONS RELATED TO THIS CHAPTER

- A firms' price and output decisions based on the marginalist principle, maximising profit when marginal revenue equals marginal costs, and on the estimated demand for its good or service.

- A firm's decision with regard to its scale of production or size and growth of the firm.

- A firm's decision with regard to the objectives it wants to pursue in the short run and in the long run.

KEY ECONOMIC AGENTS

- Producers (Firms)
- Consumers

1. FIRM, PLANT AND INDUSTRY

It is important to distinguish between some of the terms that will be used in this chapter.

DEFINITIONS:

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.

2. TRADITIONAL OBJECTIVE OF FIRMS: PROFIT MAXIMISATION

Learning Objectives

- Define and derive total revenue, total costs and profit.
- Distinguish between implicit and explicit costs.
- Distinguish between accounting profit and economic profit.
- Using the marginalist principle to explain how firms make price and output decisions in order to maximise profits.

In Chapter 2A, we assumed the supply curve is upward sloping: that a higher price will encourage firms to offer more for sale. But just how much will firms choose to supply at each price? It depends largely on the amount of profit they will make. If a firm can increase its profit by producing more, it will normally do so.

Economists typically assume that the objective of any firm is to maximise profit. Profit maximisation can be defined as striving for the highest profit. To determine the profit maximising output, we need to know the concepts of revenue, costs and profit.

2.1 TOTAL REVENUE, TOTAL COSTS AND PROFIT

DEFINITIONS & FORMULAS:

Total Revenue (TR) of a firm is the firm's total earnings from the sale of its output for a specific period of time.

Total revenue is calculated as the price (P) multiplied by the number of units sold (Q).

$$TR = P \times Q$$

Total Costs (TC) of production to the economist are technically the opportunity costs of using all the resources in the production of a product.

Total costs are calculated as the sum of explicit costs and implicit costs.

$$TC = \text{Explicit Costs} + \text{Implicit Costs}$$

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

$$\textbf{Profit } (\pi) = TR - TC$$

2.2 ACCOUNTING PROFIT AND ECONOMIC PROFIT

The total costs of production to the economist are technically the opportunity costs of using the resources in the production of a product – a measure of the sacrifice involved in using factors of production to produce that product. This definition is based on the fact that resources are scarce and they have alternative uses. To produce a product, scarce resources must be diverted from other productive activities. To use them in one process is to deny their use in other processes. The costs may be

implicit and hence need to be imputed. 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 already owned by the firm. The firm does not have to pay out money to use them, however a cost is still incurred by the firm as the firm could have benefitted from the resource if it used it for another purpose. For example, if a firm owns a building, and could have lease it to others for a rent of $5,000 per month. If the firm chose to use the building for its own business, then it forgoes the rent income of $5,000 per month, thus although it does not pay for the use of its own building, it still incurs a cost of $5,000 per month called opportunity cost.

Explicit costs are the opportunity costs of using resources that are not already owned by the firm, and the firm has to pay the factor owners to use them.

Accountants consider only the explicit costs incurred during production but economists consider both explicit as well as implicit costs hence there is a difference between accounting profit and economic profit.

In economics, profit is referred to economic profit (as opposed to accounting profit). The main difference between economic profit and accounting profit is that economic profit takes into consideration implicit costs.

An illustration of the difference between accounting and economic profit is shown in Table 1.

Table 1: Calculation of Accounting and Economic Profits

Total Revenue		**$75,000**
Less Explicit Costs:		
Worker's salary	**$15,000**	
Materials	**$18,000**	
Rental	**$2,000**	
Total Explicit Costs		**$35,000**
Accounting Profit		**$40,000**
Less Implicit Cost:		
Owner's forgone salary in his next best job		**$30,000**
Economic Profit		**$10,000**

Economic profit may either be positive, zero or negative.

NOTE:

Concepts of short run and long run will be discussed in later sections.

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 level 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.

To have a better understanding of how a firm goes about maximising its profit, we must know exactly how to measure its total revenue and total cost. It is relatively easy to calculate total revenue but to calculate total cost, it is more subtle. Before the various concepts of costs are touched on in more details, let's go through the concept of profit maximisation first.

THINK ABOUT IT

Why does the concept of economic profit facilitate better decision making than accounting profit?

Imagine that you could venture into a business, and you knew that you would earn an accounting profit of $100,000 per year. Would you venture into this business?

2.3 PROFIT MAXIMISATION AND THE MARGINALIST PRINCIPLE APPROACH

Recall that all firms aim to maximise their profits. There is a very simple profit maximising rule: in order to maximise profits, marginal revenue must equal marginal cost.

$$\mathbf{MR = MC}$$

DEFINITIONS & FORMULAS:

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

$$\mathbf{MR = \frac{\Delta TR}{\Delta Q}}$$

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

$$\mathbf{MC = \frac{\Delta TC}{\Delta Q}}$$

NOTE:

Though we do not make use of average revenue to help us determine the profit maximising level of output, it is useful to understand the Average Revenue concept as it will be relevant when we study Market Structures in Chapter 4B.

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

$$\mathbf{AR = \frac{TR}{Q}}$$

*When all output is sold at the **same** price as in the case of firms in perfectly competitive market structure, average revenue will be the same as price i.e., TR = P x Q thus TR/Q = P = AR.*

The rational decision of a profit-maximising firm is to continue producing additional units of the good/service as long as the addition to total revenue (MR) is greater than the addition to total costs (MC) at the current level of output. On the other hand, the profit maximising firm will not produce the additional unit of the good if the addition to total revenue (MR) is less than the addition to total costs (MC) at that additional unit of output. Therefore, the profit maximising point is where MR=MC. This reasoning is the same as using the marginalist principle to make rational decision taught in Chapter 1.

3. COSTS OF PRODUCTION IN THE SHORT RUN

Learning Objectives

✳ Distinguish between fixed and variable factors.

✳ Distinguish between the short run and long run.

✳ Aware of how the Law of Diminishing Returns affects output in the short run.

✳ Distinguish between fixed and variable costs.

3.1 FIXED & VARIABLE FACTORS AND THE SHORT RUN

Factors of production can be separated into two categories: fixed factors and variable factors.

DEFINITIONS:

Fixed Factors / Inputs are factors of production whose quantities **cannot be easily changed** or **cannot be changed** (with a given short period of time). Thus, firms cannot increase these factors to increase output in the short run.

Examples of fixed factors include factory buildings and heavy machinery.

Variable Factors / Inputs are factors of production whose quantities **can be easily changed** within a short period of time, such that more of these factors can be employed in order to increase the level of output, vice versa, even in the short run.

Examples of variable factors include labour/workers and materials.

In economics, the production period is divided into the short run and the long run. These time periods are defined in terms of the time required to alter the quantity of all factors of production employed and not calendar time. As long as there is at least one factor of production that is fixed, the firm is in the short run. For example, if it takes 10 months for all the factors to be varied, then the long run is 10 months and more.

DEFINITION:

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.

Let us consider the following example of a baker selling bread. He will need the following to manufacture bread:

- Raw materials such as flour, eggs, butter, sugar and yeast
- Labour/Workers
- Machinery
- A factory

Suppose the demand for bread has greatly increased, prompting him to produce more bread. He should be able to order more raw materials with little delay, so we consider raw materials to be a variable factor. He will need extra labour and he can either hire more workers or get his existing workers to work longer (more man-hours), so this is also a variable input. The machinery on the other hand, may not be a variable factor. It may take time to acquire and implement the use of additional machinery, especially those that require specific customisation. It depends on how long it would take him to buy and install the machinery and how long it would take him to train the workers to use it. Adding an extra factory is certainly not something he could do in a short period of time as it takes time to build the factory, so these would be fixed factors.

In the short run, he can increase the production of bread by adding more raw materials and more labour (both variable factors) to existing machines and factory (both fixed factors). Until the time when a whole new factory is built and more machinery are ready for production, the firm will still be operating in the short run. Hence, in the short run the firm's production decision is constrained by a fixed maximum capacity (e.g. the full productive capacity of the current factory).

The short run therefore does not correspond to a specified number of years. It varies from firm to firm and from industry to industry depending on how long it takes to change all the inputs involved.

3.2 SHORT RUN COSTS

The higher the price of the factors of production, the higher will be the total costs of production holding other things constant. In our analysis of the short-run costs of production, it is assumed that, among other things, prices of factors of production are held constant.

Recall that 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.

3.2.1 TOTAL FIXED COSTS, TOTAL VARIABLE COSTS AND TOTAL COST

DEFINITIONS:

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

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

Total Cost is the sum of total fixed costs and total variable costs at a specific level of output. TC = TFC + TVC.

NOTE:

Some costs of production can be considered to be a fixed cost in one industry but considered a variable cost in another industry e.g. generally, cost of furniture is considered a fixed cost for an office dealing with financial services but is considered a variable cost for a furniture dealer.

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. Fixed costs will be incurred even if the firm temporarily shuts down and stops production (i.e. when output is zero) in the short run. Such costs are often referred to as overheads and are beyond the firm's control in the short run.

Examples of fixed costs include:

- Depreciation of machines due to obsolescence and ageing (not due to usage)
- Fire insurance, property taxes, rentals
- Interest payments (e.g. for instalment contracts) on loans for the purchase of capital goods (e.g. factory and equipment)

NOTE:

In the short-run, fixed costs are already paid up and is not a consideration for the firm when making price and output decisions as these costs cannot be recovered by the firm even if the firm ceases to produce temporarily. Rational decision is made at the margin (refer to Chapter 1 Section 1.4) and thus only marginal costs and marginal benefits affect decision making. Referring to, for example, the formulas for marginal costs in Section 3.2.3, marginal costs are independent of fixed costs. Thus only variable costs are considered in the pricing and output decisions of the firm in the short-run (This will be further covered in Chapter 4B).

Derivation of the cost curves is not required in the syllabus.

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 since the firm needs more variable factors to increase output in the short run, vice versa. **Variable costs will be zero if the firm ceases production.** Firms can control variable costs in the short run by changing the level of output.

Examples of variable costs include:

- ✳ Costs of raw materials and fuel
- ✳ Wages of production workers

3.2.2 AVERAGE FIXED COST, AVERAGE VARIABLE COST AND AVERAGE COST

DEFINITIONS:

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

$$\mathbf{AFC} = \frac{\mathbf{TFC}}{\mathbf{Q}}$$

Average Variable Cost is the variable costs per unit of output.

$$\mathbf{AVC} = \frac{\mathbf{TVC}}{\mathbf{Q}}$$

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

$$\mathbf{AC} = \frac{\mathbf{TC}}{\mathbf{Q}}$$

Alternatively, since TC is given as the sum of TFC and TVC, and AC is given as TC dividing by Q, we have:

$$\mathbf{AC = AFC + AVC}$$

3.2.3 MARGINAL COSTS (MC)

(Recap) Marginal Cost (MC) is the additional cost of producing one more unit of output.

$$\mathbf{MC} = \frac{\Delta \mathbf{TC}}{\Delta \mathbf{Q}}$$

Because only variable costs change with output (i.e. fixed costs do not vary with output), MC can also be expressed alternatively as the change in total variable costs divided by the change in output.

$$\mathbf{MC} = \frac{\Delta \mathbf{TVC}}{\Delta \mathbf{Q}}$$

3.2.4 ILLUSTRATION OF COSTS IN A DIAGRAM

NOTE:

Illustration of the costs in a diagram (know how to draw the cost curves) ***is required*** *but explanation for the derivations/reasons for the shapes of the curves is* ***not required*** *for the syllabus.*

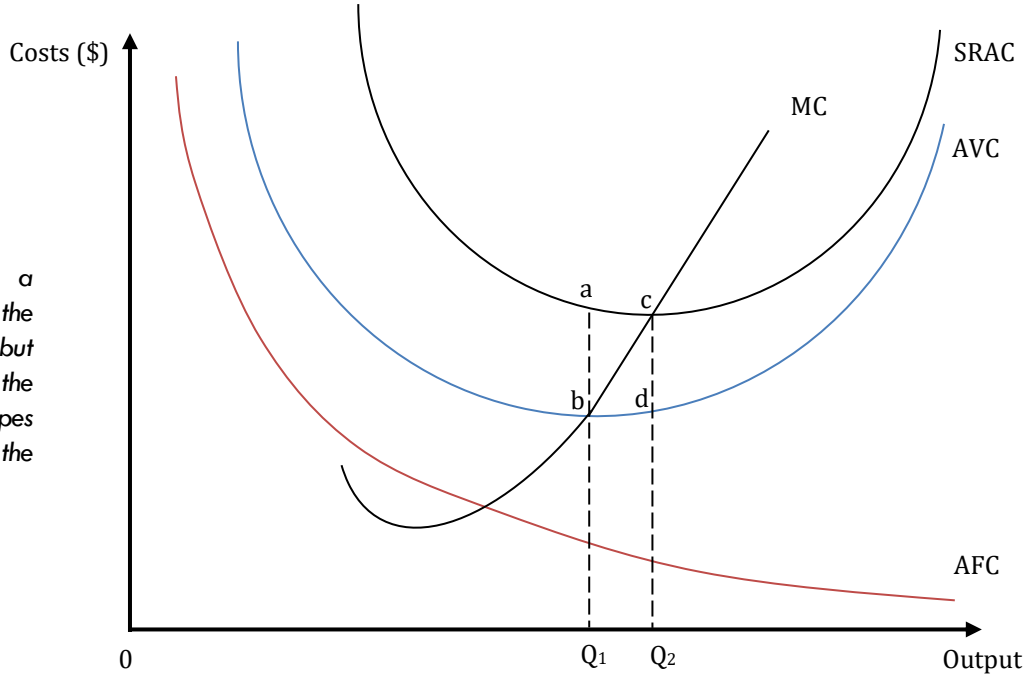


Figure 1: Short Run Cost Curves for a Typical Firm

Illustrating the costs in a diagram is important for this syllabus. Some important points to note when drawing these cost curves (explanations of the shapes of the curves can be found in the respective boxes but are for awareness only):

- ✳ The average fixed cost (AFC) curve is downward sloping and approaches the horizontal axis as output increases, but it will not meet the axis.

- ✳ The vertical distances between the SRAC and AVC curves are the AFCs at each corresponding output (eg distance ab and cd are the AFCs at Q_1 and Q_2 respectively). The AVC curve approaches the SRAC curve as output increases, but they will not intersect (or the gap between SRAC and AVC narrows as output increases but they will not meet).

- ✳ The marginal cost (MC) curve is U-shaped and cuts the short run average costs (SRAC) curve and the average variable costs (AVC) curve at their minimum.

- ✳ Both the SRAC and AVC curves are U-shaped.

- ✳ Both SRAC and AVC fall initially as output increases but at a particular output, AVC starts to increase while SRAC continues to fall. As output continues to increase, SRAC will eventually increase too.

NOTE (Explanation not required for the syllabus):
The average fixed cost (AFC) curve is downward sloping and approaches the horizontal axis as output increases, but it will not meet the axis

Since TFC is constant, thus a constant TFC divided by an increasing output results in AFC continuously declining as output increases. This is because TFC (which is constant) is being spread over a larger and larger output. AFC will never be zero as long as TFC is not zero, thus AFC curve will not meet the horizontal axis.

NOTE (Explanation not required for the syllabus):
The vertical distances between the SRAC and AVC curves are the AFCs at each corresponding output. The AVC curve approaches the SRAC curve as output increases, but they will not intersect.

The vertical distances between the SRAC and AVC curves are the AFCs at each corresponding output because SRAC = AVC + AFC. Since the vertical distance between the AVC and SRAC curves is the AFC, and AFC falls as output increases, the gap between SRAC and AVC will

be smaller as output increases, or the gap narrows as output increases because fixed costs get spread over larger and larger output levels.

They will not intersect as long as total fixed cost is positive (positive TFC ÷ Q will always be positive, regardless of how large Q is).

NOTE (Explanation not required for the syllabus):
The marginal cost (MC) curve is U-shaped.

The marginal cost first decreases and then increases as output increases. Changes in marginal cost reflect changes in the extra output produced by hiring one additional variable factor (eg labour). Marginal cost falls initially as output increases because productivity increases when the fixed factor is better utilised by the increasing quantity of variable factor employed. As output continues to increase with more variable factors employed, the law of diminishing returns sets in, thus marginal cost will increase. This results in the U-shaped marginal cost curve.

Refer to Annex to know more about the law of diminishing returns.

NOTE (Explanation not required for the syllabus):
The marginal cost (MC) curve cuts the short run average costs (SRAC) curve and the average variable costs (AVC) curve at their minimum.

The shape of the short run average cost curve (SRAC) is dependent on the shape of the marginal cost curve (MC). As long as the cost of producing new units of output is less than the average, their production must pull the average cost down. That is, if MC is less than SRAC, SRAC must be falling. Likewise, if new units cost more to produce than the average, their production must drive the average cost up. That is, if MC is greater than SRAC, AC must be rising. Therefore, the MC curve intersects SRAC curve at its minimum point.

Since all marginal costs are variable, the same relationship holds between MC and AVC.

NOTE (Explanation not required for the syllabus):

Both the SRAC and AVC curves are U-shaped.

The reason why AVC is U-shared is the same as why the MC curve is U-shaped (law of diminishing returns). AVC falls initially as output increases because productivity increases when the fixed factor is better utilised by the increasing quantity of variable factor employed. As output continues to increase with more variable factors employed, the law of diminishing returns sets in, thus AVC will increase.

Since SRAC is derived from AFC and AVC (SRAC = AFC + AVC), it follows the shape of AVC but higher.

NOTE (Explanation not required for the syllabus):

Both SRAC and AVC fall initially as output increases but at a particular output, AVC starts to increase while SRAC continues to fall. As output continues to increase, SRAC will eventually increase too.

As the firm increases its output before Q1, both AFC and AVC fall; thus SRAC must fall. Though AVC reaches minimum at Q_1 and begins to rise after Q_1, AC continues to fall because the falling AFC more than offsets the rising AVC.

As the AFC curve gets less steep, the gap between SRAC and AVC narrows. As a result, SRAC will start to rise beyond Q_2. This is due to the fact that the rising AVC more than offsets the still falling AFC.

THINK ABOUT IT:

Fill in the blanks for the various short run costs of a typical firm. Assuming the wage rate is $100 per unit of labour.

Output	TFC	Labour	TVC	TC	AFC	AVC	AC	MC
0	200	0	0	200	-	-	-	-
2		1	100		100.00	50.0	150.00	50.00
5		2			40.00	40.00		33.33
9		3			22.22			25.00
12		4						33.33
14		5						
15	200	6	600	800	13.33	40.00	53.33	100.00

Table 3: Short Run Cost Data for a Bakery

4. COST OF PRODUCTION IN THE LONG RUN

Learning Objectives:

- Distinguish between long run average costs and short run average costs.
- Explain productive efficiency.
- Explain returns to scale.
- Explain Internal and External Economies and Diseconomies of Scale.

DEFINITION:

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.

In the long run, **all** factors of production are variable. In the LR there is time for the firm to build a new factory (maybe even in a foreign country), to install new machines, to use different techniques of production, and in general to combine its input in whatever proportion and in whatever quantities it chooses. It is no longer constrained by any fixed factor.

Recall: In the short run, if the bakery wishes to increase production, it can only hire more variable factors e.g. workers.

But in the long run, if the bakery wishes to increase production, besides hiring more variable factors, he can choose what techniques of production to use, how many more factories to build, what types of machines to buy, how to organise the factory – whether to use automated process or more labour-intensive techniques.

By being able to change the fixed factor(s), e.g. the size of land and amount of capital, the firm has moved from a short run to a long run situation, however the firm must be very careful in making these decisions. Once it has built its factory and installed the machinery, these then become fixed factors of production, maybe for many years. The subsequent 'short run' time period may in practice last a very long time!

Another way to understand the long run is to consider it as the time period for which firms plan ahead to build the most appropriate scale of plant to produce the future anticipated level of output. Once the plant has been built, the firm operates in the short run again. Similar to the SR, the LR does not correspond to a specified number of years. It varies from firm to firm and from industry to industry depending on how long it takes to be able to change **all** inputs used.

The profit maximising firm will obviously want to use the least cost combination method of production in order to minimise cost at any given output level. This will allow the firm to achieve **productive efficiency at the firm level**.

In reality, whenever a firm buys new machinery, the machinery is likely to incorporate newer technology. However in theory, we **assume technology to be constant** even in the long run analysis because the impact of technological advancement is hard to predict and the primary concern here is to study the effects of increasing and decreasing scale of production in isolation. If we want to factor in technology changes, the time period is usually referred to as the **very long run** instead.

As all factors of production are variable, the law of diminishing returns does not apply. Instead, production in the long run is subjected to **economies and diseconomies to scale**. The long run also encompasses enough time for existing firms to leave or for new firms to enter the industry.

NOTE:

*Explanation of how the long run average cost curve is derived from the SRAC curves is **not required** for the syllabus.*

4.1 FROM THE SHORT RUN TO THE LONG RUN: LONG RUN AVERAGE COST (LRAC) CURVE

Recall that firms can increase the number or size of plants (fixed factors in the short run) in the long run, thus different number of plants or plant sizes are available to a firm. The firm can choose and switch to the number of plants or plant size that produces a given level of output at the lowest possible cost. Thel long run average cost (LRAC) curve indicates the least costly way to produce any given level of output i.e. any point on the LRAC curve is productively efficient.

The long run average cost curve is derived from the short run average cost curves for the various scales of production (eg different number of plants).

All the points on the LRAC are productive efficient from the firm's perspective as these points represents the lowest possible costs for the combination of FOPs corresponding to the respective levels of output. Recall that productive efficiency is producing the highest level of output with a given level of FOPs, or in another words, producing a given level of output with the least amount of FOPs, thus lowest average cost.

The LRAC curve, like the SRAC curve, is also a U-shaped curve because of the internal economies and diseconomies of scale.

NOTE (Explanation not required for the syllabus):

Derivation of the long run average cost (LRAC) curve from the short run average cost curves

For every LRAC curve drawn, the following key assumptions are made:

- ✴ Factor prices are given i.e. constant.
- ✴ The state of technology and the factor quality are given.
- ✴ Firms choose the least-cost combination of factors for each output

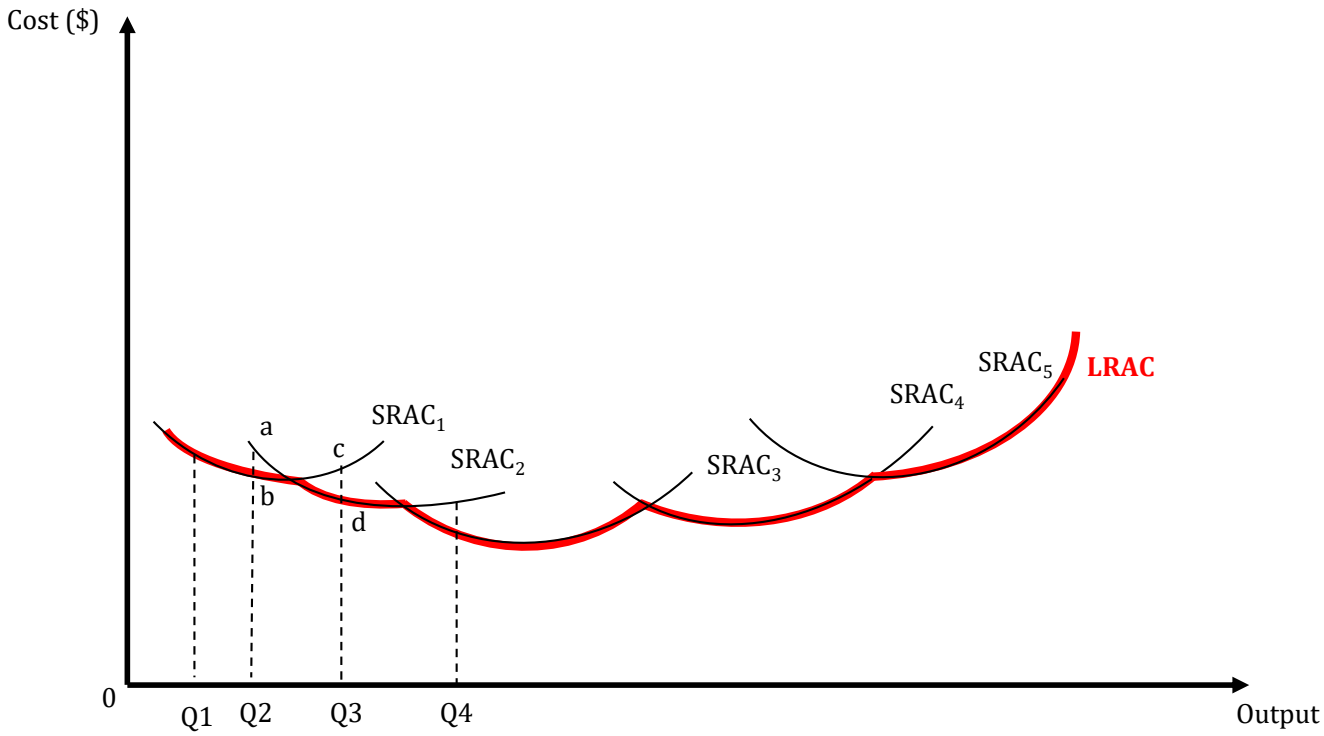


Figure 2: Derivation of the Long Run Average Cost Curve

Each SRAC curves represents the average costs at each scale of production. For example, SRAC1 corresponds to the average costs for the various output levels for 1 plant with respective levels of labour; SRAC2 represents the average costs for the various output levels for 2 plants with respective levels of labour; SRAC3 represents the average costs for the various output levels for 3 plants with respective levels of labour, and so on.

When deciding the combination of labour and capital (eg plant) for a given output, a firm will consider the lowest cost combination. For example, at Q1 the firm will choose to have 1 plant with the given level of labour to produce that output. If the firm decides to increase production to Q2, the firm will still choose to have 1 plant as the average cost of 2 plants with the same level of labour (Point a) is higher than for 1 plant (Point b). However, if the firm decides to produce at Q3, it will use 2 plants as the average cost (Point d) will be lower than if it uses 1 plant (Point c).

This decision-making process continues with every level of output (eg Q4) and thus all the lowest cost combinations will be the LRAC curve. This is why the long run average cost (LRAC) curve indicates the least costly way to produce any given level of output as it represents the lowest average cost of production at all possible levels of output.

It is usual, however, to show the LRAC curve as a smooth line tangential to the SRAC curves (as shown in Figure 3 below). This is known as the envelope curve. The assumption here is that the short-run fixed factors can be varied by any amount in the long run. For example, plants of any size can be built, or existing ones can be expanded. The result is that there will be an unlimited number of SRAC curves.

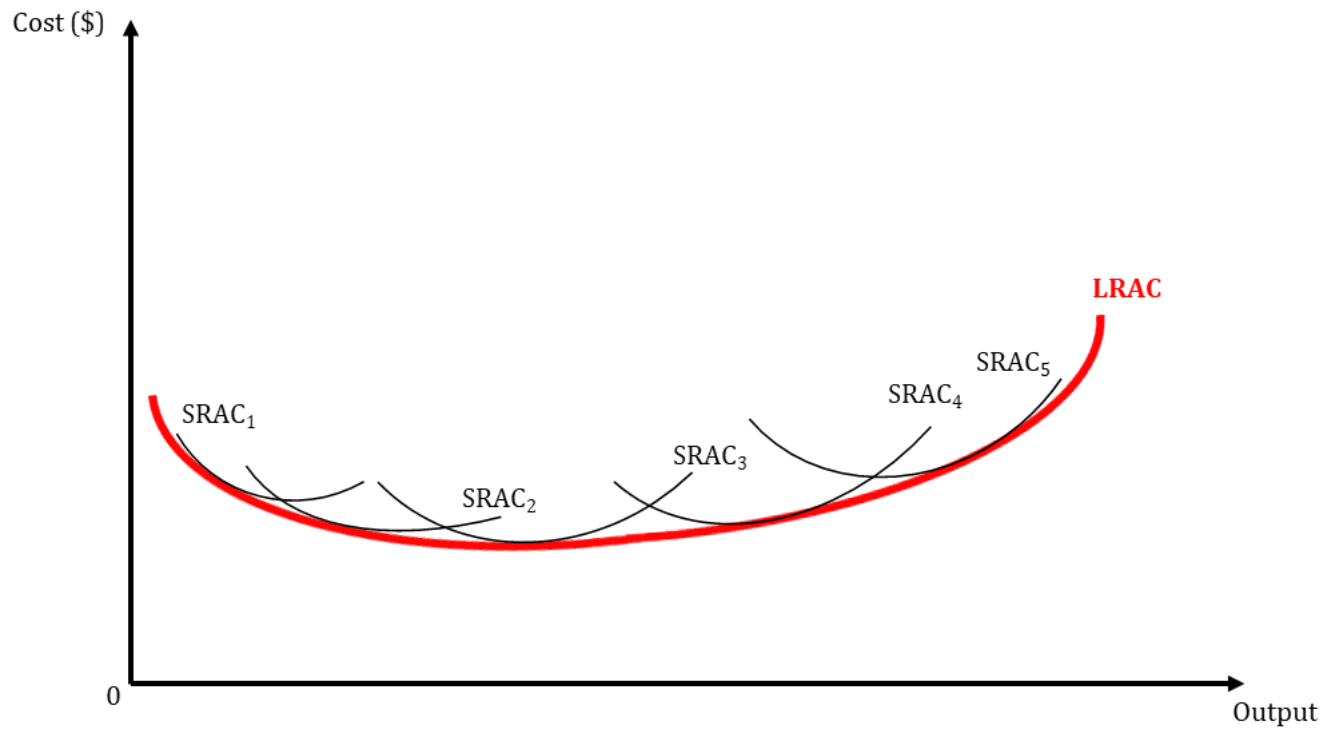


Figure 2: Derivation of the Long Run Average Cost Curve

NOTE (Explanation not required for the syllabus):

Least cost factor combination method of production and productive efficiency from firm's perspective

Firms can choose various combination of FOPs (for example, using 1 plant with 20 workers or 2 plants with 12 workers) to produce the same level of output in the LR. A rational firm will choose a productively efficient method of production which is also the least cost combination method to produce any given output level.

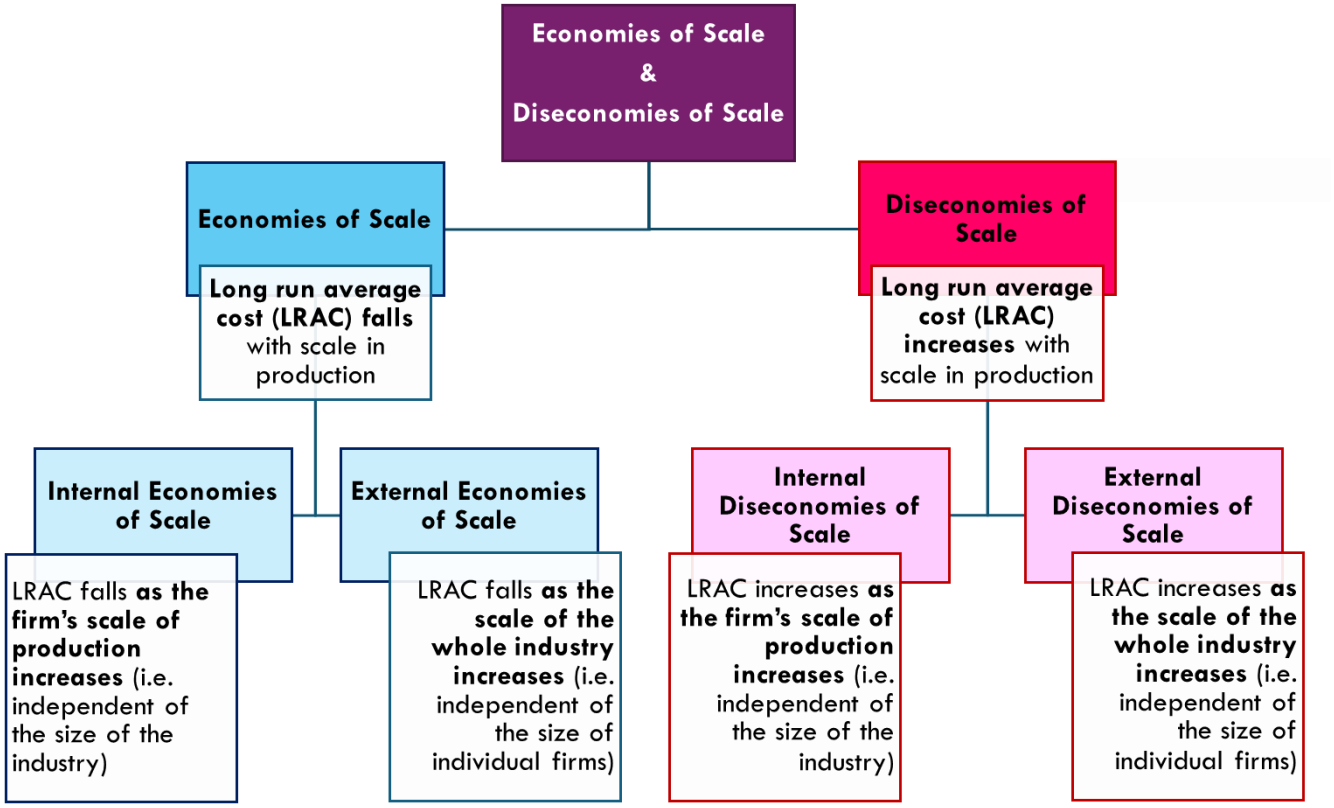
Method	Output	Labour	Capital	Unit Price of Labour ($)	Unit Price of Capital ($)	Total Cost ($)	Cost Per Unit of Output ($)
A	10	2	4	10	50	20 + 200	22
B	**10**	4	3	10	50	40 + 150	19
C	**10**	**5**	**2**	**10**	**50**	**50 + 100**	**15**

Table 4: Long Run Costs of a Firm with Different Factor Combinations

From the table, we observe that cost per unit of output produced is the lowest for method C. Therefore, method C is the productively efficient method as it makes use of relatively more of the cheaper input and relatively less of the more expensive input to produce the same output.

4.2 ECONOMIES AND DISECONOMIES OF SCALE

Economies of scale are experienced when the long run average costs (LRAC) fall with scale in production, while diseconomies of scale are experienced if long run average costs (LRAC) increases with scale in production. The scale in production may be from the firm (ie the firm increases its scale of production) or from the industry (ie the industry experiences scale of production).



4.2.1 INTERNAL ECONOMIES AND DISECONOMIES OF SCALE

Internal Economies of Scale

DEFINITION:

Internal Economies to 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 ***by an individual firm.***

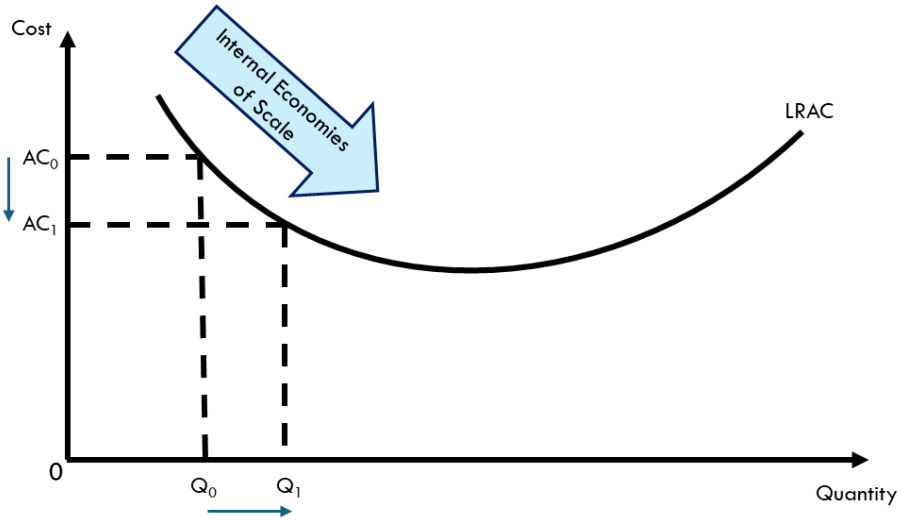


Figure 4: Internal Economies of Scale

Internal economies of scale are graphically represented by the downward movement along the LRAC curve. As the firm increases its scale of production from Q_0 to Q_1, there are cost savings (or economies of scale) which results in a fall in average cost from AC_0 to AC_1.

NOTE:

Due to the differences in the nature of production processes of different industries, producers enjoy different types of economies of scale.

There are several sources of internal economies of scale:

a. Technical Economies of Scale

Technical economies of scale is the cost saving gained from the production process because of the firm's larger scale production. There is a reduction in average costs when a firm expands its plant size. The following are examples of technical economies of scale.

✳ Division of Labour and Specialisation

As production expands, there is greater scope for division of labour leading to specialisation. Division of labour allows the production processes to be broken down into simpler processes and specialisation to be possible thereby raising productivity and lowering average costs.

By performing the same task, specialised labour can become very skilled and perform with speed and dexterity with repetition of the task, thus workers will be more productive. Less time is also lost by workers switching from one task to another. This enable more output to be produced with a given amount of inputs or cost. Training costs can also be scaled down as workers only needed to be trained in their specialised areas.

✳ **Utilisation of Technology and Capital, and Indivisibilities**

The use of technology and capital, which can enhance productivity, may only be justified with larger scale production because of their high startup or sunk cost. A larger firm will thus be in a better position to use such technology and capital, thus gaining from the cost saving from higher productivity.

Certain types of technology and capital equipment also come in a fixed productive capacity and they may be necessary even when the firm is producing a small output and thus operating below capacity i.e. underemployed equipment. A firm producing on a larger scale can optimally utilise such indivisible capital equipment and thus will be able to spread the total cost over larger output levels, thereby lowering average costs, compared to a firm producing on a smaller scale of production.

b. Marketing Economies of Scale

A large firm usually requires large amount of raw materials, thus can capitalise on this bargaining power to buy its raw materials in bulk at more favourable rates than a smaller firm. Similarly, the products of the firm can also be sold in bulk at reduced distribution costs too. For instance, it is more cost efficient for a large firm to transport larger quantities using a large truck instead of several small vans like a smaller firm. Large firms can also afford to advertise in the national press and other forms of media. Although the advertising expenditure may be substantial, the increase in average cost from advertising may be lower than that of a smaller firm because the total cost of advertising is spread over a larger output level.

c. Financial Economies of Scale

A large firm, with a higher sales volume and more valuable assets to offer as collateral is often able to obtain finance more easily and on more favourable terms than a smaller firm. Banking and financial institutions are more willing to offer loans or extend credit to large firms as larger firms with more collateral are considered to have lower risk. In addition, large firms also usually enjoy better terms (lower interest rates) when they borrow large amounts. Even a small reduction in interest rates charged on borrowing yields substantial savings where relatively large sums are involved. Besides better credit standing with the institutions,

larger firms are more likely to be limited companies that can raise capital easily through the issue of shares to the public. This is a cheaper alternative to borrowing from the banks as no interest payments are incurred and this also results in cost saving to the larger firm.

d. Organisational Economies of Scale

With a large firm, individual administrative departments can specialise in particular functions. There can also be centralised administration of the firm; for example, one human resources department could administer all the wages. This is especially true when 2 firms merge since only one finance department and one human resource department is now needed instead of two. Hence average cost savings can be enjoyed when such rationalisation takes place.

e. Managerial Economies of Scale

As a firm grows, there is greater potential for its managers to specialise in particular tasks (e.g. marketing, human resource management, finance). With greater specialisation and expertise in these areas, these managers are expected to be more effective in managing the business system and ensuring productive, thus reducing average costs.

f. Risk Bearing Economies of Scale

A large firm can diversify its risks more efficiently than that of a small firm due to its scale of operations and is more likely to have enough resources to deal with a potential contingency. This is especially true for large firms that may operate in several markets, where losses in one market can be easily overcome by gains in other markets. This reduces the average costs associated with uncertainty. For example, if there is a sudden fall in demand in one market (e.g. China), a large firm which is able to sell the excess stocks in another market (e.g. India) may not have to incur high storage costs. However, a smaller firm which may not have as extensive a network of distributors may not be able to divert their stocks quickly. Therefore, the larger firm, with a larger output, will likely incur a lower average cost.

Internal Diseconomies of Scale

DEFINITION:

Internal Diseconomies of Scale refers to the increase in long run average costs of production as a result of the expansion of the scale of production of a particular output/product ***by an individual firm.***

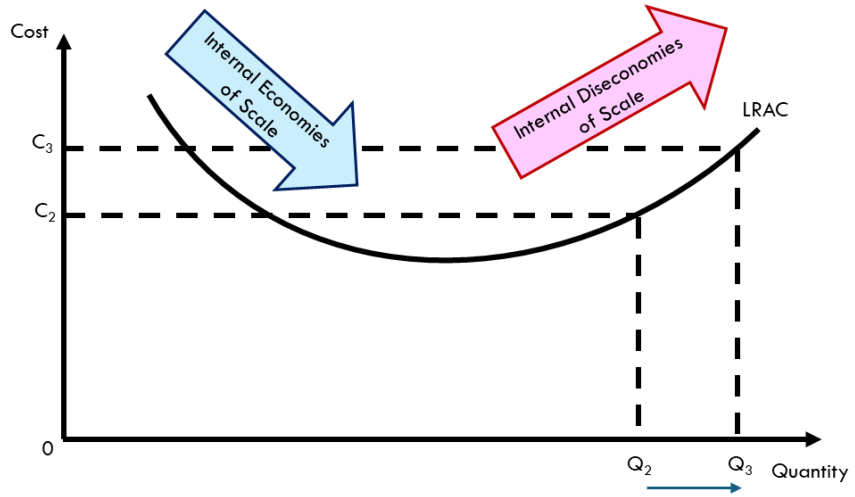


Figure 5: Internal Diseconomies of Scale

Internal diseconomies of scale are graphically represented by the upward movement along the LRAC curve. As the firm increases its scale of production from Q_2 to Q_3, there are cost "dis-saving" (or diseconomies of scale) which results in an increase in average cost from AC_2 to AC_3.

When firms expand beyond its optimal size, its average costs of production of the output may start to increase. As the size of the firm increases, management becomes more complex. It becomes increasingly more difficult to carry out the various management functions and thus results in inefficiencies and increases in average costs. The sources of internal diseconomies of scale are identified as follows:

a. Loss of Control

A large firm employing a large number of workers will face difficulties ensuring everyone is happy and equally well treated. Although large firms usually have an impressive hierarchy of authority, in practice, the problem of monitoring work performance is a very difficult task. The loss of control could mean that some workers may be underworked or not working as productively as they could. There could also be more wastage and pilferage as it is difficult to monitor all the workers all the time. Small-

to medium-sized firms often report that workers feel they 'make a difference; this may be lost in a larger firm and as a consequence lower motivation may lead to lower productivity, and thus higher average costs.

b. Lack of Coordination & Communication

Large firms are usually subdivided into many decentralised departments such as production, planning, sales, purchasing, personnel, and accounts. As these departments multiply and grow in size, a bureaucratic structure gradually takes shape and the tasks of coordinating their activities and communication becomes increasingly difficult. This leads to inefficiency and higher average cost of production.

Even within a plant, mass production processes usually involve complex interdependencies which can lead to great disruptions in output if there are any delays in just one section of the plant. Such bottlenecks in production cause a reduction in efficiency as longer time is taken to produce the same output, resulting in higher average cost.

As a firm increases in size, communication lines grow longer and as a result communication may breakdown or there may be misinterpretation of information. All these can cause delays in production which can be costly to the firm.

Minimum Efficient Scale

DEFINITION

Minimum Efficient Scale (MES) refers to the lowest level of output at which the long run average cost is at its minimum.

Minimum efficient scale (MES) occurs at the output level where LRAC stops falling. It corresponds to the lowest point on the firm's long run average cost curve. With reference to Figure 6, it is at output Q_4. The MES is the size of the firm, where further expansion will not lead to any further economies of scale.

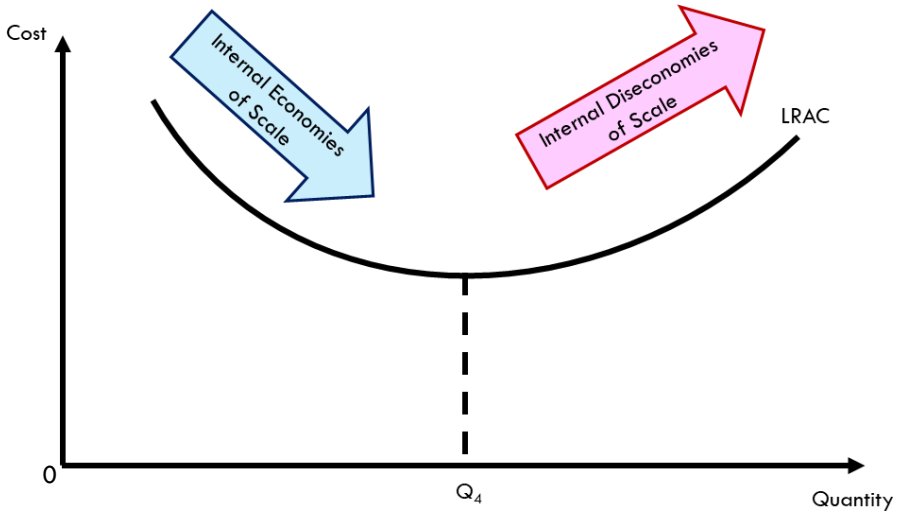


Figure 6: Minimum Efficient Scale

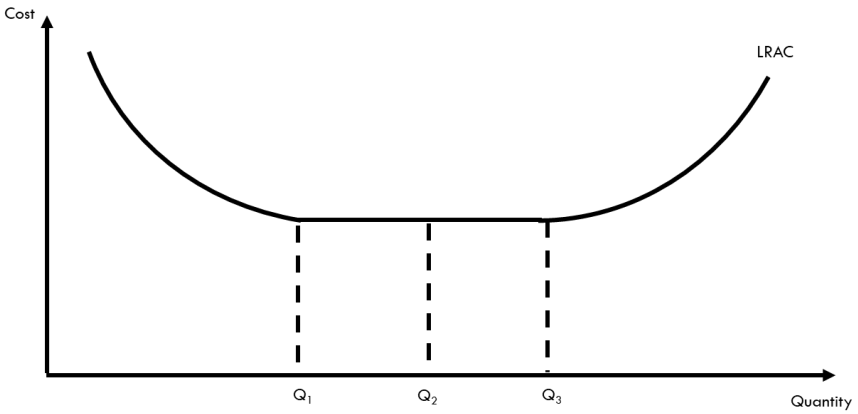
The MES varies from industry to industry depending on the nature of the cost structure in a particular sector of the economy. In general, when the cost of capital investment required is very high, there is greater potential for reducing the average cost of production as there is higher cost of capital equipment to spread around. For such industries, the MES will occur at a high level of output.

NOTE:

The concept of natural monopoly will be touched on in the next sub-chapter.

The MES is a very important indicator of possible firm size and the type of market structure that exists in different industries. An industry that requires very large output to achieve MES (eg extremely large sunk cost) may not be able to support a few firms competing with one another, thus a natural monopoly might be created.

THINK ABOUT IT



At which output level, Q_1, Q_2 or Q_3 is the MES for the above figure, and why?

4.2.2 EXTERNAL ECONOMIES AND DISECONOMIES OF SCALE

External Economies of Scale

As an industry grows in size, this can lead to external economies of scale for the firms in this industry. This is where a firm, regardless of its own individual size, benefits from the whole industry being large. It is independent of the individual firm's output. External economies of scale refers to the fall in average cost of production experienced by a firm when the industry grows as a result of other firms in the industry expanding their scale of production or more firms entering the industry.

DEFINITION:

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 scale of production.

External economies of scale are graphically represented by a downward shift of the LRAC curve from $LRAC_0$ to $LRAC_1$ as shown in Figure 7.

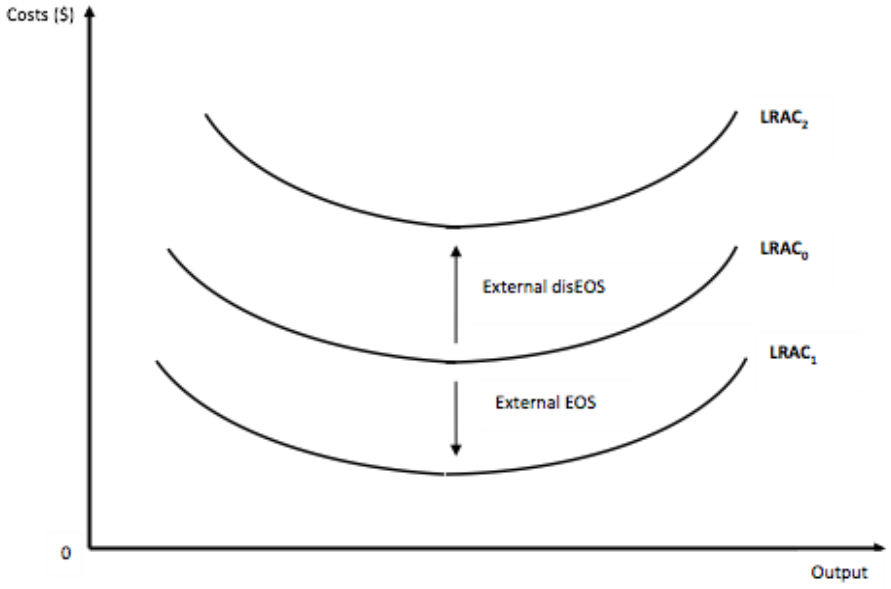


Figure 7: External Economies and Diseconomies of Scale

The following are some sources of external economies of scale:

a. Economies of Concentration

These occur when firms in a particular industry are concentrated together. Advantages might include a shared pool of skilled workers and better industry infrastructure.

✳ **Trained Workforce**

As the industry expands, the demand for labour with the necessary skills increases. Training schools may be set up to provide training and education, which means that a pool of skilled workers is readily available. Individual firms may not need to provide additional training as the manpower they hire would already have the prerequisite skills, resulting in lower average cost of production for each firm in the industry. Or each of the firm may benefit from having easier access to labour with specific skills and it is able to reduce the search cost (e.g. opportunity cost of time spent on searching, advertising cost, etc) for skilled workers. For example, SHATEC was set up by the Singapore Hotel Association to train workers for the hospitality industry. This removed the need for each hotel to incur duplicated costs to conduct their own training.

✳ **Better Industry Infrastructure**

Due to concentration of the industry in a particular region, facilities such as better transport, banking and telecommunication systems may be set up to serve the needs of the industry, thereby lowering operating costs.

It is also possible for the different firms in the same industry to share the usage of special equipment instead of each buying their own which will help them to reduce their average costs of production.

For example, the firms in the biomedical hub in Biopolis are able to share state-of-the-art facilities such as DNA-sequencing, confocal microscopy, flow cytometry, X-ray crystallography and nuclear magnetic resonance laboratories. Furthermore, education institutions in the vicinity such as National University of Singapore and Singapore Polytechnic provide specialised education and training opportunities to meet the needs of the biomedical industry.

b. Economies of Information

These economies take the form of common information services provided for the industry as a whole by trade associations or central research centres through journals and newsletters. Firms can obtain up-to-date information on production cheaply by sharing the cost of research instead of spending on costly independent research.

c. Economies of Disintegration

When an industry is heavily localised, it becomes possible for firms to split up the production process & specialise in a single process or the manufacture of a single component. As the components are mass produced to supply the whole industry, they can be produced and supplied at a much lower average cost.

If the industry has grown large enough, the industry may spawn off ancillary industries or auxiliary firms to support the main firms in the production process. The auxiliary firms will concentrate on the production of intermediate products which has been outsourced to them. This allows the main firms in the industry to specialise in the core production processes. The increase in specialisation will improve productive efficiency for both the auxiliary as well as the main firms, thus reducing average costs.

An example of such an industry is the car industry where a network of ancillary industries or auxiliary firms has emerged to support car manufacturers like Toyota, which has nearly eighty percent of its car components outsourced. A big chunk of each vehicle made by the

Japanese giant – from screws to accelerator pedals to engines – are produced by a vast network of suppliers.

External Diseconomies of Scale

DEFINITION:

External Diseconomies of Scale occur when a firm incurs higher average costs as a result of ***the whole industry growing in size,*** independent of the firm's scale of production.

External diseconomies of scale occurs when the size of the industry becomes so large that average cost of production increases. This is depicted in Figure 7 as an **upward shift** in the LRAC curve from $LRAC_0$ to $LRAC_2$. It could be due to the following factors:

a. Higher Factor Prices

As the industry expands, the demand for factors of production rises. This will result in higher input prices. This rise in prices can be significant, especially if supply of factors of production are price-inelastic. For example, it may be difficult for a firm to obtain skilled labour or specific raw materials due to other firms requiring them. The firm may thus have to bid up the prices of some of its inputs in order to secure them and this will lead to an increase in its average cost of production.

b. Increased Strain on Infrastructure

Increased congestion due to concentration and expansion of industry may result in rising transportation costs. The infrastructure will be stretched to its limit giving rise to overcrowding, pollution and other problems.

5. ALTERNATIVE OBJECTIVES OF FIRMS

Learning Objectives:

- ✳ Explain the different possible objectives a firm may adopt other than profit maximising.

The traditional objective of firms is assumed to be profit maximisation (MR=MC) in the short run as well as the long run. In reality, while firms ultimately **aim to maximise profits**, **especially in the long run,** they may **have many other objectives or goals to pursue in the short run,** some of which may conflict with the aim of maximising profits in the short run.

NOTE:

Diagrammatical analysis of these objectives are important and will be touched on in the next sub-chapter.

Other objectives of the firm may include:

a. Market Share Dominance

This objective aims to drive rivals out of the market thereby increasing the firm's market share, hence dominate the market. This strategy usually entails forgoing short run profits in pursuit of long run profits.

One way of doing this is through the strategy of predatory pricing, especially in situations where the firm has sufficient past profits to cope with the losses incurred in the process. This means the firm will sell its products below the average cost of their rivals (and may even be below its own profit maximising price or average cost) to force competitors out of the market. Over time, the rival firms might not be able to cope with the losses as a result of matching low prices set by the firm and choose to exit the market. This allows the remaining firms to increase their market share and corresponding level of market power. Any short run losses resulting from this objective can be compensated by higher prices in the long run when the competitors are out of the way.

b. Revenue Maximisation

This theory put forth by William Baumol in the 1950s states that the level of a firm's sales revenue is often taken to reflect the health of the firm, and hence the ability of the firm's sales managers and commission-based employees. For sales managers and commission-based employees whose income is dependent on the total revenue earned by firm, they might choose to maximise the firm's total revenue (TR) rather than maximise profits in the short-run. In order to maximise revenue, sales managers and commission-based employees make decisions to increase the level of production to the point where marginal revenue is zero (MR=0) and

setting a price lower than what is required for profit-maximising (MR=MC)

c. Growth Maximisation

Firms may aim to maximise growth instead of profits by internal expansion or merger. This aims to increase the size of the firm in the long-run. In the process, profits may not be maximised but the scale of production will increase. Reasons for growth include being recognised as a rapidly growing 'dynamic' organisation or gaining economies of scale. Promotion prospects are greater in an expanding organisation since new posts are created, large firms may pay higher salaries, and managers may obtain greater power in a large firm. (This can be illustrated by a movement downwards along the LRAC of the firm but the firm may no longer produce at MR=MC). Another possible reason for growth is diversification. A firm may decide to acquire firms in other product markets (e.g. conglomerate merger) in order to diversify its business and spread its risks over a range of markets. Such growth may involve a sacrifice of profits in the short term as the costs of acquiring other businesses may exceed the revenue gained in the short run.

d. Profit Satisficing Objective

The traditional view of firms assumes that they know exactly the costs of production and revenue. The cost of obtaining sufficient information to make profit-maximising decisions, however, could be significantly high in some cases such as in firms that have several production locations (possibly international) and multiple product offerings. Firms may also be contented with a profit satisficing level of output in order to avoid undue stress or perceived challenges from expansion. In addition the shareholders of the firm (who are primarily interested in maximising profits) could be far 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 shareholders even though it falls below the profit-maximising level, and enjoy other benefits such as shorter operating hours and lower levels of stress.

One form of profit satisficing behaviour is when the firm adopts cost-plus pricing. Firms may be unable to maximise profits, as they are unable to accurately estimate MR and MC due to unavailable or imperfect information. In such situations, firms may resort to simpler pricing models such as cost-plus pricing, where firms put a mark-up on top of the average cost of their goods and adjust the mark-up based on how well the goods are selling.

Cost-plus pricing (full-cost pricing) (another name for mark-up pricing) refers to a situation when firms price their product by adding a certain profit 'mark-up' to average cost. For example, if the per unit or average cost of the good is $100, and the firm decides on a mark-up of 80%, it will sell the good at a price of $180.

Many firms set prices by adding a profit mark-up to average cost. This cost-plus pricing is most likely when firms are profit satisficers or when they do not have the information to find the price that will equate marginal cost and marginal revenue. The mark-up could be based on achieving a target level of either total profit or profit per unit, which the firm considers to be reasonable or satisfactory. In either case, a supply curve can be derived by adding the corresponding mark-up to the average cost curve. For firms keen to increase profit, the size of the mark-up can be varied as market conditions permit the target profit to be increased.

e. **Managerial Utility Maximisation**

Developed by O.E. Williamson in the 1960s, this theory proposed that once the satisfactory levels of profit are reached, managers running the firm have the discretion to choose what policies to pursue. In other words, they are free to pursue their own interests and may seek to maximise their own utility. For example, they may hire more staff than necessary or spend excessive amounts on office furnishings to enhance their sense of power or prestige. The four main factors affecting a manager's utility are salary, job security, dominance (including status, power and prestige) and professional excellence. When managers are allowed to pursue their own interests, the average costs of the firm are likely to be higher, and hence profits may not be maximised. This usually occurs in larger firms where the control (or management) of the firm is separated from the ownership of the firm, thus giving rise to the principal-agent problem

where the agents (e.g. salaried managers) who run or manage the firm may not have the same objectives as the owners (principal) of the firm.

6. ANNEX

Law of Diminishing Returns

In the short run, the level of output can only vary by changing the quantities of variable factors, with the amount of fixed factors remaining constant.

The law of diminishing returns is a short run concept as it assumes the existence of at least one fixed factor. It describes the relationship between increases in the amount of variable factors and the resulting changes in output in the short run.

DEFINITION:

The **Law of Diminishing 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.

To understand the law of diminishing returns better, let us use the same example of the bakery as illustrated in Table 2. In this case, we will assume the fixed factor to be the factory (the size of the factory is fixed) while the variable factor to be labour. 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 or identical). Technology is also assumed to be constant in the short run.

No. of Factories	No. of Bakers	Quantity of Bread Baked	Addition to Output	Capital/Labour
1	0	0	0	-
1	1	30	30	1
1	2	70	40	1/2
1	3	120	50	1/3
1	4	180	60	1/4
1	5	220	40	1/5
1	6	240	20	1/6
1	7	240	0	1/7
1	8	230	-10	1/8

Table 5: Contribution to Output by Each Additional Baker (Illustrating the Law of Diminishing Returns)

With reference to Table 5, initially, with only one baker working, factory space (i.e. the fixed factor) is not fully utilised yet (**under-utilised**). From the first to the fourth baker, each additional baker contributes more and more to the total output as there is greater efficiency in the use of the factory space for productive purposes.

However, as more and more bakers are added, the factory became congested and mobility became restricted. It will reach a point when any additional baker will contribute a smaller amount of output due to the discomfort working within a constrained space or having to wait their turn to use the oven- this is when the bakery experiences the **onset of diminishing returns.**

From Table 5, we observe that 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!

Therefore, from Table 5, 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.

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. Thus, the Law of Diminishing Returns is also known as the Law of Variable (Factor) Proportions.

THINK ABOUT IT:

Imagine that you were in a group completing a craft project to cut strips of papers. Your group consisted of 2 members and was given 4 pairs of scissors.

Explain how the law of diminishing returns set in as your group added more members.