BLOCK 3 PRODUCTION AND COSTS

BLOCK 3 PRODUCTION AND COSTS

Block 3 develops the theory of the firm and explains the laws that are observed in course of production. This will enable you to know how firms combined inputs such as capital, labour and raw materials to produce goods and services in a way that minimises costs of production. In this process various concepts like production function, Iso product curves, Iso-cost lines etc have been explained.

The block comprises four units. **Unit 7** throws light on production function with one variable input, and discusses the law of variable proportions. **Unit 8** deals with the Properties of isoquants and optimal combination of factors and producer's equilibrium. The economic region of production and ridge lines and the expansion path have also been discussed. **Unit 9** covers the production function in the event all the inputs vary and hence application of returns to scale. **Unit 10** discusses the cost side of production considering different types of costs.

UNIT 7 PRODUCTION WITH ONE VARIABLE INPUT

Structure

7.0 OBJECTIVES

After going through this unit, you will be able to :

- state the concept of total product, average product and marginal product;
- explain the nature and relationship of total, average and marginal product curves;
- analyse the operation of the law of variable proportions; and
- identify the three stages of production.

7.1 INTRODUCTION

For the purpose of production, we require a combination of various inputs or factors of productions. It is only with the joint efforts of these inputs (like labour, machines, land, raw materials etc.) that output is produced. Normally, production is carried out under conditions of variable proportions which implies that the rate of input quantities may vary. Fixed proportions production means that there is only one ratio of inputs that can be used to produce a good. For example, only one driver can work one truck. In this case, the ratio of driver and truck is technologically determined and is fixed. It is beyond the capabilities of the producer to change it. However, the ratio of land and labour in agriculture can be changed and is thus regarded as variable. In the short run, not all inputs are variable. In the long run, however, all inputs are variable and the ratio of inputs may also vary. This is the case of technological Progress. In this unit, we shall focus only on short run production. In the short run, for the

*Dr. V.K. Puri, Associate Professor of Economics, Shyam Lal College (University of Delhi) Delhi.

purpose of analysis, it is often assumed that only one input is variable and all other inputs are fixed. We shall follow this convention.

7.2 TOTAL, AVERAGE AND MARGINAL PRODUCTS

At the outset we shall explain the concept of total, average and marginal products. The short run production function, whether it is shown as a table, a graph or as a mathematical equation, gives the total output obtainable from different quantities of the variable inputs given a specified amount of the fixed input. Let us now consider the case in which capital is fixed, but labour is variable, so that the firm can produce more output by increasing the labour input. For example, consider a firm manufacturing garments. It has a fixed amount of equipment, but it can hire more or less labour to operate the machines. For decision making, the firm's manager (or owner) must know how the amount of total output or product (Q) increases (if at all) as the labour input (L) increases. Table 7.1 provides this information about the production function.

Table 7.1 shows the output that can be produced with different amounts of labour and with capital fixed at 5 units. The first column shows the fixed amount of capital, the second shows the amounts of labour from zero to 10 units and the third shows total product or output. From the table, it is clear that when labour input is zero, output is zero because capital alone cannot produce anything. Then, upto a labour input of seven units output increases first at an increasing rate and then at a decreasing rate in response to increased use of labour. The eighth unit of labour input does not raise output. Whether firm applies 7 or 8 units of labour input to a fixed amount of capital input, total output remains 224 units. Beyond this point using more units of labour input is counter productive because output declines as use of labour is increased.

Table 7.1: Production with One Variable Input

Amount of Capital (K)	Amount of Labour (L)	Total Product or Output (Q)	Average Product (Q/L)	Marginal Product (ΔQ/ΔL)
5	0	0	--	--
5	1	20	20	20
5	2	60	30	40
5	3	120	40	60
5	4	160	40	40
5	5	190	38	30
5	6	216	36	26
5	7	224	32	8
5	8	224	28	0
5	9	216	24	-8
5	10	200	20	-16

Although the figures provided in Table 7.1 are hypothetical, the general relationship they indicate is common. To examine the relationship further, we introduce the concepts of average product and marginal product of an input.

The average product (or average physical product) of an input can be defined as total output (or total product) divided by the amount of input used to produce that output. For example, 4 units of labour input produce 160 units of output, so the average product of labour is 40 units of output per worker at that level of employment. In a more general way, we may express

$$AP_L = \frac{Q}{L}$$

where, AP_L = average product of labour

Q = total output or total product

L = amount of labour

The fourth column in Table 7.1 shows the average product of labour (AP_L). The average product for each quantity of labour is derived by dividing total output shown in column 3 by corresponding amount of labour in column 2 that produces each output level. In our illustration, the average product of labour increases initially but when labour input exceeds 4 units, it tends to fall.

The marginal product (or marginal physical product) of an input is defined as the change in total output due to a unit change in the use of an input while quantities of other inputs are held constant. For example, with capital fixed at 5 units when the amount of labour increase from 3 to 4 units, total output rises from 120 to 160 units or by 40 units. So the marginal product of labour, when fourth unit of labour input is employed, is 40 units of output. We may thus generalise,

$$MP_L = \frac{\Delta Q}{\Delta L}$$

where, MP_L = Marginal product of labour

ΔQ = Change in output

ΔL = Change in labour input

In Table 7.1, the fifth column shows the marginal product of labour. It may be noted that like the average product, the marginal product increases initially and then falls and finally becomes negative. In the present example, the marginal product of labour becomes negative when labour input exceeds 8 units. This happens when the variable input is used too intensively with the fixed input.

The marginal product is greater than average product when average product is rising, equals average product when average product is at maximum, and is less than average product when average product is falling.

This proposition is, in fact, true of all marginal and average relationships.

7.3 TOTAL, AVERAGE AND MARGINAL PRODUCT CURVES

Fig. 7.1 plots the information provided in Table 7.1 (it has been assumed in drawing the graphs that both labour input and the product are divisible into smaller units and thus the relationships are smooth curves rather than discrete points). The total product curve shown in Fig. 7.1 indicates how the total product varies with the quantity of labour input used. As indicated in Table 7.1, Fig. 7.1 a also shows that first the total output increases at an increasing rate upto point E as more labour is used. The point E where total product stops increasing at an increasing rate and begins increasing at a decreasing rate is called the **point of inflexion**. Total product reaches a maximum at 224 units when 7 units of labour input are used. The use of an additional unit of labour input at this stage does not lead to any increase in total product. Beyond this point, further use of labour input results in a fall in total product.

That portion of total product curve (TP) is shown by dashed segment which indicates a decline in output as a result of increased employment of labour. In Fig. 7.1 a when labour input is expanded beyond eighth unit, output falls which means that production is not technically efficient and is thus not a part of the production function.

Fig. 7.1 b shows the average and marginal product curves for labour. (The units of the vertical axis have been changed from output per period of time to output per unit of labour). Hence, average product and marginal product curves measure the output per unit of labour. It may be noted that as the use of labour input increases, initially the marginal product of labour increases, reaches a maximum at 3 units of labour, and then declines. The marginal product of labour in our example becomes zero at 8 units of labour and thereafter turns negative. However, technical efficiency rules out the possibility of negative marginal products and is, therefore, not a part of the production function. The average product of labour also increases initially, reaches a maximum at 4 units of labour input, and then declines.

Relationship between MP and AP Curves:

Let us now consider the relationship between the marginal and average product curves. As is true of all marginal and average curves, there are definite relationships between the marginal and average product curves.

i) When marginal product increases, average product also increases though at a rate lower than that of the marginal product. It is important to note in this context that even when marginal product starts declining but remains greater than the average product, the latter shows a tendency to increase.

ii) When the average product is maximum, the marginal product is equal to it. This is the reason why the marginal product curve intersects the average product curve at its highest point.

iii) Beyond this point, when the marginal product declines, it also pulls down the average product. However, the rate of decline in the average product is less than that of the marginal product.

Relationship between TP and MP Curves

The relationship between the total product curve and the marginal product curve can be stated as under:

i) As long as marginal product is positive, total product curve will continue to rise.

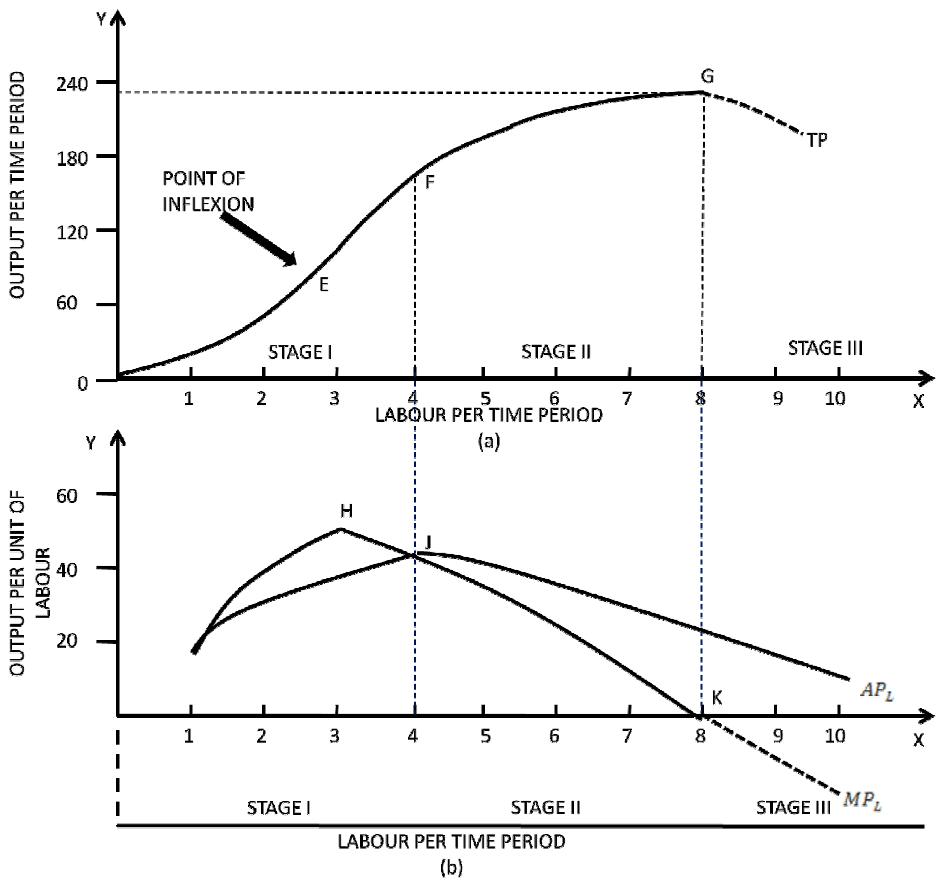

Fig 7.1: Production with one variable input (labour). In the upper part of the figure, the total product curve (TP) of labour is shown. The lower part of the figure shows how average product curve (AP) of labour and marginal product curve (MP) of labour are obtained with the help of information contained in the upper part

ii) When marginal product is zero, total product curve reaches its highest point. It may be noted that when eighth unit of labour input is employed, marginal product of labour becomes zero and total product is at the maximum.

iii) Thereafter, marginal product of labour is negative and total product curve has a downward slope which means that total product falls.

Check Your Progress 1

1) Indicate the following statement as true (T) or false (F):

   i) The marginal product is greater than average product when average product is falling.

   ii) As long as marginal product is rising, total product curve will continue to rise.

2) Discuss the relationship between the marginal and average product curves.

....

....

....

7.4 THE LAW OF VARIABLE PROPORTIONS: RETURNS TO A FACTOR

Knowledge regarding the conditions of production reveals that as more and more of some input is employed, all other input quantities being held constant, normally marginal and average product (of the variable input) increase upto a point. Thereafter, marginal product starts declining and this pulls down the average product also. In the production process generally land, capital equipment and buildings remain fixed in the short run while quantities of labour and raw materials can be conveniently varied. However, we may consider a case where amount of capital is fixed and the quantity of labour is increased.

i) In this case, initially the marginal product of labour will increase as its amount is increased and the marginal product will also pull up average product with it. In this situation, total product increases at an increasing rate.

ii) If the variable input, say, labour is further increased, marginal product stops increasing after a point. Therefore, the rate of increase of total product also shows a tendency to fall.

iii) Ultimately marginal product turns negative and this causes a fall in total product itself.

Since in the short run, changes in technology are ruled out, the tendency of marginal product to decline after a point is inevitable. This statement of trends in marginal product in response to changes in the quantities of a variable factor applied to a given quantity of a fixed factor is called the **law of diminishing returns**. It is also called the law of variable proportions because it predicts the consequences of varying the proportions in which factors of production are used. we can sum up the law of variable proportions as follows:

> **"As equal increments of one input are added, the inputs of other productive services being held constant, beyond a certain point the resulting increments of product will decrease, i.e, the marginal product will diminish."**

The law of variable proportions can be easily followed with the help of Table 7.1 and Fig. 7.1 which has been drawn on the basis of illustration given in Table 7.1. In Table 7.1, it has been assumed that capital is a fixed factor and its quantity remains unchanged at 5 units. Labour is the variable factor and its quantity increases from 1 to 10. It can be seen from Table 7.1.

i) As the amount of labour employed increases, the total output also increases until the seventh unit of labour is employed. Initially the increase in output takes place at an increasing rate because marginal product rises. This tendency is observed upto the point E where marginal product reaches a maximum. At point E, which is the point of inflexion, the rate of increase in total product switches from increasing to decreasing because marginal product begins to diminish. However, average product continues to increase until it reaches a maximum at point F on total product curve (point J on average product curve).

ii) When the amount of labour is further expanded, total product continues to increase though at a diminishing rate. Both marginal product and average product remain positive, but both continue to diminish. Eventually, total product reaches a maximum at point G and the marginal product becomes zero (note point K in Fig. 7.1 b). The average product, however, remains positive but continues to diminish.

iii) Any attempt to increase output beyond this point by employing more units of labour will not be fruitful. In fact, it will be counter-productive because marginal product is negative which implies that total product diminishes.

Product curves such as the one shown in Fig. 7.1 are general representations of production function with fixed and variable inputs. To illustrate particular instances, similar product curves could be drawn, though each different from others in some way. The stage of increasing marginal product may be long or brief or can be totally absent. Moreover, when marginal product diminishes, the rate at which it happens may be different in each case. Table 7.2 sums up the law of variable proportions.

Table 7.2: Properties of Product Curves

Total Product	Marginal Product	Average Product	Figure 7.1
Stage I first increases at increasing rate	Increases	Increases	to point E
then rate of increase changes from increasing to diminishing	reaches a maximum, and then starts diminishing	continues increasing	at points E and H
Stage II continues to increase at diminishing rate	continues diminishing	reaches a maximum where it equals MP and then starts diminishing	at points F and J
reaches a maximum and then starts diminishing	becomes zero	continues diminishing	at points G and K
Stage III diminishes	is negative	continues diminishing	to right of points J and K

7.4.1 The Three Stages of Production

Normally when the amount of a variable input is expanded, the marginal product first rises and then falls and the product curves have the shapes shown in Fig. 7.1. Conventionally, these product curves are partitioned into three regions, shown as Stages I, II and III in Fig. 7.1.

Stage I is characterised particularly by the rising average product. In our example, Stage I occurs when labour is employed from 1 to 4 units. In Stage 1, total product first increases at an increasing rate and thus marginal product rises. It reaches a maximum at labour input of 3 units. When fourth unit of labour input is employed, diminishing returns set in implying that total product increases at a diminishing rate and the marginal product falls.

In **Stage II**, total product increases at a diminishing rate and thus both marginal product and average product decline. Marginal product being below the average product, pulls the latter down. The right-hand boundary of Stage II is at maximum total product where marginal product reaches zero. In our example, Stage II ranges from 4 to 8 units of labour.

In **Stage III**, total product falls and marginal product is negative. In our example, stage III occurs when labour is employed in excess of 8 units.

Actual Stage of Operation

The rational producer will operate in Stage II. It is not difficult to follow why production will not be done in Stage III. In Stage III, less output is produced by using more of the variable input which means that production costs would be higher in Stage III than they were in Stage II. Obviously, any rational producer will always avoid such inefficiencies in the use of production inputs.

In Stage I, average product of the variable input is increasing. Therefore, if the amount of variable input is doubled, the output more than doubles and the unit cost of producing output decreases. If a firm is operating in a competitive market, it would avoid producing in this stage because by expanding output it reduces the unit costs while the price it receives remains same for each additional unit sold. This means that total profits increase if production is expanded beyond the region of rising average product.

To sum up we can say: Initially, the variable factor-labour is not able to use all the capacities of the fixed factor, hence MP and AP remain low. For instance, one worker may not be able to make full use of the potential of a one hectare plot of land. But two workers, together are is a better position to work on that field. Hence rise in MP as Labour increases from 1 to 2.

Thus, any rational producer will operate in the second stage only when the law of diminishing marginal return operates. This is why the law of variable proportions is also called the Law of Diminishing Marginal Returns to a factor.

7.4.2 Explanation of Increasing Returns

According to modern economists, when in the initial stage of production quantity of the variable factor is increased, the tendency of increasing returns in production operates. The classical economists had also observed this tendency and had termed it as the Law of Increasing Returns. However, they felt that this law operated only in manufacturing industries. As against this, the modern economists believe that this law can operate in any area of economic activity. Below we give the views of Marshall (representing the former position) and Joan Robinson (representing the latter position) in this regard.

Marshall opined that the tendency of increasing returns operates only in the manufacturing industries. He believed that when the quantity of labour and capital employed in the manufacturing industries is increased, the scale of production expands and this leads to a better organisation of production. In Marshall's own words:

"An increase in labour and capital leads generally to improved organisation, which increases the efficiency of the work of labour and capital... Therefore, in those industries which are not engaged in raising raw produce, an increase in labour and capital generally gives a return increased more than in proportion."

Joan Robinson's explanation of the tendency of increasing returns is more scientific. She states:

"When an increased amount of any factor of production is devoted to a certain use, it is often the case that improvements in organisation can be introduced which will make natural units of the factor (men, acres or money capital) more efficient, so that an increase in output does not require a proportionate increase in the physical amount of the factors."

1) The tendency of increasing returns operates not only in manufacturing industries but in all productive activities. Limiting the application of this tendency to manufacturing industries alone is wrong.

2) The tendency of increasing returns comes into operation because the efficiency of the factors of production is improved.

Let us now examine in detail why the tendency of increasing returns operates.

1) **Optimum combination of factors of production:** According to Joan Robinson, full exploitation of some indivisible factors of production is not possible until increased quantities of some other factors of production are employed. Therefore, when the producer engages a small quantity of different factors of production, an optimum proportion among them is not established and the level of production remains low. When he increases the quantities of those factors of production, which were employed less (in relation to the requirements of optimum production), marginal product increases till the point is reached where the factors are combined in optimum proportion. Naturally, at this point, output level is the maximum.

2) **Large size of fixed factors:** When the size of the fixed factors used for producing a given good is very large while the quantity of the variable factor used is very small, the level of efficiency remains very low. As more and more quantities of the variable factors are employed, marginal productivity increases (since the level of efficiency increases). For example, if only one person is working on a ten hectare plot of land, his productivity will be very low. As the number of workers increases, division of labour and specialisation will lead to increasing returns as marginal product will rise rapidly.

7.4.3 Explanation of Constant Returns

If even on continuously increasing the quantity of variable factors of production in a firm, the marginal product neither increases nor decreases but remains constant, the tendency of constant returns is in operation. In fact, there is no industry in which increase in the quantity of variable factors of production yields constant returns permanently. According to Marshall, "if the actions of the law of increasing and diminishing returns are balanced, we have the law of constant returns."

Marshall feels that the operation of the law of constant returns is very limited. According to him, this law can operate only when there is a balance between the tendencies of increasing returns and diminishing returns. However, modern economists regard the area of operation of constant returns as fairly large. According to them, tendency of constant returns is generally found to operate before the tendency of diminishing returns sets in. In no field of productive activity increasing returns are obtained forever. Whether it is agriculture, manufacturing, industry or any other productive activity, the tendency of increasing returns can operate only up to a certain limit. After this limit is reached, constant returns operate for some time. From the point of view of the producer, this is an important stage because it exhibits an optimum combination of the factors of production. In this stage, marginal cost is the minimum. This is due to two reasons. First, the stage of constant returns is reached only when the tendency of increasing returns comes to an end so that there is no possibility of a further decline in marginal cost. Second, after the stage of constant returns, the stage of diminishing returns sets in. Therefore, the stage of constant returns is very significant from the point of view of the producers.

7.4.4 Explanation of Diminishing Returns

The diminishing returns stage is the most important of the three stages of the law of variable proportions. In Economics, the explanation of the law of diminishing returns is presented in two ways. The classical economists believed that this law applies only to agriculture. Basically accepting this position of the classical economists, the neo-classical economist Marshall had stated, "We say broadly that while the part which nature plays in production shows a tendency of diminishing returns, that part which man plays shows a tendency of increasing returns."

Modern economists like Joan Robinson, Stigler, etc. constitute the second category of economists. These economists regard the law of diminishing returns of far greater applicability than the classical economists. According to them, this law operates in all areas of productive activity.

Marshall had argued that this law operated only in agriculture. Therefore, he discussed it only in reference to agriculture. According to him,

> **"An increase in the capital and labour applied in the cultivation of land causes in general a less than proportionate increase in the amount of produce raised unless it happens to coincide with an improvement in the arts of agriculture."**

The implication is that when land is kept fixed in agriculture while the quantity of labour and capital applied on that land is increased, total production increases but not in the same proportion as the factors of production are increased. It increases by a lesser proportion. For example, if an agriculturist doubles the amount of labour and capital employed on a fixed plot of land, the total production will undoubtedly increase but it will not double itself. Due to this reason agriculturists do not consider it profitable to continuously increase the application of other factors of production on their fixed plots of land. They know from their experience that unless there is some improvement in agricultural techniques, increased application of labour and capital on a fixed quantity of land leads to a situation of continuously declining marginal product.

Marshall has accepted two limitations of the law of diminishing returns as applied to agriculture:

1) **The law generally operates in agriculture:** Marshall was aware of the fact that the law of diminishing returns does not always operate in agriculture (hence the qualification that it generally operates in agriculture). In some cases when the agriculturist applies the first unit of labour and capital on his fixed plot of land, the fertility of the soil is not properly exploited. Accordingly, the level of production remains low. When the second unit of labour and capital is applied, output increases in a greater proportion. However, this tendency does not remain for long because the agriculturist soon finds that additional units of labour and capital start yielding a lower and lower marginal product. On account of the above reasons, Marshall was careful in pointing out that the law of diminishing returns operates generally in agriculture. However, in certain exceptional cases, it may not operate.

2) **There should be no improvement in agricultural techniques:** The law of diminishing returns operates only if there is no improvement in agricultural techniques. It is a law of static agriculture. If the agriculturist is able to expand irrigation facilities on his land, or make use of better seeds, better agricultural implements, more fertilisers, etc. or use new scientific methods in production, he can stall the operation of this law. Generally, an improvement in agricultural techniques leads to a more than proportionate increase in output corresponding to an increase in labour and capital.

As against the view of Marshall, modern economists like Joan Robinson, Stigler and Boulding regard the law of diminishing returns as more pervasive and universal. According to these economists, this law operates in all branches of productive activity. Accordingly, they have presented this law in a general fashion as would be clear from the definition of this law presented by Joan Robinson:

> **"The Law of Diminishing Returns, as it is usually formulated, states that, with fixed amount of any one factor of production, successive increases in the amount of other factors will after a point yield a diminishing increment of the product."**

From the above definition of the law by Joan Robinson, it is clear that she regards this law as of universal value and does not restrict its application to agriculture alone. According to her, this law operates in all branches of productive activity and the principal reason behind the operation of this law is that the optimum proportion between different factors of production breaks down sooner or later.

The law of diminishing returns is a logical necessity. When in any productive activity, the quantity of the variable factors of production employed with given quantity of fixed factor of production is increased, the law of diminishing returns sets in after the point of optimum proportion has been reached. Initially, application of variable factors was sub-optional, given the size of fixed factor. Later, the expansion in use of variable factors leads to sub-optimality of a different kind: each doze or unit of variable factors have sub-optional quantity of fixed factor to work on.

Another important reason for the operation of the law of diminishing returns is that one factor of production (out of the various factors of production) is used in a fixed quantity. Had all the factors of production been available in abundance and had it been possible to increase their use in production to all conceivable limits, the law of diminishing returns would not operate. However, all factors of production land, labour, capital, enterprise, organisation, etc. are scarce and often the supply of one of these is taken to be fixed. It is this factor that results in diminishing returns.

Check Your Progress 2

1) Indicate the following statement as true (T) or false (F):

    i) In statge II of production, both marginal product and average product decline.

    ii) In stage III of production, marginal product is negative.

    iii) The law of diminishing returns operates only in agriculture.

2) State the law of diminishing marginal returns. There is a provision to the law that other things be held constant. What are these things?

    ....

    ....

    ....

3) Explain the three stages of production. Why should a rational producer under competitive conditions produce in stage II?

    ....

    ....

    ....

4) Explain the (i) law of increasing returns, (ii) law of constant returns.

    ....

    ....

    ....

7.5 LET US SUM UP

In this unit we have focused on short run production assuming that only one input is variable and all other inputs are fixed. We then define total product, average product of an input and the marginal product of an input. We note that total product in the case of production with one variable input first increases at an increasing rate as the amount of variable input expands and then switches to increasing with decreasing rate. Having reached a maximum, it eventually declines. We then explain the law of variable proportions. Conventionally the product curves drawn to depict the law of variable proportion are partitioned into three stages. In stage I, average product increase throughout, in stage II marginal product from the point where it equals average product falls throughout but remains positive; and in stage III total product fall and marginal product is negative. The diminishing returns stage is the most important of the three stages of the law of variable proportions.

7.6 REFERENCES

1) Robert S.P rindyck, Daniel L. Rubinfeld and Prem L. Mehta, *Microeconomics* (Pearson Education, Seventh edition, 2009), Chapter 5, Section 5.1.

2) Dominick Salvatore, *Principles of Microeconomics* (Oxford University Press, Fifth edition, 2010), Chapter 7, Section 7.2.

3) A.Kontsoyianmis, *Modern Microeconomics* (The Macmillan Press Ltd., Second Edition, 1982/, Chapter 3.

4) John P Gould and Edward P Lazar, *Microeconomic Theory* (All India Traveller Bookseller, Sixth edition, 1996), Chapter 6.

7.7 ANSWERS OR HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) (i) (F); (ii) (T)

2) See Section 7.3

Check Your Progress 2

1) (i) F; (ii) (T); (iii) (F)

2) See Sub-section 7.4.4

3) See Sub-section 7.4.1

4) See Sub-section 7.4.2. for law of increasing returns and Sub-section 7.4.3 for law of constant returns.

UNIT 8 PRODUCTION WITH TWO VARIABLE INPUTS

Structure

8.0 OBJECTIVES

After going through this unit, you should be able to:

- know the meaning and nature of isoquants;
- identify the economic region in which production is bound to take place;
- find out the level at which output will be maximised subject to a given cost;
- find the point on the isoquant where cost will be minimised; and
- describe the nature of optimal expansion path both in long run and short run.

*Dr. V.K. Puri, Associate Professor of Economics, Shyam Lal College (University of Delhi) Delhi.

8.1 INTRODUCTION

As stated in Unit 7, production requires the use of certain resources often called factors of productions or inputs. When resources are broadly defined, they are known as factors of production and classified as labour, land and capital. The relationship between the inputs to the production process and the resulting output is described by a production function. In this unit we shall start with a definition of production function and then proceed to discuss the concept of isoquants. This will be followed by a discussion on the economic region of production and then, finally, on how the optimum combination of factors and producer's equilibrium is obtained.

8.2 PRODUCTION FUNCTION

The theory of production begins with some prior knowledge of the technical and/or engineering information. For instance, if a firm has a given quantity of labour, land and machinery, the level of production will be determined by the technical and engineering conditions and cannot be predicted by the economist. The level of production depends on technical conditions. If there is an improvement in the technique of production, increased output can be obtained even with the same (fixed) quantity of factors. However, at a given point of time, there is only one maximum level of output that can be obtained with a given combination of factors of production. This technical law which expresses the relationship between factor inputs is termed as production function.

The production function describes the laws of production, that is, the transformation of factor inputs into products (outputs) at any particular period of time. Further, the production function includes only the technically efficient methods of production. This is because no rational entrepreneur will use inefficient methods.

For the sake of simplicity, we may assume that there are two inputs, labour (L) and capital (K). We can, then, write the production function as

$$Q = F(L,K)$$

This equation relates the quantity of output Q to the quantities of the two inputs, labour and capital.

A popular production function in economics is Cobb Douglas production function which is given as

$$Q = AL^{\alpha}K^{\beta}$$

A special class of production functions is linear homogenous production function. In this case, *when all inputs are expanded in the same proportion, output expands in that proportion.* In this case, Cobb Douglas production function becomes

$$Q = AL^{\alpha}K^{1-\alpha}$$

i.e. $\beta = 1-\alpha$

Here we can see that when labour and capital are increased λ times, output Q also increased λ times as

$$A(\lambda L)^{\alpha}(\lambda K)^{1-\alpha} = A[\lambda^{\alpha+(1-\alpha)} L^{\alpha}K^{1-\alpha}] = \lambda[AL^{\alpha}K^{1-\alpha}] = \lambda Q$$

8.3 WHAT ARE ISOQUANTS?

8.3.1 Definition

An isoquant is the locus of all the combinations of two factors of production that yield the same level of output.

It is easy to understand the concept of an isoquant with the help of an example. Let us suppose that a firm wants to produce 100 units of commodity X and for that purpose can use any one of the six processes indicated in Table 8.1.

Table 8.1: Isoquant table showing combinations of Labour and Capital Producing 100 Units of X

Process	Units of Labour	Units of Capital
1	2	7
2	4	4
3	6	3

From Table 8.1, it is clear that all the three processes yield the same level of output, that is, 100 units of X. The first process is clearly capital-intensive. Since we assume possibilities of factor substitution, we find that there are two more processes available to the firm and in each of them factor intensities differ. The third process is the most labour-intensive or the least capital-intensive. Graphically, we can construct an isoquant conveniently for two factors of production, say labour and capital. One such isoquant is shown in Fig. 8.1. It has been constructed on the basis of information provided in Table 8.1.

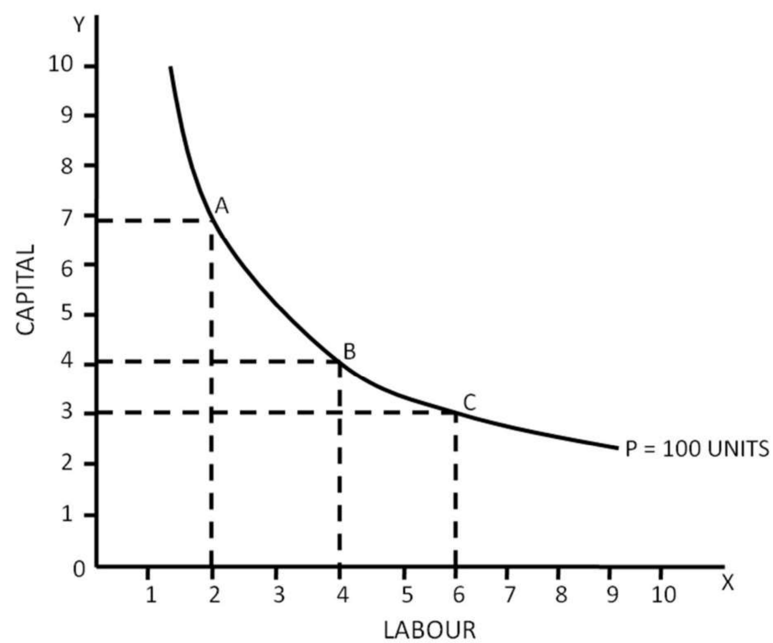

Fig. 8.1: This figure shows that at point A, B and C same level of output (=100 units) is obtained by using different combinations of labour and capital. Curve p is known as isoquant

8.3.2 Types of Isoquants

1) Convex isoquant

2) Linear isoquant

3) Input-output isoquant

The traditional economic theory has mostly used the convex isoquant as shown in Fig. 8.1. However, the isoquant can assume some other shapes depending on the degree of the substitutability of factors. The two other possible production isoquants are linear isoquant and input-output isoquant.

Linear Isoquant: In case of perfect substitutability of the factors of production, the isoquant will assume the shape of a straight line sloping downwards from left to right as in Fig. 8.2. In Fig. 8.2 it is shown that when quantity of labour is increased by RS, the quantity of capital can be reduced by JK to produce a constant output level, i.e., 50 units of X. Likewise, on increasing the quantity of labour by ST, it is possible to reduce the quantity of capital by KL, and on increasing the quantity of labour by TU, quantity of capital can be reduced by LM for producing 50 units of X. Since in respect of labour RS = ST = TU and in respect of capital JK = KL = LM, it is clear that a constant quantity of labour substitutes a constant quantity of capital. It implies that a given commodity can be produced by using only labour or only capital or by infinite combinations of labour and capital. In the real world of production, this seldom happens. Therefore, a linear downward sloping isoquant can be taken only as an exception.

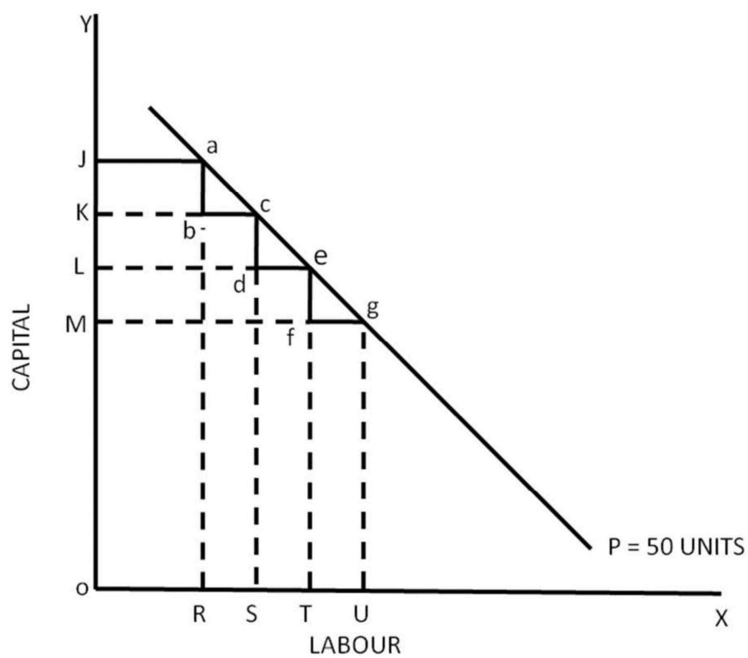

Fig. 8.2: In the case of perfect substitutability of factors of production, the isoquant becomes a straight line and is, therefore, known as linear isoquant

Input-Output Isoquant: When factors of production are not substitutes but complementary, technical coefficients are fixed. The meaning of this statement is that optimum output is obtained only when the factors of production are used in a fixed proportion. In this situation, if a producer uses one factor of production in excess of what is required by fixed proportion, there will be no increase in output. In the case of complementarily of factors of production, the shape of the isoquant is right angled or like the letter 'L' as shown in Fig. 8 .3. As would be clear from the figure, the isoquant is formed by two straight lines, one vertical and the other horizontal, and these two lines are perpendicular to each other. The common point of these lines is convex to the origin.

This type of isoquant is also called Leontief isoquant after Wassily Leontief who did pioneer work in the field of input-output analysis. Input-output isoquant does not imply that by increasing the quantities of the two factors of production, viz., labour and capital the output will increase proportionately; it implies only that for producing any quantity of a commodity, capital and labour must be used in a fixed proportion. Fig. 8.3, slope of Isoquant P_1 and P_2 indicates the capital-labour ratio has to be maintained for ensuring efficiency in production.

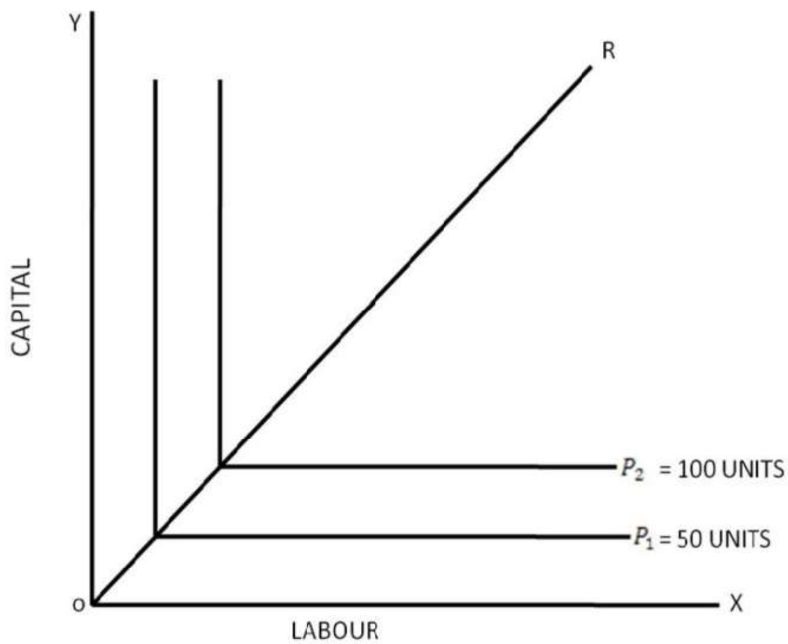

Fig. 8.3: If factors of production can be used only in a fixed proportion, the isoquant is 'L' shaped and is known as an input-output isoquant

8.3.3 Isoquants Map

The production function shows how output varies as the factor inputs change. Therefore, there are always a number of isoquants for a producer depicting levels of production (one isoquant depicting one particular level of production). Isoquants nearer the point of origin represent relatively lower level of production. The level of production increases as one moves away from the origin and goes to higher isoquants. A complete set of isoquants for the producer is called an isoquant map. One such isoquant map showing four isoquants is shown in Fig. 8.4.

In Fig. 8.4, P_4 is the highest isoquant and it represents the highest level of output, i.e., 400 units. P_3, P_2 and P_1 represent lower output levels in that order.

It may, however, be noted that the distance between two isoquants on an isoquant map does not measure the absolute difference between output levels.

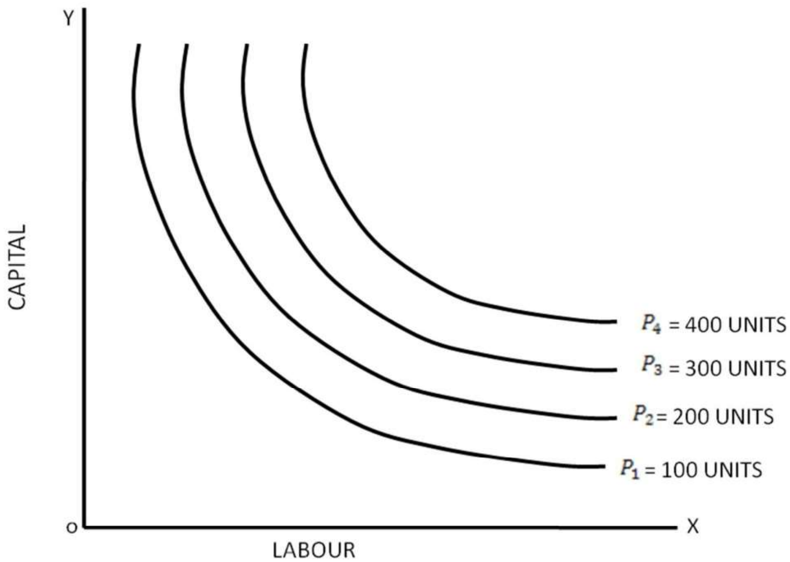

Fig. 8.4: When a number of isoquants are depicted together, we get an isoquant map

8.3.4 Assumptions of lsoquants

Isoquant analysis is normally based on the following assumptions:

1) It is generally assumed that there are only two factors or inputs of production. This makes the geometric exhibition of the concept easy since we can easily draw a diagram. If we abandon this assumption and consider four or five factors of production (in keeping with the reality) we would not be able to make use of the diagrammatic representation and would have to resort to the algebraic method.

2) The second assumption of the isoquant analysis is that the factors of production are divisible into small units and can be used in various proportions.

3) Technical conditions of production are given and it is not possible to change them at any point of time.

4) Given the technical conditions of production, different factors of production are used in the most efficient way. If this assumption is abandoned, then any one combination of the factors of production will yield a number of different levels of production of which the highest level obtained would be efficient (and all lower levels of production inefficient).

8.4 CHARACTERISTICS OR PROPERTIES OF ISOQUANTS

A smooth continuous isoquant that has been adopted in the traditional economic theory possesses the following characteristics:

1) lsoquants are negatively sloped

2) A higher isoquant represents a larger output

3) No two isoquants intersect or touch each other

4) lsoquants are convex to the origin.

1) **lsoquants are negatively sloped**

Normally, isoquants slope downwards from left to right implying that they are negatively sloped. The reason for this characteristic of the isoquant is that when the quantity of one factor is reduced, the same level of output can be achieved only when the quantity of the other is increased. This characteristic of the isoquant, however, assumes that in no case marginal productivity of a factor will be negative. In a more realistic case when this assumption is dropped, one may find an isoquant which bends back upon itself or has a positively sloped segment. in Fig. 8.5, such an isoquant is shown. AB and CD segments of this isoquant are positively sloped.

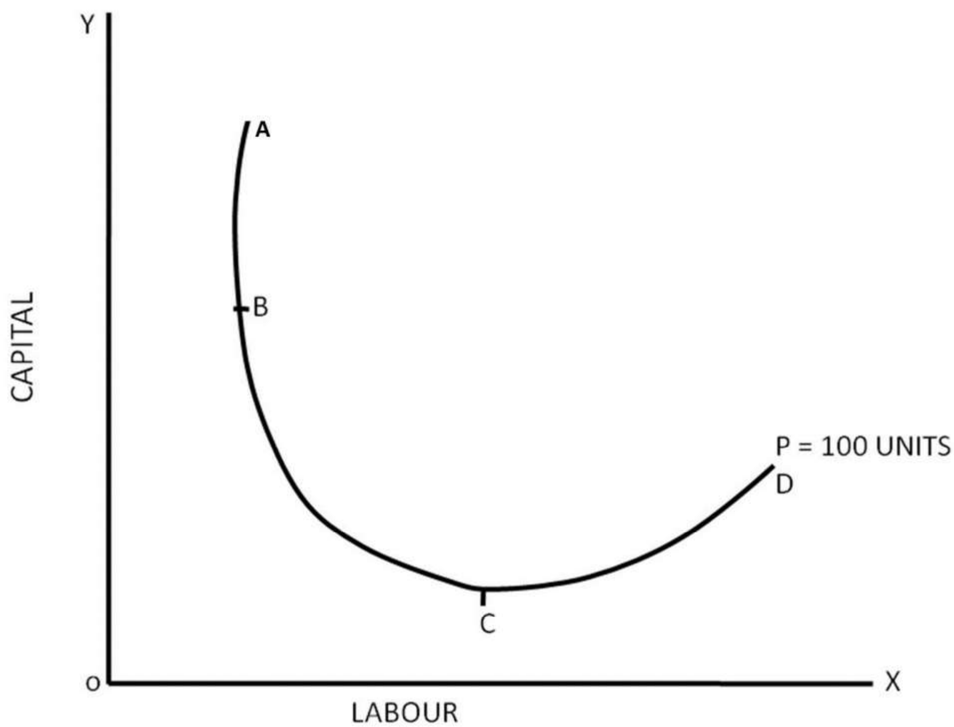

Fig. 8.5: Isoquant having positively sloped segments

2) **A higher isoquant represents a larger output**

A higher isoquant is one that is farther from the point of origin. It represents a larger output that is obtained by using either the same amount of one factor and the greater amount of the other factor or the greater amounts of both the factors. Two isoquants P_1 and P_2 have been shown in Fig. 8.6. They depict output levels of 100 units and 200 units. Obviously, the output level represented by isoquant P_2 can be reached only by using more of factor inputs as compared to the amount of factor inputs required to reach output level represented by isoquant P_1.

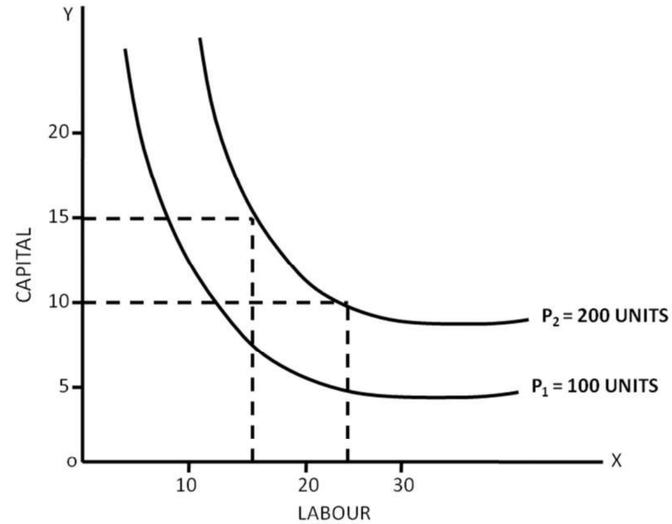

Fig. 8.6: Two Isoquants representing different output levels. A higher isoquant depicts a higher amount of output

3) No two isoquants intersect or touch each other

Isoquants do not intersect or touch each other because they represent different levels of output. If, for example, isoquants P_1 and P_2 (Fig. 8.7) represent output levels of 100 and 200 units respectively, their intersection at some point, say A would mean that two output levels (i.e, 100 and 200 units) will be reached by using the same amount of capital and labour which is not likely to happen. For the same reason, no two isoquants will touch each other.

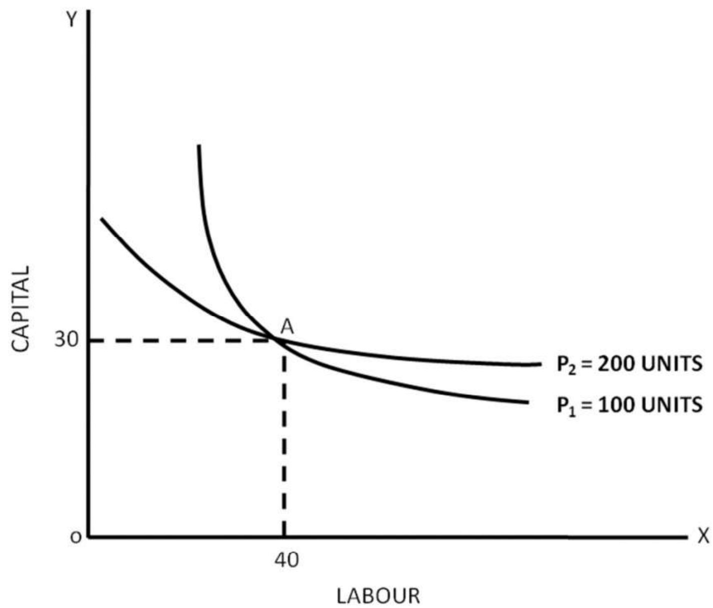

Fig. 8.7: No two isoquants intersect one another because each isoquant depicts a different level of output

4) Isoquants are convex to the origin

In most production processes, the factors of production have substitutability. Often, labour can be substituted for capital and vice versa. However, the rate at which one factor of production is substituted for the other in a production process, that is, the marginal rate of technical substitution **(MRTS)** often tends to fall.

Marginal rate of technical substitution of factor L for factor K ($MRTS_{L,K}$) is the quantity of K that is to be reduced on increasing the quantity of L by one unit for keeping the output level unchanged.

The isoquants are convex to the origin precisely because the marginal rate of technical substitution tends to fall. Let us explain why this happens with the help of Fig. 8 8. Here, the isoquant is curve P. Let us suppose that the producer is at point 'a' of the curve. The meaning of this is that he uses OJ units of capital and OR units of labour to produce 100 units of output. We shall assume that one unit of labour is OR = RS = ST = TU = UV. Now, if he wants to increase the amount of labour by RS, and keep the output at 100 units, he must reduce the use of capital by JK. Similarly, when he increases the amount of labour by ST, TU and UV, he must reduce the application of capital by KL, LM and MN respectively if output has to be kept at the same level (i.e., 100 units). It is clear from the figure that JK > KL > LM > MV. In other words, as additional units of labour are employed it becomes progressively more and more difficult to substitute labour in place of capital so that lesser and lesser units of capital can be replaced by additional units of labour. This means that the marginal rate of technical substitution tends to fall. This is due to the reason that factors of production are not perfect substitutes for one another. When the quantity of one factor is reduced, it becomes necessary to increase the quantity of the other at an increasing rate. For example, let us suppose that in a particular productive activity two factors of production – labour and capital– are employed. When the quantity of labour employed is reduced by one unit, it is possible to undertake the activity by employing one more unit of capital initially. However, when one more unit of labour is reduced, it might become necessary to compensate this by employing, say, two units of capital. As the quantity of labour employed is reduced successively at each stage, we would require more and more units of capital to compensate for the loss of each additional unit of labour.

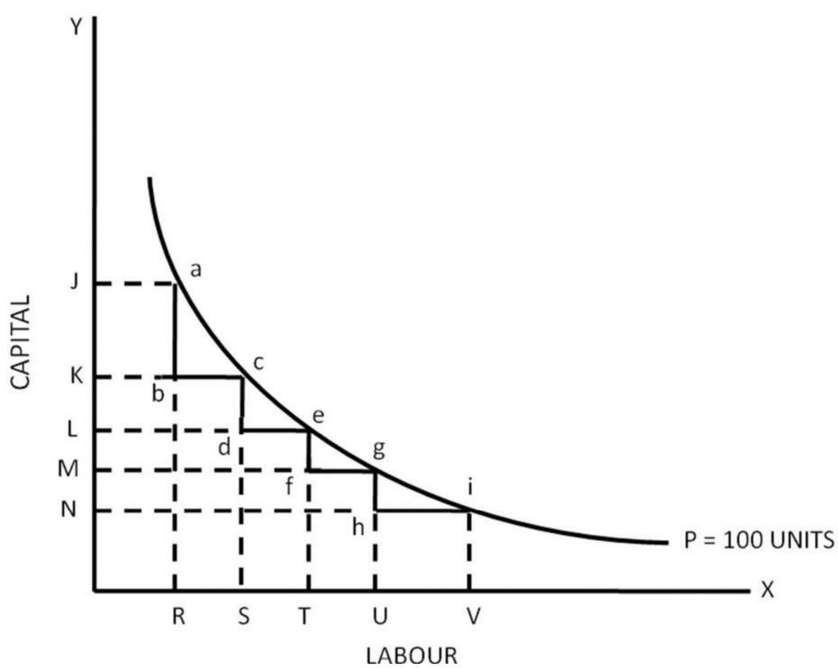

Fig. 8.8: An isoquant is covex from below because the marginal rates of technical substitution tends to fall

If the factors of production are perfect substitutes, the marginal rate of technical substitution between them would be constant and the isoquant will be linear and sloping downwards from left, to right as in Fig. 8.2. In the case of strict complementarity, that is, zero substitutability of the factors of production the isoquant will be right angled or we may say that it will assume the shape of 'L' as in Fig. 8.3. However, the linear and right angled isoquants are the limiting cases in the production processes.

Check Your Progress 1

1) Indicate the following statements as true (T) or false (F):

   i) In case of perfect substitability of the factors of production, the isoquant is convex from below.

   ii) Isoquants are positively sloped.

   iii) A higher isoquant represents a larger output.

   iv) No two isoquants intersect each other.

2) Define isoquant. Discuss its properties.

   ....

   ....

   ....

3) Discuss the possible shapes which the isoquants may assume depending on the degree of substitutability.

   ....

   ....

   ....

8.5 ECONOMIC REGION OF PRODUCTION AND RIDGE LINES

Generally, production functions generate isoquants which are convex to the origin, negatively sloped throughout, do not intersect each other and the higher the isoquants, greater the level of output. However, there are some production functions which yield isoquants having all the properties of a normal isoquant except that they are not negatively sloped throughout. In other words, they have positively sloped segments. In Fig. 8.9, the production function is depicted in the form of a set of isoquants which have positively sloped segments.

Let us consider isoquant P_3. AB segment of this isoquant has a negative slope. Beyond points A and B, this isoquant is positively sloped. Similarly, other isoquants have the points where they bend back upon themselves implying that they become positively sloped. The lines OK and OL joining these points are called ridge lines. They form the boundaries for the economic region of production. A careful interpretation of any of the isquants in Fig. 8 .9 will make this point clear.

Suppose the output represented by isoquant P_3 is to be produced. For producing this quantity, a minimum of OK_2 amount of capital is required because any smaller amount will not allow the producer to attain the P_3 level of output. With OK_2 amount of capital, OL_2 amount of labour must be employed. In case the producer uses an amount of labour less than OL_2 together with OK_2 amount of capital, his output level would be lower than the one represented by isoquant P_3. This is quite normal, because use of inputs in smaller amounts would yield a smaller output. But combining labour input in an amount larger than OL_2 with OK_2 amount of capital would also result in output smaller than that is represented by the isoquant P_3. In order to maintain the P_3 level of output with a larger labour input, capital input also in a larger amount has to be used. Obviously, this is something which no rational producer would attempt because it involves uneconomic use of resources.

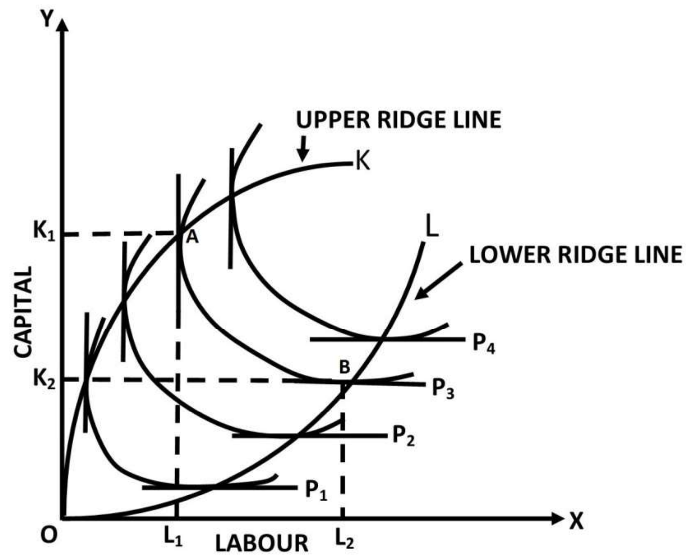

Fig. 8.9: Area enclosed within the upper side line OK and the lower side lint OL indicates the economic region of production

Point B on isoquant P_3 represents the intensive margin for labour because an increase in the amount of labour input beyond OL_2 with a fixed amount of capital input OK_2 does not increase the output level. At this point, marginal product of labour is zero and thus the marginal rate of technical substitution of labour for capital ($MRTS_{LK}$) is zero. This implies that at point B labour has been substituted for capital to the maximum extent. Thus, to the right of ridge line OL in Fig. 8.9, we have Stage III for labour.

Similarly, for producing P_3 level of output, a minimum of OL_1 amount of labour input is required. A smaller amount of labour input will not allow the producer to attain P_3 level of output. With OL_1 amount of labour, OK_1 amount of capital must be used and any additions to capital input beyond OK_1would not increase output. This point represents intensive margin for capital because an increase in the amount of capital input beyond OK_1 with a fixed labour input of OL_1 does not augment output. At point A on P_3, capital has been substituted for labour to the maximum extent. Thus, above ridge line OK in Fig. 8.9, we have Stage III for capital. The marginal rate of technical substitution of capital for labour (MRTSKL) is zero, which means that the marginal rate of technical substitution of labour for capital (MRTSKL) is infinite or undefined.

The line OK in Fig. 8.9 connects the point of zero marginal product of capital. We have designated it as the **upper ridge line.** Similarly, the line OL designated as the **lower ridge line** joins the points of zero marginal product of labour.

The combinations of labour and capital inputs comprising the area between ridge lines OK and OL constitute the generalised Stage II of production for both resources. These are the combinations that are relevant for production decisions.

8.6 THE OPTIMUM COMBINATION OF FACTORS AND PRODUCER'S EQUILIBRIUM

So far, we have explained as to how different combinations of inputs allow a producer to attain a certain level of output. The producer is free to choose any of these input combinations. However, his choice cannot be arbitrary if he wishes to minimise cost of producing a stipulated output. Our task now is to explain how the producer selects a particular input combination.

8.6.1 Input Prices and Isocost Lines

A producer may attempt maximisation of output subject to a given cost or alternatively, he may seek to minimise cost subject to a given level of output. In both cases, for choosing optimum quantities of two inputs, viz., labour and capital, he must consider their physical productivities as well as their prices. While isoquants represent the productivities of the inputs, their prices are shown by isocost lines.

An isocost line represents various combinations of inputs that may be purchased for a given amount of expenditure; that is, the producer's budget.

The firm or the producer has to purchase factors or inputs from the market. How the prices of labour and capital are determined in the market is not our present concern. Moreover, the firm is in no position to influesence the input prices unless it is a monopsonist or oligopsonist. In other words, prices of labour and capital have to be taken as given by the firm operating in a competitive factor market. Let us now suppose that the firm's total cost outlay on labour and capital is Rs. 1000. The firm is free to spend this entire amount on labour or capital or it may spend it on a combination of both labour and capital. In Fig. 8.10, we have shown that if the firm chooses to spent the entire amount of Rs. 1,000 on labour input, it can employ OL_2 amount of labour, and if the entire amount is to be spent on capital, it can get OK_2 amount of capital. The straight line $K_2 L_2$ is an isocost line representing all the combinations of capital and labour which the firm can obtain for Rs. 1,000. In the figure, the length of OL_2 is twice the length of OK_2 which means that the price of a unit of labour is half that of a unit of capital. The slope of the line $K_2 L_2$ shows the ratio of input prices. Hence, the slope of an isocost line is (w/r), which is the ratio of the price of labour (w) to the price of capital (r) when X-axis denotes labour input and Y-axis denotes capital input. We can thus generalise that for any isocost line which is always linear because the firm has no control over the prices of inputs and the prices remain the same, no matter how much quantity of these inputs, the firm buys,

$$\textbf{Slope} = \frac{\Delta K}{\Delta L} = \frac{K}{L} = \frac{\text{expenditure}}{r} / \frac{\text{expenditure}}{w} = \frac{w}{r}$$

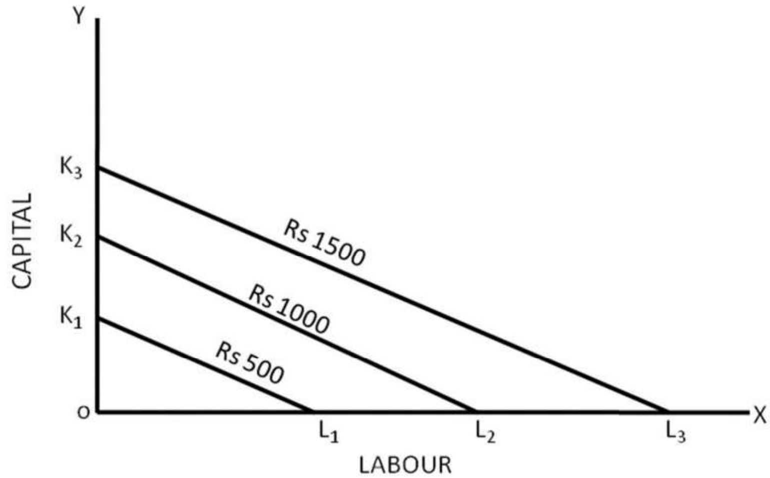

Fig. 8.10: Isocost Lines - A higher cost line indicates a higher cost

This property of an isocost line is similar to that of the budget line of the consumer. However, there is an important difference between the two lines. Since the consumer's budget is invariably fixed, he has a single budget line. The firm generally has no such constraint and thus has more than one isocost lines. In Fig. 8.10, we have shown three isocost lines. There can be many more of them corresponding to firm's cost outlay plans to attain various output levels.

An isocost line farther to the right reflects higher costs; the one closer to the origin reflects lower costs.

8.6.2 Maximisation of Output for a Given Cost

A rational producer is expected to maximise output for a given cost. Alternatively, he may attempt to minimise cost subject to a given level of output.

In this section, we shall explain how a producer maximises his output for a given cost. Suppose the producer's cost outlay is C and the prices of capital and labour are r and w respectively. Subject to these cost conditions, the producer would attempt to attain the maximum output level.

Let KL isocost line in Fig. 8.11 represent the given cost outlay at input prices r and w. P_1, P_2 and P_3, are isoquants representing three different levels of output. It may be noted that P_3 level of output is not attainable because the available factor resources (various labour-capital combinations represented by isocost line KL) are insufficient to reach that output level. In fact, any output level beyond isocost line KL is not attainable. The producer, however, can attain any output level in the region OKL, but that would not require all the resources (labour and capital inputs) that are available to the producer for his cost outlay.

Therefore, in the case of a given cost, the producer's attempt would be to reach the isoquant which represents the maximum output level. The producer can operate at points such as R and T. At these two points, the combinations of labour and capital to produce P_1 level of output are available for a given cost represented by isocost line KL. In contrast, at point S, the combination of labour and capital available for the same cost (as it is also on isocost line KL) enables the producer to reach isoquant P_2 which represents an output level higher than that represented by P_1. Since at point S on isoquant P_2 is jus tangent to isocost line, a greater output than P_2 is not obtainable for the given level of cost. A lesser output is not efficient because production can be raised without incurring additional cost. Hence, the optimal combination of factors of production, viz., capital and labour is $OK_2$of capital plus OL_2 of labour as it enables the producer to reach the highest level of production possible given the cost conditions.

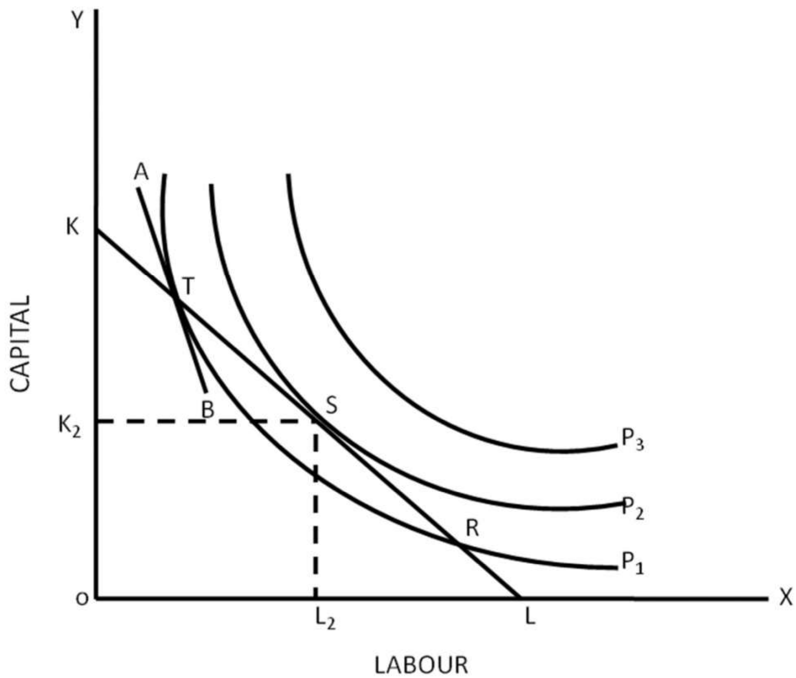

Fig. 8.11: With the given cost line KL, the highest isoquant that a producer can reach is P_2. Point S on this isoquant, therefore, indicates producer's equilibrium

The above proposition should be obvious to those who have studied the theory of consumer behaviour. At the same time, the reason that lies behind it must be followed carefully. Let us suppose that the producer wishes to produce at point T. The marginal rate of technical substitution of labour for capital indicated by the slope of tangent AB at point T is relatively high. Suppose ΔK is equal to 3 and ΔL is equal to 1. Thus, the slope of tangent AB is 3:1 which implies that at point T one unit of labour can replace 3 units of capital. However, the relative factor price indicated by the slope of KL is less, say, 0.7:1 which means that the cost of 1 unit of labour is the same as the cost of 0.7 unit of capital. Therefore, it would be rational on the part of the producer that he substitutes labour for capital so long as the marginal rate of substitution of labour for capital is not equal to the factor price ratio, that is, the ratio of the price of labour to the price of capital. At point R, the opposite situation prevails because the marginal rate of technical substitution is less than the factor price ratio.

The producer maximises output for a given cost (reaches equilibrium) only when the marginal rate of technical substitution of labour for capital is equal to the ratio of the price of labour to the price of capital.

Thus,

$$MRTS_{LK} = \frac{w}{r} = \frac{MP_L}{MP_K}$$

8.6.3 Minimisation of Cost for a Given Level of Output

If a producer seeks to minimise the cost of producing a given amount of output rather than maximising output for a stipulated cost, the condition of his equilibrium remains formally the same. That is, the marginal rate of technical substitution must be equal to the factor price ratio.

This can be easily followed graphically. In Fig. 8.12, we have a single isoquant P which denotes the desired level of output, but there is a set of isocost lines representing various levels of total cost outlay. An isocost line closer to origin indicates a lower total cost outlay. The isocost lines are parallel and thus have the same slope w/r because they have been drawn on the assumption of constant prices of factors.

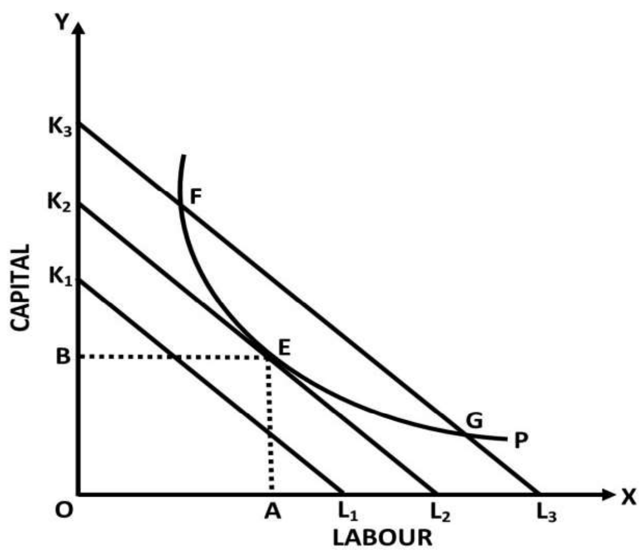

Fig. 8.12: To obtain a level of production indicated by isoquant P, the minimum cost that must be incurred is given by point E on the isocost line K_2L_2. Therefore, point E indicates the point of producer's equilibrium

It may be noted that isocost line $K_1 L_1$ is just not relevant because the output level represented by the isoquant P is not producible by any factor combination available on this isocost line. However, the P level output can be produced by the factor combinations represented by the points F and G which are on isocost line $K_3 L_3$. Alternatively, the producer can attain the P level output by the factor combination represented by the point E which is on isocost line $K_2 L_2$. Since the isocost line $K_2 L_2$ is closer to the origin as compared to the isocost line $K_3 L_3$, it represents relatively lower cost. Therefore, by moving either from F to E or from G to E, the producer attains the same output level at a lower cost. The producer thus minimises his costs by employing OB amount of capital plus OA amount of labour determined by the tangency of the isoquant P with the isocost line $K_2 L_2$. Points representing factor combinations below E are certainly preferable because they represent lower costs but they cannot be considered as they cannot help in producing the output level represented by the isoquant P. Points above E represent higher costs. Hence, point E denotes the least cost combination of the factors, viz., labour and capital for producing output shown by isoquant P. This discussion thus leads us to the principle that in the case of producer's equilibrium, the marginal rate of technical substitution of labour for capital must be equal to the ratio of the price of labour to the price of capital. We can now sum up the whole discussion as follows:

1) **The optimal combination of factors, whether the producer seeks to maximise output for a given cost or he wishes to minimise cost for a stipulated output, is that where marginal rate of technical substitution and the factor price ratio are equal.**
2) **The producer is in equilibrium when there is optimal combination of factors.**

8.7 THE EXPANSION PATH

Producers expand their outputs both in the long run and in the short run. In the long run, output expands with all factors variable, while in the short run, expansion of output is possible with some factor(s) constant and some others variable. We shall consider both cases.

8.7.1 Optimal Expansion Path in the Long Run

In the long run, there is no limitation to the expansion of output as all the factors of production are variable. The firm's goal being maximisation of its profits, it seeks to expand outputs in the optimal way. With given factor prices, the optimal expansion path is the locus of the points of tangency of successive isocost lines and successive isoquants.

Consider now Fig. 8.13. Given the factor prices, the output corresponding to isoquant P_1 is producible at the lowest cost at point A where isocost line $K_1 L_1$ is tangent to the isoquant P_1. This is the initial position of producer equilibrium. Assuming that factor prices remain constant, suppose the producer desires to expand output to the level indicated by the isoquant P_2. This will cause a shift in the isocost line from $K_1 L_1$ to $K_2 L_2$. The new equilibrium is found at point B where isocost line $K_2 L_2$ is tangent to the isoquant P_2. Further expansion in output to the level corresponding to the isoquant P_3 will shift equilibrium to point C where isocost line $K_3 L_3$ is tangent to the isoquant P_3.

On connecting all points of producer equilibrium, such as A B and C, we get the curve OE which is called the expansion path. **Since every point of the expansion path denotes an equilibrium point of the producer, it indicates the optimum combination of factors of production of some particular level**

of output. It may be recalled that each point of producer equilibrium is defined by equality between the marginal rate of technical substitution and the factor price ratio. Since the latter has been assumed to remain constant, the former also remains constant. Hence, OE is an isocline along which output expands when factor prices remain constant.

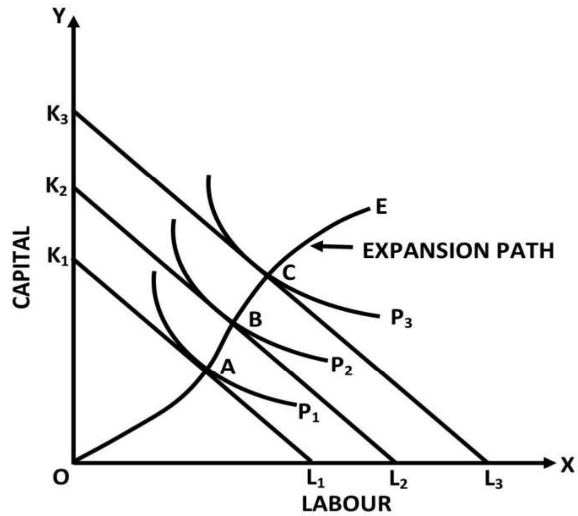

Fig. 8.13: Expansion path in the case of non-linear production function

In the case of linear homogeneous production function, the isoclines are straight lines through the origin. Therefore, the expansion path will also be a straight line as shown in Fig. 8.14. This means that given the prices of the factors of production, the optimal proportion of the inputs of the firm will not change with the size of the firm's output or input budget.

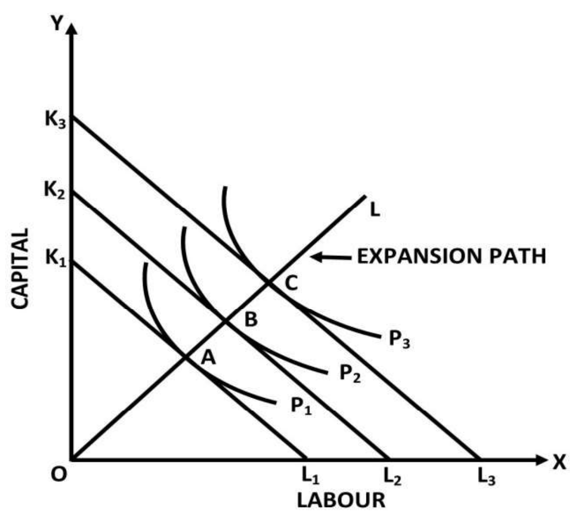

Fig. 8.14: Expansion path in the case of linear homogeneous production function is a straight line

The line formed by connecting the points determined by the tangency between the successive isoquants and the successive isocost lines is the firm's expansion path. It identifies the least costly input combination for each level of output and will slope upward in the long-run setting. This means that the firm will expand use of both inputs as it expands its output.

8.7.2 Optimal Expansion Path in the Short Run

In the short run, capital is a fixed factor and thus its amount remains constant. Labour is, however, variable and the producer can expand his output by increasing the amount of labour along a straight line parallel to the axis on which this factor is measured. In Fig. 8.15, the straight line AB indicates the expansion path as the total amount of capital is fixed at OA in the short run.

With the prices of the factors of production remaining constant, the firm cannot maximise its profits while it expands its output in the short run, on account of the constraint of the fixed amount of capital. This can be followed from Fig. 8.15. The firm's initial equilibrium is at point E where isocost line K_1L_1 is tangent to the isoquant P_1. If the firm wishes to raise its output level corresponding to the isoquant P_2, it reaches the point F which, given the factor prices, is not the least cost situation. Further expansion of output to the level corresponding to the isoquant P_3 leads the firm to reach the point G which again does not represent the least cost situation. The optimal expansion path would be OR, were it possible for the firm to increase the quantity of capital. However, given the amount of capital, the firm has no choice but to expand along the straight line AB in the short run.

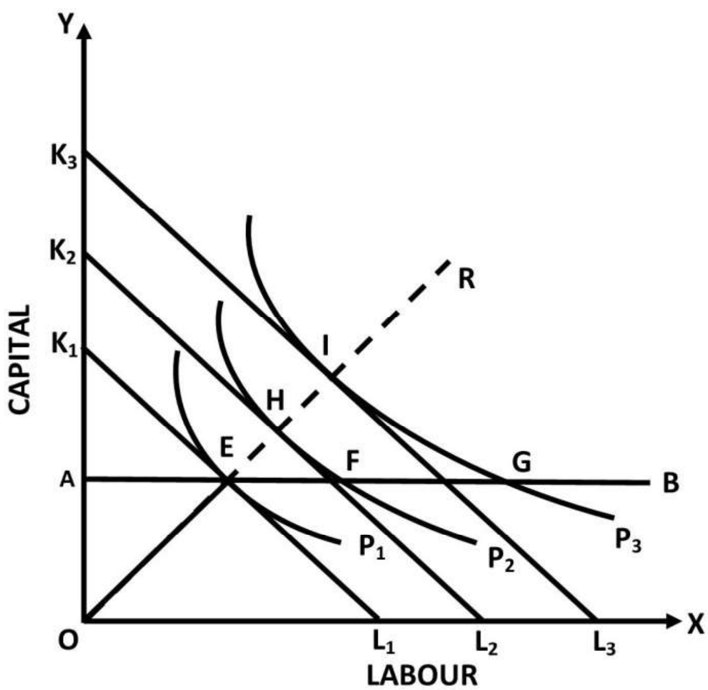

Fig. 8.15: Expansion path in the short run in the case of linear homogeneous production function

Check Your Progress 2

1) Indicate the following statements as True (T) or False (F):

   i) The condition for optimal combination is that marginal rate of technical substitution is greater than factor price ratio.

ii) The area between ridge lines constitutes the Stage II of production for both resources.

iii) An isocost line represents various combinations of input that may be purchased for a given amount of expenditure.

iv) An isocost line farther to the right reflects higher cost.

v) Every point on the expansion path denotes an equilibrium point of the producer.

vi) The line formed by connecting the points determined by the tangancy between the successive isoquants and the successive iocost lines is the firm's expansion path.

2) Explain the condition of a producer's equilibrium.

....

....

....

3) Suppose that P_K= Rs. 10, P_L = Rs. 20 and TO (total outlay) = Rs. 160.

i) What is the slope of the isocost ?

ii) Write the equation of the isocost?

....

....

....

4) Explain the significance of tangency between an isoquant and an isocost line.

....

....

....

5) Explain why, for a least cost combination of inputs, a firm requires that the marginal rate of technical substitution be equal to the input ratio.

....

....

....

6) What is meant by a firm's expansion path? Distinguish between the expansion path in respect of a linear homogeneous production function from the expansion path in respect of a non-linear production function.

....

....

....

8.8 LET US SUM UP

The unit begins with the concept of Production function which refers to functional relationship between inputs and output. This is followed by the definition of an isoquant and the explanation of three types of isoquant- (i) convex isoquant, (ii) linear isoquant, and (iii) input-output isoquant. The properties of isoquants are: (i) isoquant are negatively sloped (ii) a higher isoquant represents a larger output, (iii) no two isoquants intersect or touch each other, and (iv) isoquants are convex to the origin. From here we proceed to a discussion of the concept of the economic region of production and ridge lines. The next section is devoted to a discussion of the optimum combination of factors and producer's equilibrium. In this section, we first consider the concept of isocost lines and then consider (i) maxmisation of output for a given cost, and (ii) minimisaton of cost for a given level of output. The last section of the chapter discusses the optimal expansion path for a firm both under long run and short run.

8.9 REFERENCES

1) Robert S Pindyck, Daniel L. Rubinfld and Prem L Mehta, *Microeconomics* (Pearson Education, Seventh Edition, 2009), Chapter 5, Section 5.1 and Section 5.3.
2) Dominick Salvatore, *Principles of Microeconomics* (Oxford University Press, Fifth Edition, 2010), Chapter 7, Section 7.1, Section 7.3 and Section 7.4.
3) A.Koutsoyiannis, *Modern Microeconomics* (The Macmillan Ltd., Second edition, 1982). Chapter 3.
4) John P. Gould and Edward P. Lazear, *Microeconomic Theory* (All India Traveller Bookseller, Sixth edition, 1996), Chapter 7.

8.10 ANSWERS OR HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) i) F) ii) F iii) T iv) T

2) See Sub-section 8.3.1 of Section 8.3 and Section 8.4

3) See Sub-section 8.3.2 of Section 8.3

Check Your Progress 2

1) i) F ii) T iii) T iv) T v) T vi) T

2) See Section 8.6

3) (i) Slope of isocost is $-P_L/P_K = -2$ and the eqution is $160 = 10K + 20L$ or $16 = K + 2L$ or $K = 16 - 2L$

4) See Section 8.6

5) See Sub-section 8.6.3 of Section 8.6

6) See Section 8.7

UNIT 9 RETURNS TO SCALE

Structure

9.0 OBJECTIVES

After going through this unit, you should be able to :

- state the concept of returns to scale;
- distinguish between the stage of increasing, constant and diminishing returns to scale; and
- explain the concepts of economies and diseconomies of scale (both internal and external).

9.1 INTRODUCTION

Sometimes to increase the level of output, all factors are increased simultaneously and factor proportions are held constant. This is known as expansion in scale. In this context, three phases of production are discussed: increasing returns to scale, constant returns to scale, and diminishing returns to scale. Expansion of scale confers a number of economies i.e. advantages on the firm – both internal and external. Internal economies, in turn, can be divided into real internal economies of scale and pecuniary internal economies. If the scale of production is continuously expanded, a stage of internal diseconomies of scale sets in i.e. after a certain point, increase in production is less than proportionate increase in the factors of production. In this unit, we propose to discuss all these issues. We shall also explain the concept of external economies and external diseconomies.

*Dr. V.K. Puri, Associate Professor of Economics, Shyam Lal College (University of Delhi) Delhi.

9.2 CONCEPT OF RETURNS TO SCALE

The concept of returns to scale is associated with the tendency of production that is observed when the ratio between the factors is kept constant but the scale is expanded, that is use of all the factors is changed in same proportion.

When all the factors of production (labour, capital, etc.) are increased in the conditions of constant techniques, three possibilities arise:

1) Output increases in a greater proportion as compared to the increase in the factors of production. This is the case of **increasing returns to scale.**

2) Output increases in the same proportion as the increase in the amount of the factors of production. This is the case of **constant returns to scale.**

3) Output increases in a smaller proportion as compared to the increase in the amounts of the factors of production. This is the case of **diminishing returns to scale.**

We can illustrate these three situations with the help of numerical examples as follows:

OUTPUT SCHEDULE-1				
Input-X	**Input -Y**	**% Change in Inputs**	**Output**	**% Change in Output**
2	4	100	1000	-
4	8	100	3000	200
8	16	100	10000	233
16	32	100	35000	250

It can be observed from the given output schedule that:

1) At all the stages, we are increasing the quantity of inputs by 100%.
2) With increase in the quantity of inputs, the quantity of output is increasing by more than 100% at all stages. In other words output is increasing proportionately more than the increase in input.

Compare this situation with the illustrations given below:

OUTPUT SCHEDULE-2				
Input-X	**Input -Y**	**% Change in Inputs**	**Output**	**% Change in Output**
2	4	100	1000	-
4	8	100	2000	100
8	16	100	4000	100
16	32	100	8000	100

Output shedule-2 indicates that:

1) Inputs increase by 100% at each stage.
2) Output also increased by 100% at each stage.

In this illustration, output increased by the same proportion in which inputs have been increased.

Compare this situation with the one given below:

OUTPUT SCHEDULE-3				
Input-X	**Input -Y**	**% Change in Inputs**	**Output**	**% Change in Output**
2	4	100	1000	-
4	8	100	1800	80
8	16	100	2500	39
16	32	100	3000	20

It would be observed that the total output at all stages increases less proportionately than the increase in inputs.

9.2.1 Increasing Returns to Scale

When the ratio between the factors of production is kept fixed and the scale is expanded, initially output increases in a greater proportion than the increase in the factors of production.

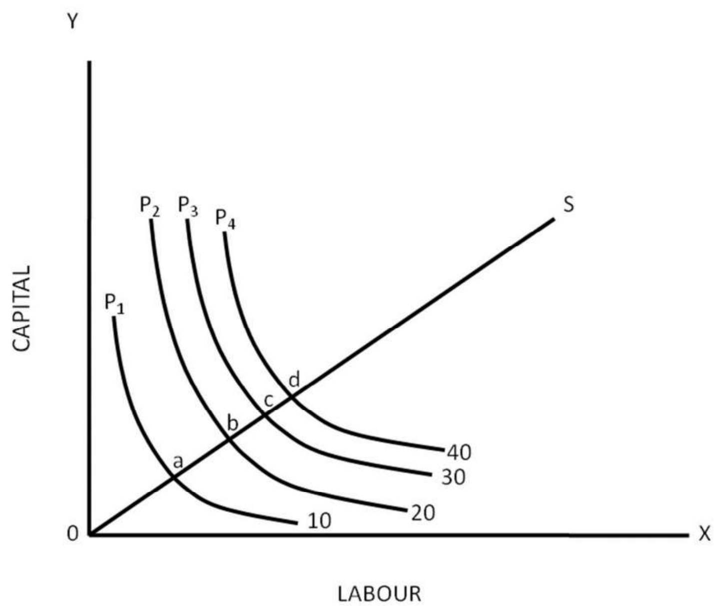

Fig. 9.1: Increasing returns to scale output increases in a greater proportion than the increase in the factors of production

For example, if factors are doubled the output is more than doubled. In other words, to double the quantity of the output, it is not necessary to double the quantity of the factors of production. This can be understood with the help of Fig. 9.1. In this figure P_1, P_2, P_3, P_4 are isoquants. They show 10, 20, 30, 40 units of output respectively. OS is the scale line which is cut by the isoquants at unequal distances. In the figure, it can be seen that cd < bc < ab < oa. This means that to enable the firm to rise from isoquant P_1 to P_2 (so that production increases from 10 to 20 units), the amount of factors of production required is less than the amount required to produce the initial l0 units of output.

Similarly, to increase the output further by 10 units so as to reach isoquant P_3, the amount of factors of production required is less than the amount required to produce the earlier 10 units of output as bc < ab. This position seems to hold true till isoquant P_4. There are three main factors which account for increasing returns to scale:

1) **Indivisibility:** The most important reason of increasing returns to scale is the 'technical and managerial indivisibilities'. The meaning of an indivisible factor of production is that there is a certain minimum size of the factor and even if it is large in relation to the size of the output, it has to be used (i.e., it cannot be divided). For example, even if only 10-15 letters are to be despatched from an office, it would be necessary to keep a typewriter. It is not possible to purchase only half the typewriter since only a small number of letters have to be typed daily. We would, therefore, say that typewriter is not divisible. In a similar way, plants and managerial services in modern factories are not divisible. Accordingly, when the scale of production is enlarged initially, there is no equi-proportionate increase in the demand for the factors of the production.

2) **Specialisation:** Chamberlin does not regard indivisibility as an important cause of 'increasing returns to scale'. According to him, the main reason of increasing returns to scale is specialisation. When due to division of labour, workers are given jobs according to their ability, their productivity increases while cost declines. According to Donald S. Watson, acknowledgement of this fact contradicts the assumption that the ratio of different factors of production remains constant. Accordingly, he casts doubts whether specialisation can be regarded as leading to increasing returns to scale. The importance of specialisation can be accepted only if we assume that although an increase by an equal amount in quantity of labour and capital employed is necessary for an expansion in scale, this increase does not mean the doubling or trebling their units employed but it does mean an increase in their fixed money cost. But this can lead to technical changes and it is very much possible that increasing returns emerge not due to an expansion in scale but due to technical reasons.

9.2.2 Constant Returns to Scale

Increasing returns to scale can be obtained only upto a point. After this point is reached, expansion of scale only leads to equal proportionate change in output.

Empirical evidence suggests that the phase of constant returns is a fairly long one and is observed in the case of a number of commodities. In a scientific sense, constant returns to scale implies that when the quantity of the factors of production is increased in such a way that the ratio of the factors remains unchanged, output increases in the same proportion in which the factors are increased. In other words, when the quantity of the factors is doubled, the output also doubles. Such a production function is often called linear homogeneous production function or homogeneous production function of the first degree. The phase of constant returns to scale can be understood with the help of Fig. 9.2. In this figure, when the firm goes from isoquant P_3 to P_4, or from isoquant P_4 to P_5 or from isoquant P_5 to P_6, constant returns to scale are obtained. The fact cd = de = ef on the scale line indicates this phenomenon.

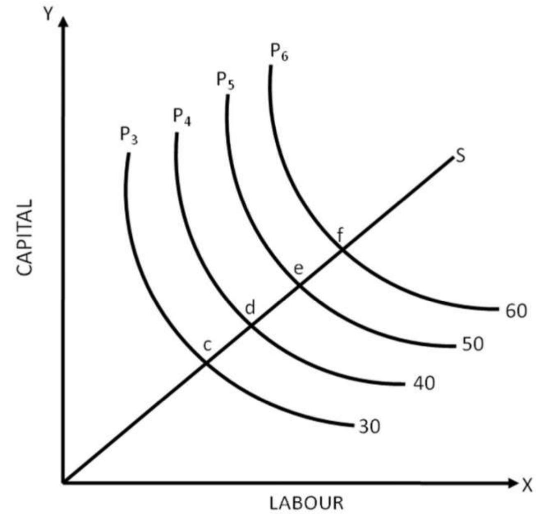

Fig. 9.2: Constant returns to scale-output increases in the same proportion in which inputs are increased

The question that now arises is what are the reasons which account for constant returns to scale. Generally when inefficiencies of production on a small scale are overcome and no problems regarding technical and managerial indivisibilities remain, expansion in scale leads to a situation where returns increase in the same proportion as the factors of production. Some economists are of the view that when benefits of specialisation of a factor in the unit of production are small or when such benefits have already been reaped at a small level of production, then for a considerable period of time, production increases according to the law of constant returns to scale.

Economists have argued that if the factors of production are perfectly divisible, the production function must exhibit constant returns to scale. In their opinion, if constant returns to scale does not prevail in some industries, it is because in these industries either due to scarcity or indivisibility of some factors, it is not possible to vary all them in the same proportion. Indivisibility of a factor often results in its under-utilisation at lower levels of output. When a producer for obtaining a larger output increases quantities of other factors, the amount of the lumpy factor which had not been fully utilised at lower levels of output, will not be increased. These economists do not think that economies of scale will be available when the factors of production are perfectly divisible. They however, stress the role of optimum factor proportionality in production. When factors of production are perfectly divisible, they can be increased or decreased in such amounts that an optimum proportion between factors is achieved. The output can be increased or decreased by increasing, or decreasing the amounts of the factors in the optimum proportion without any economies or diseconomies of scale which means that constant returns to scale will necessarily prevail.

9.2.3 Diminishing Returns to Scale

Diminishing returns to scale ensure that the size of the productive firms cannot be infinitely large. Generally after a limit when the quantity of the factors of production is increased in such a way that the proportion of the factors remains unchanged, output increases in a smaller proportion as compared to increases in the amounts of the factors of production. For example, it may happen that an increase in amount of labour and capital by 100 per cent leads to an increase in output by only 75 per cent. In other words, if output has to be doubled, the factors of production will have to be more than doubled. We can understand this phenomenon with the help of Fig. 9.3. In this figure, when the firm is at isoquant P_6, the tendency of constant returns to scale has come to an end. From here, the increasing distance between two consecutive isoquants is an indication that to obtain the same increase in output, factors of production will have to be increased at a higher and higher rate. On the scale line OS, ij > hi > gh >fg > ef indicate this phenomenon very explicitly.

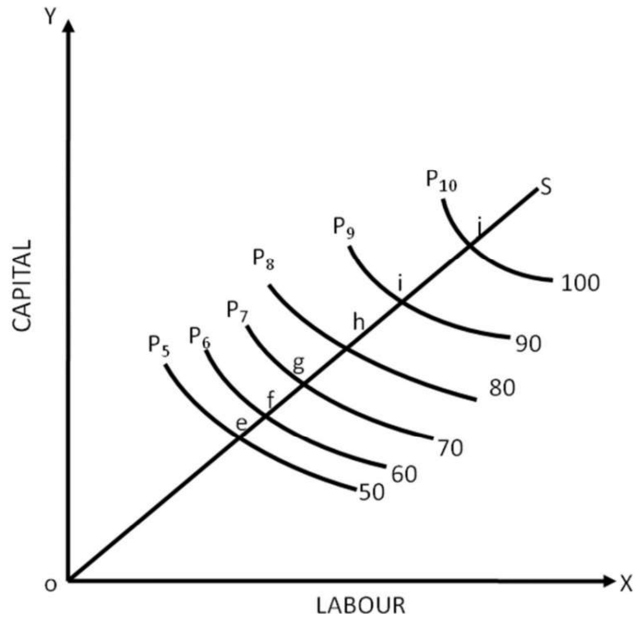

Fig. 9.3: Diminishing returns to scale – output increases proportionally less than inputs

Economists do not agree on the causes which leads to operation of diminishing returns to scale. Nevertheless, the two causes that are often mentioned are as follows:

1) **Enterprise:** Some economists emphasise that enterprise is a constant and indivisible factor of production and its supply cannot be increased even in the long run. Accordingly, when the quantity of other factors is increased and the scale of production expanded in a bid to boost up production, the proportion of other factors in relation to enterprise increases. Beyond a certain point, this results in diminishing returns as enterprise becomes scarce in relation to other factors.

2) **Managerial difficulties:** According to some other economists, the main reason for the operation of diminishing returns to scale is managerial difficulties. When the scale of production expands, the co-ordination and control on different factors of production tend to become weak and therefore output fails to increase in the same proportion as the factors of production increase. This results in diminishing returns to scale.

9.3 ECONOMIES AND DISECONOMIES OF SCALE

Expansion of the scale confers a number of economies on the firm. Some of these are in 'real terms' while others are in 'pecuniary terms'. Economies that are obtained in production work, marketing, management, transport, etc. are in real terms, while economies that are obtained in terms of, say, purchase of inputs at wholesale rate, availability of finance at lower rate of interest, saving on advertisement costs, etc. are in money terms. Then, there are certain economies that do not accrue to the firm whose scale of operation is large but accrue to certain other firms which benefit from the large scale of this firm.

In Economics, those economies which accrue to a firm on expansion of its own size are known as internal economies. As against this, those economies which accrue to a firm not due to its own operations but due to the operations of other firms are termed external economies.

9.3.1 Internal Economies of Scale

Generally, when the scale of production is sought to be enlarged, the firm replaces its small plant by a larger plant. This increases the efficiency of production. However, it is not always necessary to change the plant for expanding the scale of production. The firm can keep the old plant in a running condition and either establish a new plant of the same type or a new plant of some new type. In all these alternatives, the firm obtains many different kinds of economies. The fact is that it is the economies of scale that determine the nature of the long-run average cost curve.

9.3.1.1 Real Internal Economies of Scale

When expansion in the scale of production takes place, the firm obtains some real internal economies. These economies accrue in the form of saving in the physical quantities of raw materials, labour, fixed and variable capital, and other inputs. Broadly speaking, real internal economies are of the following four types: (i) production economies, (ii) selling or marketing economies, (iii) managerial economies, and (iv) economies in transport and storage.

1) **Production economies:** When the scale of production expands, a number of economies accrue to the firm in the production process itself. First, opportunities for obtaining various types of economies emerge in the workshop of the factory. Production on a large scale enables the firm to carry out extensive division of labour and employ large automatic machines. The capacity of the machines is also fully used on account of the large volume of production. Instead of depending on others for carrying out repairs of machines and machine tools, the firm can itself employ technicians and workmen for the purpose. Techniques of production are changing so rapidly in the modern world that every producer has to remain ever alert. Large size of the firm and large-scale of operations is distinctly better in this regard since a larger firm can easily make use of its big financial resources to conduct research in its laboratories and/or adapt technology discovered elsewhere to suit its own requirements. The firm is, thus, able to discover better and less expensive techniques of production.

Whatever be the scale of operations, some waste material is invariably left out in each factory. If the scale of operations is small, a relatively larger quantity of this material goes unutilised. However, if the scale of operations is large, some useful goods can be prepared even from the waste material. For example, from the syrup left out in the sugar factory, liquor can be prepared. Similarly, in bangles factory, a number of small glass goods can be prepared out of the broken bangles.

When the scale of production is small, the producer generally cannot afford a packaging department. Therefore, he has to depend on others for obtaining packaging material like boxes, labels, etc. This leads to a substantial expenditure on packaging. However, if the scale of production is large, the firm can setup its own packaging department which is economical and also leads to lower per unit packaging costs.

2) **Selling or marketing economies:** Every producer produces with the purpose of selling. Therefore, he has to incur some expenditure in making his goods available to the consumer. When the scale of production is large, the per unit expenditure of the producer on marketing of goods is reduced substantially due to a number of reasons. All firms advertise their products in a number of ways. Even very small firms have to spend a certain minimum amount on advertising, though this expenditure of the small firms is considerably less than the expenditure of the large firms, yet the per unit cost of the large firms is smaller due to the fact that advertising cost is not required to be increased proportionately as the volume of production increases. Also, when the scale of production is large, the firm can economise on the expenditure on salesmen, agents, etc. The large firms can also enter into such contracts with the wholesalers and distributors that they take more interest in selling the products of the firm. Naturally, a small firm is deprived of this benefit.

3) **Managerial economies:** Managerial costs are partially production costs and partially selling costs. However, they are generally considered separately since it is convenient to do so. Managerial economies are obtained on account of the following two basic reasons: First, benefits of specialisation in the field of management can be obtained only when the scale of operations is considerably large. When the scale of production is small, all managerial responsibilities regarding production, marketing, finance, etc. will have to be borne by one person only. However, as the scale of operations expands, separate managers are appointed to look after these tasks. This raises the level and quality of management. At the same time, cost does not increase in proportion to the increase in the scale of operations. Large firms are in a position to use a number of machines for purposes of management. The use of computers, telephone, fax, etc. can be made only by a sufficiently large firm. If small firms use these machines, the total costs incurred on them would be very much higher in relation to the level of production attained.

The economists are, however, not in complete agreement on the managerial economies. Some economists argue that with the expansion of scale, managerial economies are obtained only upto a limit. After this limit, costs on management increase in a greater proportion. This is due to two reasons. First, the managerial structure in large companies is bureaucratic and when the scale of production expands, delays in decision making creep in. This weakens managerial efficiency. Second, the degree of uncertainty increases as the size of the firm increases. On account of this reason, various difficulties have to be encountered in decision making leading to an increase in managerial costs.

4) **Economies in transport and storage:** When the scale of production expands, economies in transport and storage accrue to the firm. Small firms have usually to depend on public transport and therefore their per unit transport cost is higher. As the scale of operation expands, the firm can purchase its own truck, lorry, etc. This will reduce the per unit transport cost for the firm. If the scale of operation expands still further, the firm can go in for larger trucks and lorries. The railways also give siding facilities to large producers and this reduces their loading costs. In reality, the transport cost is partly production cost and partly selling and marketing cost. When the firm purchases raw material, the loading cost is a part of its production cost. On the other hand, when finished goods are transported to the market, it is a part of selling and marketing cost. However, for convenience in analysis, the economists prefer to treat transport costs separately.

Like transport costs, storage costs are also partly production costs and partly selling and marketing costs. For example, expenditure on storing the raw material is a production cost whereas expenditure on finished and semi-finished goods is a part of marketing costs. From the point of view of the size of the warehouse, an important thing to remember is that larger the size of warehouse, larger will be the economies accruing to the firm. The reason is that the cost of construction of the warehouse does not increase in the same proportion as the increase in the storage capacity of the warehouse.

9.3.1.2 Pecuniary Internal Economies of Scale

Some pure pecuniary economies accrue to a firm as its scale of operation expands. The more important ones are the following:

1) A large sized firm can ask the suppliers of raw materials to give specific concessions and discounts. No raw material supplier usually ignores such requests (or pressures) of the large firm.
2) Perfect competition generally does not prevail in the capital market. Since the large companies have greater goodwill in the capital market, they are in a position to obtain loans at lower rates of interest from the banks and financial institutions.
3) Transport companies are also willing to provide discounts and concessions if the cargo is substantially large. This enables the firm to obtain monetary economies in transport costs by expanding its scale of operations.
4) When production is large, the firm is required to spend a large amount on advertising as well. However, advertising on a large scale attracts discounts and concessions from the media in which the advertisements appear.

9.3.2 Internal Diseconomies of Scale

If the scale of production is continuously expanded, is it possible that after a certain point, increase in production is less than proportionate than increase in the factors of production? Many economists believe that such a situation can and does arise if production is pushed beyond the point of optimum scale. The reasons that they advance are as follows:

1) **Limitations on the availability of factors of production:** The factors of production are always available in limited supply at the place of production. When the scale of production is increased beyond a certain point, it no longer remains possible to meet the requirements of some factors from local sources and, accordingly, factors have to be transported from other regions. This is generally possible only at higher prices. Let us suppose that an engineering factory is to be set up at Rudrapur in the Terai region. When the scale of production is small, it would be possible to meet the demand for some materials from local sources. As the scale of production expands, it will become more and more difficult to get even the labour from local sources and after a certain point, workers will have to be attracted from other regions by offering them higher wages.

2) **Problems in management:** When the scale of production is very large, the task of management at the top level becomes increasingly more and more burdensome and some inefficiency is bound to creep in. At times, information vital for taking a decision does not reach the top managers of the company in time. This delay, in turn, leads to a delay in decision making and increases the per unit cost.

3) **Technical factors:** When the scale of production is expanded, per unit cost increases due to a number of technical reasons. The establishment cost of large and sophisticated plants and machinery is generally high. The buildings of large factories should also have stronger foundations and the factory itself must be equipped with coolers, air-conditioners, etc. All these factors lead to an increase in per unit cost.

9.3.3 External Economies

External economies were discussed first of all by Alfred Marshall. According to him, when a firm enters production, it obtains a number of economics for which the firm's own production strategy, managerial arrangements, etc. are not responsible. In fact, these are economies external to the firm. For example, let us suppose that a firm is established at a place where transport, advertising facilities, etc. are not available. If the size of the firm remains small, it is possible that these facilities are not locally available in the future as well. However, if the size of the firm increases significantly, these facilities will themselves start coming to the firm. These are, in fact, external economies.

When a firm expands its scale of production, other firms also earn many economies. For example, when a large factory attracts various factors of production fairly regularly, many other factories set up in the neighbourhood, that could not have attracted these factors on their own, also stand to gain. They obtain these factors at practically the same prices at which the large factory obtained them.

Because of external economies of large-scale production, there is a gap between private and social returns. When a firm expands its scale of production, it becomes possible for the other firms to reduce their cost of production. However, there is no method available in the prevalent price mechanism to the firm expanding its scale of operations to charge for the benefits it confers on the other firms.

9.3.4 External Diseconomies

When the scale of operations is expanded, many such diseconomies emerge that have no particular ill-effect on the firm itself. In fact, their burden falls on the other firms. On account of this reason, they are termed external diseconomies. The smoke rising from the chimney of a factory pollutes the atmosphere. When the firm is of a small size, the pollution is less and its ill-effects on the people living in colony nearby is limited. However, if the scale of the firm is large, the smoke will be very dense and can cause serious health hazard to the people. Similarly, as the scale of production of the factories increases, employment rises sharply. This creates problems of traffic congestion and overcrowding in the city where these factories are located. In agriculture, increase in the scale of production leads to problems of soil erosion and this reduces the fertility of the adjoining fields as well. From the above illustrations, it is clear that external economies and diseconomies can be both pecuniary and technological.

Check Your Progress 1

1) Indicate the following statements as true (T) or false (F):

   i) When output increases in a greater proportion as compared to the increase in the amount of the factors of productions, we have the stage of increasing returns to scale.

   ii) Those economies which accrue to a firm an account of the other firms are known as external economies.

   iii) Production economies are a part of pecuniary internal economies.

   iv) In the case of linear homogenous production function, we have constant returns to scale.

2) Discuss the factors which account for increasing returns to scale.

   ....

   ....

3) Explains how decreasing returns to scale arise.

   ....

   ....

4) Discuss the internal economies of scale.

   ....

   ....

5) What do you mean by external economies and external diseconomies?

   ....

   ....

9.4 LET US SUM UP

In this unit, the concept of returns to scale has been explained. As noted in the beginning itself, this concept is associated with the tendency of production that is observed when the ratio between the factors is kept constant but the scale is expanded. This, in turn, can give rise to three possibilities – increasing returns to scale, constant returns to scale, and diminishing returns to scale. After discussing all these possibilities, we shift our focus to a discussion of economies and diseconomies of scale. Economies of scale, in turn, are divided into two parts – internal economies of scale and external economies of scale. Economies which accrue to a firm on expansion of its own size are known as internal economies while economies which accrue to a firm not due to its own operations but due to operations of other firms are termed as external economies. We have discussed in detail all the causes which can result in the generation of such economies. In the end, we have focussed our attention on diseconomies of scale – both internal and external.

9.5 REFERENCES

1) Robert S. Pindyck, Daniel L Rubinfeld and Prem L. Mehta, *Microeconomics* (Pearson Education, Seventh edition, 2009), Chapter 5, Secton 5.4.

2) Dominick Salvotore, *Principles of Microeconomics* (Oxford University Press, Fifth edition, 2010) Chapter 7, Section 7.5.

3) A. Koutsoyianms, *Modern Microeconomics* (The Macmillan Press Ltd, Second edition, 1982) Chapter 3.

4) John P. Gould and Edward P Lazear, *Microeconomic Theory* (All India Traveller Bookseller, Sixth edition, 1996), Chapter 8, Section 8.6.

9.6 ANSWERS OR HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) (i) (T), (ii) (T), (iii) (F), (iv) (T)

2) See Sub-section 9.2.1 of Section 9.2

3) See Sub-section 9.2.3 of Section 9.2

4) See Sub-section 9.3.1 of Section 9.3

5) See Sub-section 9.3.3 and 9.3.4 of Section 9.3.

UNIT 10 THE COST OF PORDUCTION

Structure

10.0 OBJECTIVES

After going through this unit, you should be able to:

- state the various concepts of costs like private cost, social cost, money cost, sunk cost, economic cost, accounting cost etc.;

*Dr. V.K. Puri, Associate Professor of Economics, Shyam Lal College (University of Delhi) Delhi.

- differentiate between short-run and long-run cost functions;
- know the difference between fixed cost and variable cost and the nature of total cost curve;
- explain the concept of average fixed cost, average variable cost, average total cost and marginal cost and nature of these curves;
- discuss the relationship between marginal cost curve and average cost curve;
- appreciate the difference between short-run and long-run cost curves; and
- describe the relationship between long-run marginal cost and short marginal cost.

10.1 INTRODUCTION

The decision of a firm regarding production of a good depends on two factors: First, the demand for the good, and second, the cost of production of the good. Accordingly, the concept of cost of production is basic to the understanding of the price theory and requires a thorough discussion. A price taker firm wishes to maximise its profits will be able to do so if it is able to minimise its costs. Obviously a firm is interested in minimising what economists call the *private cost*. The concept of *social cost* that is being often referred to in the context of social welfare is not relevant for the theory of firm. However, it is necessary to understand the distinction between the concepts of the private cost and the social cost. In economic analysis, we often distinguish between *money cost* and the *opportunity cost*. From analytical point of view both the concepts are relevant and thus must be understood carefully. The concept of money cost may be interpreted from the point of view of an accountant or an economist. The two approaches differ on the treatment of implicit costs.

After settling these conceptual issues in the theory of costs, one has to analyse the nature of costs in both the short-run and the long-run. In the short-run since we have some fixed inputs and some other inputs are variable, one has to draw the distinction between the *fixed costs* and the *variable costs*. However, in the long-run because the amounts of all the inputs can be varied, all costs are considered together. Finally, the theory of costs attempts to explain as to how cost changes occur in response to changes in the size of production. In the last two units we have discussed the theory of production at some length. This discussion should help us to understand that the cost changes depend largely on how changes in production take place as a result of changes in the amounts of inputs.

10.2 THE CONCEPTS OF COSTS

10.2.1 Private Costs and Social Costs

In microeconomic theory, the concepts of both private cost and social cost are used. The firm, in its attempt to attain the goal of profit maximisation, is guided entirely by the private cost considerations. In its decision making, it ignores all those costs which it may be imposing on others while carrying out its production programme. However, in welfare studies, together with the firm's both explicit and implicit costs, all such costs are taken into account which are external to the 'narrow economy' of the firm.

Private costs: Every firm requires various inputs to produce a good. In order to secure a command over these inputs, the firm has to pay some price for each of these inputs. In common parlance, the amount of money so paid is known as cost. Economists, however, include in the private cost not only the expenditure incurred by the producer on purchasing (or hiring) of factors of production (or inputs) from the market, but also the imputed cost of all those services which the producer himself provides. The private cost of production of any output may thus be defined as either the purchase or the imputed value of all productive services used in producing the output and is equivalent to the total monetary sacrifice of the firm made to secure it.

Generally, economists include the following expenditures in the cost: (i) cost of the raw materials, (ii) wages of the labourers, (iii) interest payments on capital loans, (iv) rent of the land and the buildings, (v) repairing costs of machines and depreciation, (vi) tax payments to the government and local bodies, (vii) imputed wage payment to the producer for the work performed by him, (viii) imputed interest payment for the capital invested by the producer himself, (ix) rent of land and buildings owned by the producer himself and (x) normal profits of the firm.

This shows that three types of expenditures are included in the private cost: (i) the purchase price of the factors of production employed in the production process, (ii) imputed price of the resources provided by the producer himself, and (iii) normal profits.

Social costs: Social costs differ from private costs on account of two reasons:

First, externalities are not included in private costs. For example, a factory located in the residential area by polluting the atmosphere will expose the residents of the colony to various ailments and will thereby raise their medical expenditures. Though these costs are quite relevant from the point of view of the society, they will never be considered by the firm as part of its costs.

Secondly, market prices of goods may not reflect their social value and thus there may be divergence between private and social costs. The imposition of government taxes, subsidies, and controls of various kinds distort free market prices. Further, prices of factors of production may overstate or understate the opportunity cost of using those factors. In heavily populated countries where widespread disguised unemployment is to be found in the agricultural sector, the industrial wage often exceeds the opportunity cost of the labour which is drawn from the agricultural sector. In computing the social costs, adjusted market prices for goods and factors of production are used. While the adjusted prices for factors of production are called shadow prices, the adjusted prices for goods are termed as social prices.

10.2.2 Money Cost: Explicit and Implicit Costs

The concept of the money cost in contrast to the concept of opportunity cost is simple.

The money cost of production of any output is considered to be equivalent to the total monetary sacrifice made to obtain that output.

Thus, costs are not sacrificed alternatives but monetary payments. This conception of money cost is rather narrow and is used for accounting purposes. From the point of view of the economists, this concept of cost is not very relevant. Since economists wish to study as to how costs affect output choices, employment decisions, and the like, costs should include imputed value of all the inputs provided by the producer himself in addition to outright money expenses. Hence, costs can be classified as explicit costs and implicit costs. Explicit costs arise from transactions between the firm and other parties in which the former purchases inputs or services of inputs for carrying out the production. These costs are usually the costs shown in the accounting statements and include wage payments, raw materials costs, interest on loans, payments for insurance, electricity and so on. Implicit costs are the costs associated with the use of the firm's own resources. Since these resources will bring return if employed elsewhere, their imputed values constitute the implicit costs. Implicit costs are however difficult to measure. Economists nonetheless assert that they must be taken into account in analysing the activities of a firm.

10.2.3 Real Costs

The concept of real cost was developed by Alfred Marshall. In his opinion, a worker suffers discomfort while he renders his services for productive purposes. Similarly, a person makes some sacrifice when he saves his income and lends it to investors who use it for carrying out production. These discomforts and sacrifices are in the nature of real costs of production. In Marshall's own words, "The exertions of all the different kinds of labour that are directly or indirectly involved in making it; together with the abstinences or rather the waitings required, for saving the capital used in making it; all these efforts and sacrifices together will be called the real costs of the production of the commodity."

The concept of real cost is, however, based on subjectivity and cannot be used for precise measurement of production cost. It is this reason why modern economists do not consider it to be of much relevance in the price theory. They admit that most of the labour involves hard work and is definitely unpleasant. It, therefore, has a heavy real cost. In contrast, the real cost of simple and less arduous work is generally low. But this fact is not at all relevant from the point of view of price determination in a free enterprise economy. Moreover, to modern economists, savings do not involve any sacrifice. Hence, these economists regard the concept of real cost as inappropriate.

10.2.4 Sunk Cost and Incremental Cost

In economics and business decision-making, a sunk cost is a cost that has already been incurred and cannot be recovered. Sunk costs (also known as retrospective costs) are sometimes contrasted with prospective costs, which are future costs that may be incurred or changed if an action is taken. In traditional microeconomic theory, only prospective (future) costs are relevant for decision making. Since sunk costs have already been incurred and cannot be recovered, therefore they should not influence the rational decision-maker's choices.

An incremental cost is the increase in total costs resulting from an increase in production or other activity. For instance, if a company's total costs increase from Rs. 5.6 lakh to Rs. 6.0 lakh as a result of increasing its machine hours from 7,000 to 8,000, the incremental cost of the 1,000 machine hours is Rs. 40,000.

10.2.5 Economic Cost and Accounting Cost

Economists and accountants view costs from different angles. Accountants are concerned with the firm's financial statement and tend to take a retrospective look at the firm's finances because they have to keep track of assets and liabilities and evaluate past performance. Accounting cost includes depreciation expenses for capital equipment at rates allowed by the tax authorities.

Economists, on the other hand, are concerned with what cost is expected to be in the future, and with how the firm might be able to rearrange its resources to lower its cost and improve its profitability. Thus, they take a forward looking view and must therefore be concerned with opportunity costs.

As stated earlier, there is a difference regarding the treatment of explicit and implicit costs as well. Both, the economists and the accountants consider explicit costs (like payment of wages and salaries, cost of raw material, property rentals, etc.) because these involve direct payments by a company to other firms and individuals that it does business with. However, while economists also take into account the implicit costs, accountants ignore them. For example, consider the owner of a retail store who manages his own retail store but does not pay any salary to himself. Since no monetary transaction has taken place, accountant will not include it in the accounting cost. However, the economist will include this implicit cost in total cost as the retail store owner could have earned a competitive salary by working elsewhere (that is, the implicit cost of the owner will be his opportunity cost).

The treatment of depreciation is also different. When estimating the future profitability of a business, an economist is concerned with the capital cost of plant and machinery. This involves not only the explicit cost of buying and the running of the machinery, but also the cost associated with wear and tear. On the other hand, accountants use depreciation rates on different assets as allowed under the tax laws in their cost and profit calculations. These depreciation rates need not reflect the actual wear and tear of the equipment, which is likely to vary asset by asset.

The above discussion shows that there are some important differences in the methods of calculating costs as used by the economists and the accountants. Accordingly, the calculation of profit will also differ. To illustrate, consider a retail store owner who has invested Rs. 1,00,000 as equity in a store and inventory. His monthly sales revenue is Rs. 2,60,000. After deduction of cost of goods sold, salaries of hired labour, and depreciation of equipment and buildings, the accounting profit to the store owner is Rs. 25,000 (see Table 10.1).

Table 10.1: Accounting income statement for the Retail-Store Owner

Sales		Rs. 2,60,000
Cost of goods sold	Rs. 1,80,000	
Salaries	30,000	
Depreciation expense	25,000	Rs. 2,35,000
Accounting profit		Rs. 25,000

In Table 10.2 we consider the economic statement of profit of the same store. The cost of goods sold and salaries remain the same. Let us suppose that the market values of the equipment and building in fact declined by Rs. 25,000 over the current year and that the depreciation charge, therefore, reflects the opportunity costs of these resources. Thus, depreciation expense is taken to be Rs. 25,000 as in Table 10.1. However, the economist will add two items relating to the implicit cost in the cost of production. Suppose that the owner-manager could earn Rs. 25,000 per month as a departmental manager in a large store and that this is his best opportunity for salary. Then we would add Rs. 25,000 as the imputed salary of the owner-manager to the cost of production. Similarly, the owner-manager has Rs.1,00,000 equity in the store and inventory – a sum he could have easily invested elsewhere. Let us suppose that he could have earned 10 per cent interest on this amount had he invested it elsewhere. Thus, imputed interest cost on equity will be Rs. 10,000. Thus, as can be seen from Table 10.2, the total economic costs, or the opportunity costs of all resources used in the production process will add up to Rs. 2,70,000. This implies an economic loss of Rs.11,000 to the owner-manager of the store against the accounting profit of Rs. 25,000 depicted in Table 10.1.

Table 10.2: Economic statement of profit to the Retail-Store Owner

	Rs.	Rs.
Sale		2,60,000
Cost of goods sold	1, 80,000	
Salaries	30,000	
Depreciation expense	25,000	
Imputed salary to owner-manager	25,000	
Imputed interest cost on equity	10,000	2,70,000
Economic Profit		-10,0000

In addition to the above differences in the calculation of profits by the economists and the accountants, it is also important to point out that while for economists, profits and losses are the driving force, business accounting does not stop here. Business accounts also include the balance sheet, which is a picture of financial conditions on a particular date. This statement records what a firm is worth at a given point of time. On one side of the balance sheet are recorded the 'assets' and on the other side are recorded the 'liabilities' and 'net worth'. A balance sheet must always balance because net worth is a residual defined as assets minus liabilities.

The business accounting concepts can be summarised as follows:

1) **The income statement shows the flow of sales, cost, and revenue over the year or accounting period. It measures the flow of money into and out of the firm over a specified period of time.**

2) **The balance sheet indicates an instantaneous financial picture or snapshot. It is like a measure of the stock of water in a lake. The major items are assets, liabilities and net worth**.

10.2.6 Historical Cost and Replacement Cost

The historical cost is the cost that was actually incurred at the time of the purchase of an asset. As against this, replacement cost is the cost that will have to be incurred now to replace that asset (i.e., replacement cost is the current cost of the new asset of the same type).

These two costs differ because of changes in prices over a period of time. Naturally, if prices remain unchanged over time, both the costs will be the same. But this seldom happens. Accordingly, historical cost and replacement cost of an asset always differ. If the price rises over a period of time, replacement cost will be higher than the historical cost. On the other hand, if the price of the asset declines over a period of time, replacement cost will be lower than the historical cost.

Because of the requirements of tax laws and the laws governing financial reporting to shareholders, accountants generally express many costs in terms of the actual or historic costs paid for the resources used in the production process in accordance with the convention of financial accounts. However, both economists and accountants agree on the fact that for decision making purposes, it is not the historical cost that is relevant but the replacement cost. This is due to the reason that for all decision making purposes, it is the 'current' (or the replacement) cost that is important and not the cost that was incurred some years earlier at the time of the purchase of the asset.

Check Your Progress 1

1) Indicate the following statements as true (T) or false (F):

   i) Externalities are not a part of private cost ( )

   ii) Implicit costs are the costs associated with the use of firm's own resource ( )

   iii) Retrospective costs are relevant for decision making ( )

   iv) Accountants tend to take a retrospective look at the firm's finances ( )

   v) Economists are concerned with opportunity costs ( )

   vi) The historical cost is the current cost of the new asset of the same type ( )

2) Explain the difference between explicit cost and implicit cost.

   ....

   ....

   ....

3) Distinguish between private cost and social cost.

   ....

   ....

   ....

4) What is the difference between sunk cost and incremental cost?

....

....

....

5) Explain the difference between economic cost and accounting cost.

....

....

....

10.3 COST FUNCTIONS: SHORT-RUN AND LONG-RUN

The relationship between product and costs is known as the cost function.

There are two elements in determining the cost function of a firm. First, the production of the firm, and second, the prices paid by the firm for the factors used.

In practice, production functions can be of various types. At times, one factor of production is variable and other factors fixed. It is also possible for some factors to be variable. On account of this reason, cost function can also be of various types. In economics, generally two types of cost functions are considered under the price theory:

i) The short-run cost function, and

ii) The long-run cost function.

Cost functions can be illustrated in diagrammatic forms as cost curves.

10.3.1 Cost Function and the Time Element

To understand the theory of cost, it is necessary to be clear about the meaning of short-run and long-run. In common usage, these terms may be used for weeks, months and years but for the economist they indicate conditions of production and have no reference to the calendar year. Even then, the concept of time does creep in indirectly when the terms short-run and long-run are discussed.

Generally, economists regard that period of time as short-run in which some factors of production are fixed (at least one factor is fixed) and the firm depends only on the variable factors of production to increase the level of output. If the firm does not employ the variable factors at all, the output will be zero in the short-run. However, the maximum quantity of output that can be produced depends upon the quantity of the fixed factors of production. In the long-run, all factors are variable and the quantity of the output can be increased to any limit. For example, in a manufacturing industry the plants, machinery, building of the factory, etc. are fixed resources in the short-run while the raw materials, labour, power, etc. are variable. Therefore, to increase the amount of output in this period, it will become necessary to employ more units of the variable resources in conjunction with the fixed resources. Obviously, the maximum output that can be obtained in this period will depend to a great extent upon the total quantity of the fixed resources of production.

10.3.2 Long-Run Cost Function

In the long-run, total cost is a multivariable function which implies that total cost is determined by many factors. The long-run cost function may be written as

$$C = f(Q, T, P_f)$$

Where, C = total cost of production

Q = output

T = technology

P_f = prices of the relevant factors of production.

Graphically, the long-run cost function is shown on two dimensional diagram as C=f(Q), ceteris paribus. With the assumption that the technology and the prices of relevant factors of production remain constant, the long-run cost function may be written as

$$C = f(Q, \overline{T}, \overline{P}_f \text{ or } C = f(Q)$$

However, the technology and the factor prices need not remain constant. When these factors change, their effect on cost is shown by a shift of the cost curve. It is this reason why the factors other than output are known as shift factors. Theoretically there is no difference between the various factors which determine the costs, and the distinction we have drawn above between the output level and other factors determining costs can sometimes be misleading. However, for showing costs on two dimensional diagrams this distinction has to be made.

10.3.3 Short-Run Cost Function

In the short-run, in addition to output level, technology and factor prices, the fixed factors such as capital equipment, land, etc. also determine costs of production. Therefore, the short-run cost function is written as

$$C = f(Q, \overline{T}, \overline{P}_f, \overline{K})$$

Where, $\overline{K}$ indicates fixed factors. In the discussion on the production function, it has been stated that in the short-run certain factors like capital equipment, land, factory building and top managerial staff remain constant. $\overline{K}$ underlines the fact of the constancy of the fixed factors. Since the amount of fixed factors does not change in the short-run under any circumstances, $\overline{K}$ is not a shift factor like technology and factor prices.

10.4 THEORY OF COST IN THE SHORT-RUN

The short-run costs of a firm are divided into fixed and variable costs. Therefore,

$$TC = TFC + TVC$$

where, TC = total cost

TFC = total fixed cost

TVC = total variable cost

10.4.1 Fixed Cost

Fixed cost is also known as supplementary cost. While engaging in productive activity, the producer always has to incur some expenditure which remains fixed whatever the level of production, so much so that even if the producer stops production altogether, these costs have to be incurred.

This is known as fixed cost of production. Interest paid by the producer on the capital borrowed for purchasing plant and machinery, rent of the factory building, depreciation of the machinery, the wages of foremen and organisers, etc. are all fixed costs. These costs remain fixed even when the level of output is varied. Even if the producer decides to close down production, he has to bear these costs since the factory rent, wages of managers, interest on capital, etc. have to be paid. This discussion makes it clear that larger the level of production in a firm, the lower will be the per unit fixed cost (or average fixed cost).

10.4.2 Variable Cost

The cost which keeps on changing with the changes in the quantity of output produced is known as variable cost.

For instance, raw material has to be used in the process of production in a manufacturing industry, labour has to be employed for running machines, and energy (electricity) has to be arranged. Generally expenditure on these inputs increases or decreases due to changes in the level of production. It is important to remember in this context that when the producer abandons production in the short run, these costs also vanish completely. In fact, it is due to this direct relationship between expenditure on such inputs and the level of production that these expenditures are known as variable costs.

The concepts of total cost, total variable cost and total fixed cost in the short-run can be easily followed with the help of Table 10.3.

Table 10.3 : Short-Run Costs of a Hypothetical Firm

Output (Unit)	Total Fixed Cost (Rupees)	Total Variable Cost (Rupees)	Total Cost (Rupees)
0	240	0	240
1	240	120	360
2	240	160	400
3	240	180	420
4	240	212	452
5	240	280	520
6	240	420	660

10.4.3 Total Fixed Cost

Total fixed cost is the total expenditure by the firm on fixed inputs.

From Table 10.3, it is clear that the total fixed cost of the firm remains constant at Rs. 240 irrespective of the level of output. In our illustration, output varies from 1 unit to 6 units, but the total fixed cost remains 240 in each case. Even when the firm stops production altogether, implying that output is at zero level, the total fixed cost remains unchanged. The firm's total fixed cost function is shown in Fig. 10.1.

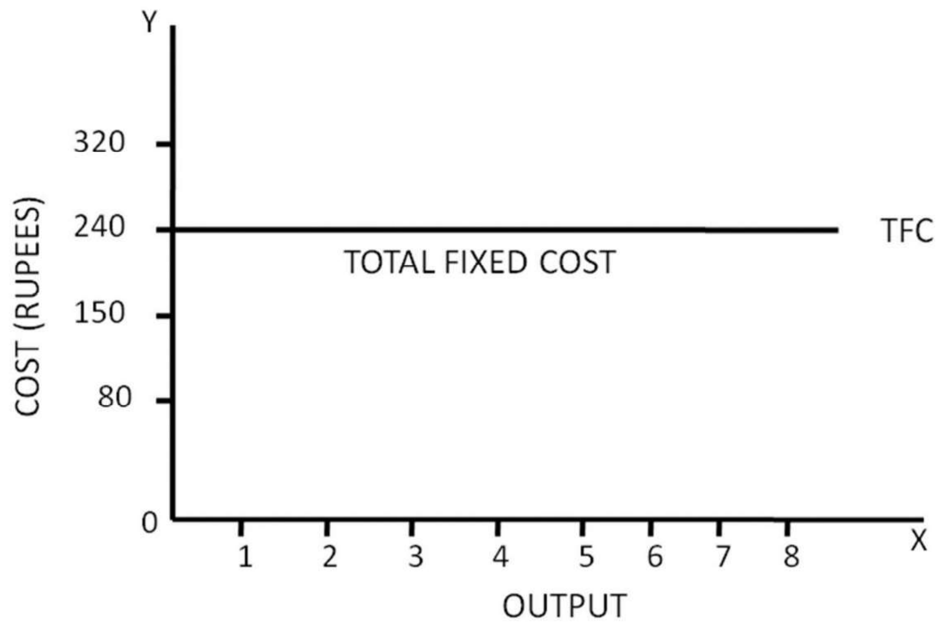

Fig. 10.1 : Total Fixed Cost curve is parallel to X axis as total fixed cost remains the same for all levels of output

10.4.4 Total Variable Cost

Total variable cost is firm's total expenditure on variable inputs used to carry out production.

Since higher output levels require greater utilisation of variable inputs, they mean higher total variable cost. Table 10.3 shows that the total variable cost of the firm increases as its output increases. However, when the firm stops its production altogether, it does not require any variable input and, therefore, its total variable cost is zero. Fig. 10.2 shows the firm's total variable cost function. Notice one peculiar feature of TVC – initially it rises sharply, then, there is a moderation in its rate of rise and ultimately it resumes rising at a faster pace.

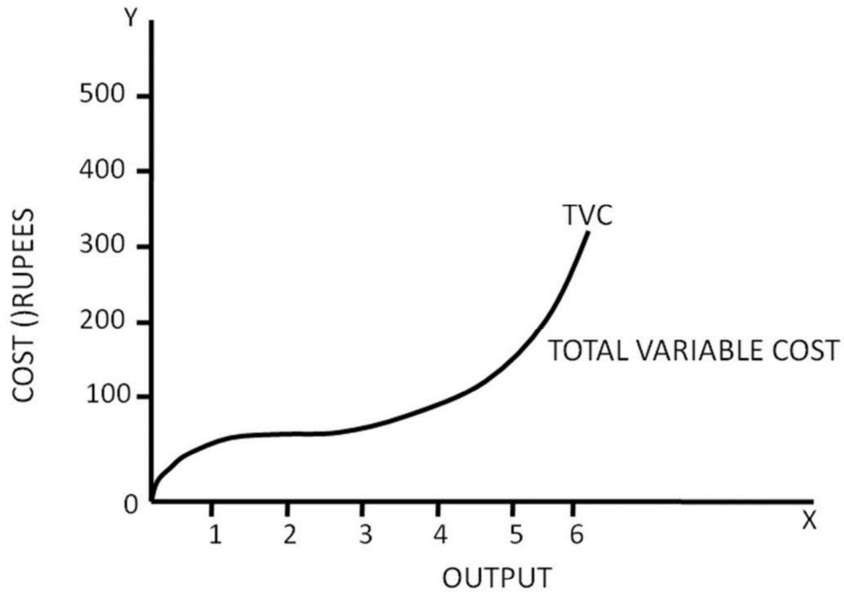

Fig. 10.2 : Total Variable Cost Curve rises from left to right

10.4.5 Total Cost

Total cost is the sum of total fixed cost and total variable cost.

Thus, to obtain the firm's total cost at a given output, we have only to add its total fixed cost and its total variable cost at that output. The result is shown in Table 10.3 and the total cost function is shown in Fig. 10.3. Since the total cost function and the total variable cost function differ by only the amount of total fixed cost which is constant, they have the same shape.

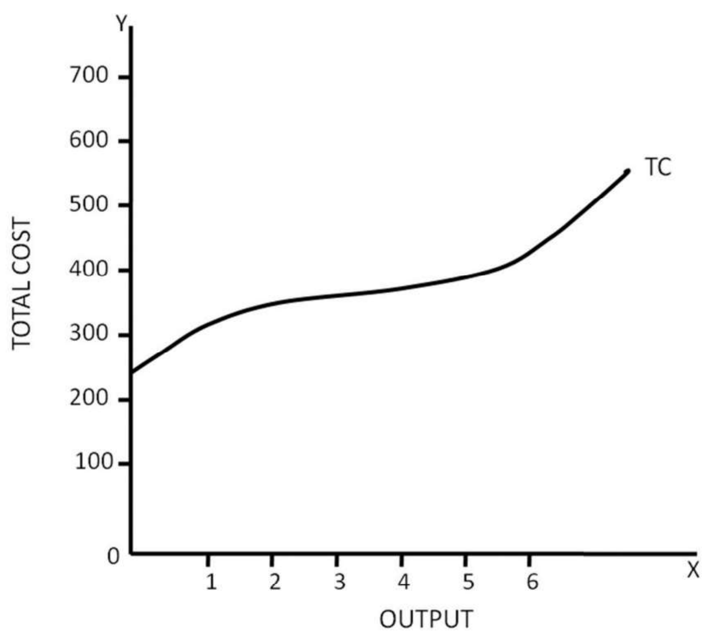

Fig. 10.3: Total Cost curve is obtained by adding the total fixed cost to total variable cost

In Fig. 10.4, all the three cost functions discussed above (total fixed cost function, total variable cost function and total cost function) have been shown together. Cost functions, when depicted graphically, are often called cost curves.

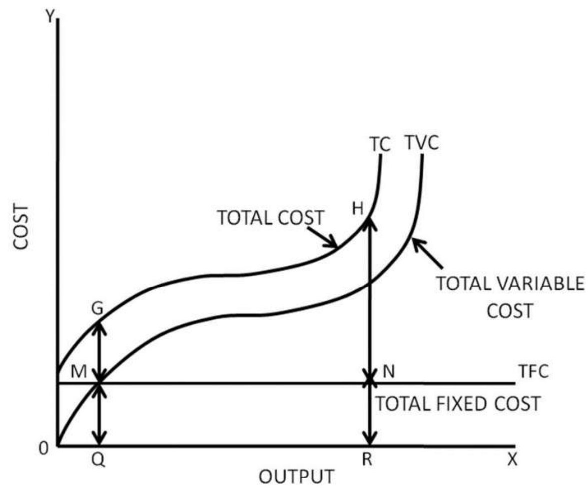

Fig. 10.4 : Total Fixed Cost, Total Variable Cost and Total Cost

In Fig. 10.4, TFC is the total fixed cost curve. Since it is parallel to X-axis, it indicates that whatever be the level of output the total fixed cost remains the same (i.e., it does not change in response to a change in the level of production). TC is total cost curve. It indicates the sum of total fixed cost and total variable cost for the various output levels. If the level of production is to be raised, the use of variable inputs will have to be increased and this will push up the costs. The rising total cost curve TC from left to right (the positive slope of TC curve) indicates this fact. The vertical distance between the total cost curve TC and the total fixed cost curve TFC indicates total variable cost. For example, if the firm wishes to produce OQ units of output, the total variable cost will be GQ – MQ = GM and if the level of output is OR, the total variable cost will be HR – NR = HN. The total variable cost has been depicted by the curve TVC in Fig. 10.4. This is parallel to the total cost curve TC and the vertical distance between the two curves (TC and TVC) indicates total fixed cost.

Check Your Progress 2

1) Indicate the following statements as true (T) or false (F):

   i) Cost function explains the relationship between product and costs ( )

   ii) In the long run all factors are variable ( )

   iii) Fixed cost is also known as supplementary cost ( )

   iv) Total variable cost is the total expenditure by the firm for fixed input ( )

2) Define and distinguish between long run cost function and short run cost function.

   ....

   ....

   ....

3) Distinguish between fixed cost and variable cost.

....

....

....

4) Define total fixed cost and total variable cost and trace the nature of the total cost curve.

....

....

....

10.5 SHORT-RUN COST CURVES

To find out the per unit profit, the firm has to compare the per unit cost (or average cost) with per unit price. Therefore, it is necessary for us to understand the concepts of average fixed cost, average variable cost and average total cost.

10.5.1 Average Fixed Cost

Generally, all those firms whose total costs of production include a significant proportion of fixed costs try to increase the level of production to such an extent that per unit fixed cost which is often known as average fixed cost, is reduced substantially. To find out the average fixed cost, total fixed cost has to be divided by the output.

In the form of a formula,

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

where, AFC is the average fixed cost

TFC is the total fixed cost

Q is the output

Table 10.4: Average Fixed Cost, Average Variable Cost and Average Total Cost of the Firm

Output (Units)	Average Fixed Cost TFC ÷ Q	Average Variable Cost TVC ÷ Q	Average Total Cost TC ÷ Q
1	240 ÷ 1=240	120 ÷ 1=120	360 ÷ 1=360
2	240 ÷ 2=120	160 ÷ 2=80	400 ÷ 2=200
3	240 ÷ 3=80	180 ÷ 3=60	420 ÷ 3=140
4	240 ÷ 4=60	212 ÷ 4=53	452 ÷ 4=113
5	240 ÷ 5=48	280 ÷ 5=56	520 ÷ 5=104
6	240 ÷ 6=40	420 ÷ 6=70	660 ÷ 6=110

A mere look at Table 10.4 will show how the average fixed cost declines with a rise in the level of output. When the firm produces only 1 unit, average fixed cost is Rs. 240. As the ouput is expanded, there is a sharp decline in average fixed cost and it is as low as Rs. 40 when 6 units of the commodity are produced.

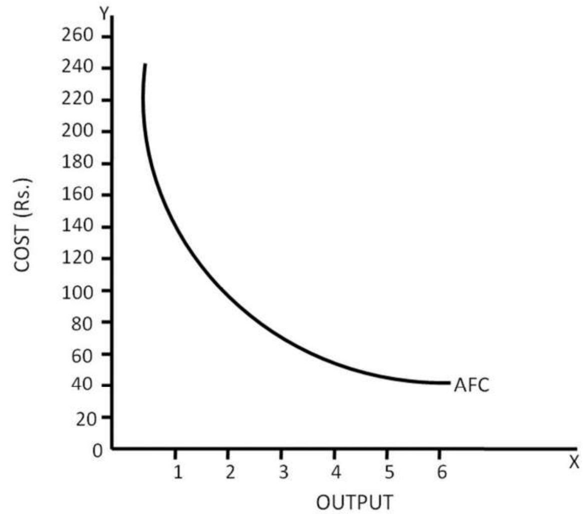

Fig. 10.5: Average Fixed Cost curve is a rectangular hyperbole

The fact that average fixed cost must decline with increases in output can be easily understood with the help of average fixed cost curve in Fig. 10.5. In this figure, when output is 1 unit, the average fixed cost is Rs. 240. When the output is increased to 3 units and then to 6 units, average fixed cost declines first to Rs. 80 and then to Rs. 40.

The average fixed cost curve (AFC) is a rectangular hyperbole because multiplication of average fixed cost with the quantity of output produced always yields a fixed value (the area under the curve is always same and is equal to the total fixed cost).

10.5.2 Average Variable Cost

To obtain the average variable cost, we divide the total variable cost by the output. In the form of formula:

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

where, AVC = the average variable cost

TVC = the total variable cost

and Q = the output.

In fact, the average variable cost curve (AVC) gives us the same information in money terms that we obtain from the average product curve of the variable factor in physical terms.

With an increase in the amount of variable factor, the efficiency in production increases (resulting in an increase in average product) and the average variable cost declines. If average productivity remains constant, average variable cost will also remain constant. If it declines, average variable cost increases.

Thus, average variable cost curve is the reciprocal of the average variable (factor) product curve.

After having understood the relationship between average variable factor productivity and average cost, it is easy to understand the nature of the AVC curve. While discussing the laws of production, we had stated that if other factors are kept constant and only the quantity of one factor is increased, then initially the tendency of increasing returns is observed. Later on, it is followed by constant returns and diminishing returns in that order. This means that in the initial stages, average variable cost declines and, after reaching a minimum point, starts increasing. This increase is due to the operation of the law of diminishing returns. From Table 10.4 we learn that at the output level of 1 unit the firm's average variable cost is Rs. 120. It declines when output is increased and is Rs. 53 when 4 units of the commodity are produced. Thereafter, it increases and is Rs. 70 when output level is raised to 6 units. The average variable cost curve is thus U-shaped as in Fig. 10.6.

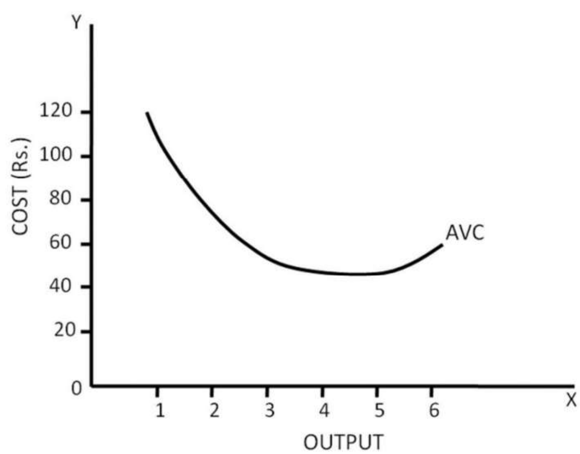

Fig. 10.6: Average variable cost curve is a U-shaped curve

10.5.3 Average Total Cost

The average total cost is also known as average cost. To find out average cost, we divide total cost (which is the sum of total fixed cost and total variable cost) by the output. In the form of a formula:

$$\text{AC or ATC} = \frac{\textbf{TC}}{\textbf{Q}} = \frac{\textbf{TFC}}{\textbf{Q}} + \frac{\textbf{TVF}}{\textbf{Q}}$$

The modern economists are generally agreed that in all areas of economic activity, average total cost declines initially. The reasons are the same which lead to increasing returns in the initial stages. Average cost declines initially because some of the resources are indivisible and there are possibilities of specialisation in the production process. As long as the indivisible factors are not fully utilised, the average total cost falls and when expansion in output leads to a stage where the indivisible resources are fully utilised, an optimum proportion is established between the factors of production. Output obtained at this point is the optimum output. Here, the average total cost is minimum. If the output is expanded beyond this point (which denotes an optimum combination of resources) by increasing the amount of variable inputs, then total production increases at a diminishing rate. This leads to a rise in average total cost. This shows why the average total cost curve is U-shaped as shown in Fig. 10.7. The illustration given in Table 10.4 also makes this point abundantly clear.

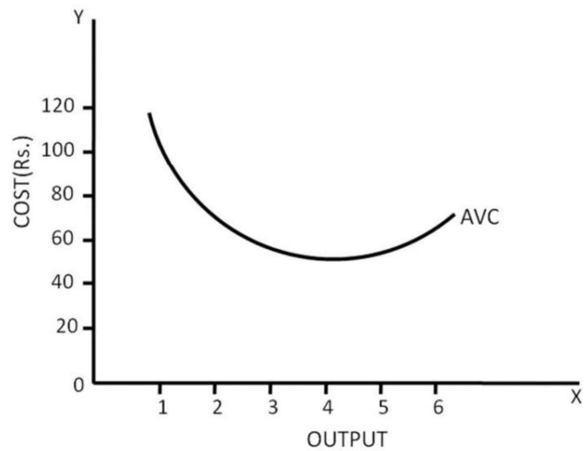

Fig. 10.7: Average Total Cost curve is obtained by dividing total cost by the output

We can understand the shape of average total cost curve ATC better with the help of average variable cost curve AVC and average fixed cost curve AFC drawn in Fig. 10.8. Since the ATC curve is obtained by vertically summing up the AVC and AFC curves, when both AVC and AFC curves slope downward, the ATC curve also slopes downwards. The point R on the AVC curve shows the minimum average variable cost. After this point, the average variable cost starts increasing and thus the AVC curve is sloping upward. However, the fall in the average fixed cost more than compensates for the rise in average variable cost. Hence, the ATC curve slopes downward. Since at point T on the AVC curve the rate of increase of the average variable cost is the same as the rate at which the average fixed cost falls corresponding to this level of output, average total cost is minimum at this output level. As the level of output increases beyond this point, the average variable cost rises far more rapidly than the rate at which average fixed cost falls. Therefore, the ATC curve slopes upward.

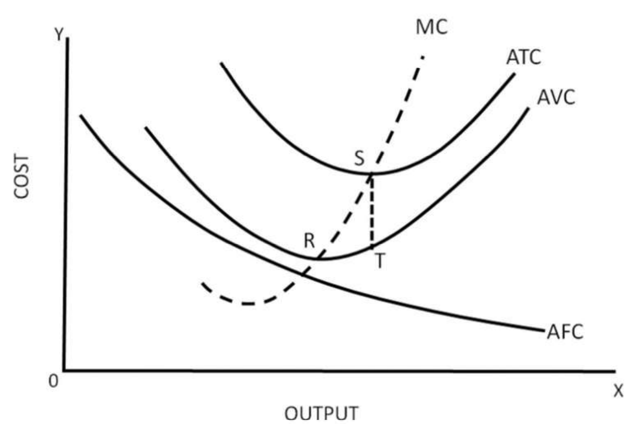

Fig. 10.8: Average total cost is the vertical sum of AFC and AVC

10.5.4 Marginal Cost

The marginal cost is the increase in the total cost owing to a small increase in output.

In symbols,

$$MC = \frac{\Delta TC}{\Delta Q} \text{ or } \frac{\Delta TVC}{\Delta Q}$$

where, MC is marginal cost

ΔTC is change in total cost associated with a small change in output

ΔTVC is change in total variable cost associated with a small change in output

ΔQ is small change in output

The concept of marginal cost can be understood with the help of an example. In Table 10.5, the total cost of producing 2 units of output is Rs. 400 and the total cost of producing 2 + l or 3 units of output is Rs. 420. Therefore, marginal cost is Rs. 20 which is Rs. 420 – Rs. 400.

Table 10.5: Calculation of Marginal Cost

Output Units	Total Cost (Rs.)	Total Variable Cost (Rs.)	Marginal Cost (Rs.)
0	240	0	-
1	360	120	120
2	400	160	40
3	420	180	20
4	452	212	32
5	520	280	68
6	660	420	140

Since fixed cost remains unchanged in the short run, marginal cost can also be defined as the increase in total variable cost consequent upon a small increase in output. From Table 10.5, we learn that the variable cost of producing 2 units is Rs. 160 and that of 3 units Rs. 180. The marginal cost, thus, will be Rs. 180 – Rs. 160 = Rs. 20.

The marginal cost (MC) curve as it would be clear from Fig. 10.9 is U-shaped. This implies that the marginal cost curve MC first slopes downward and then at the point where marginal cost is minimum, it starts sloping upward because marginal cost after decreasing with increases in output at low output levels, increases with further increases in output. The shape of marginal cost curve is in fact attributable to the law of variable proportions. According to the law of variable proportions, the marginal product of the variable input rises at low output levels and then falls with the expansion in output. Hence, the marginal cost curve will first fall and then rise. There are two important points to remember about the marginal cost curve:

i) The MC curve reaches its minimum point before the ATC and the AVC curves reach their minimum points; and

ii) When the MC curves rises, it cuts the AVC and the ATC curves at their minimum points.

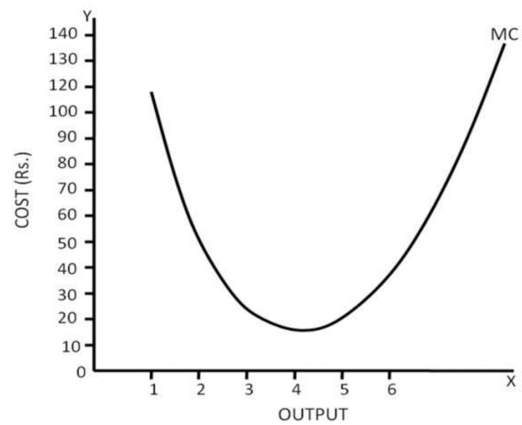

Fig. 10.9 : Marginal Cost Curve is a U-shaped Curve

10.5.5 Relationship between Marginal Cost and Average Cost

There is a close relationship between the marginal cost (MC) curve and the average total cost (ATC) and average variable cost (AVC) curves. We shall explain the relationship only between the MC curve and the ATC curve, but the relationship between the MC curve and the AVC curve can be explained along the same lines of reasoning.

Fig. 10.10 shows the MC curve together with the ATC curve and the AVC curve. The relationship between the ATC curve and the MC curve is as follows:

1) When the MC curve is below the AC curve (which means marginal cost is less than average cost), the latter falls.

2) When the MC curve is above the AC curve (which means marginal cost is more than average cost), the latter rises.

3) The MC curve intersects the AC curve at its minimum point.

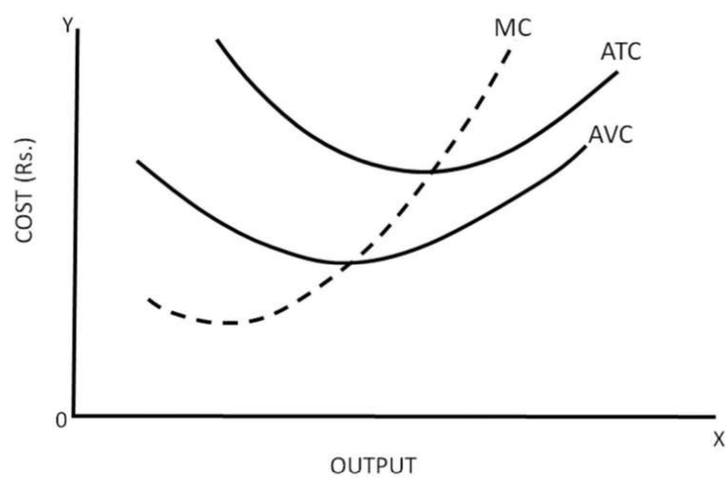

Fig. 10.10: MC curve intersects both AVC curve and ATC curve at their minimum points

The reason for the above stated relationship between the MC curve and the ATC curve is simple. So long as the MC curve lies below the ATC curve, it pulls the latter downwards; when the MC curve rises above the ATC curve, it pulls the latter upwards. Consequently, marginal cost and average total cost are equal where the MC curve intersects the ATC curve. Further when output is small, marginal cost remains lower than average total cost; but when output is expanded, marginal cost exceeds average total cost. Thus, it is natural that the MC curve intersects the ATC curve at its minimum point.

Another important feature of the relationship between MC and AC curves is that MC is affected only by variable costs. Fixed costs do not affect marginal costs. This can be proved algebraically as follows:

$$MC_N = TC_N - TC_{N-1}$$

$$= (TFC_N + TVC_N) - (TFC_{N-1} + TVC_{N-1})$$

Since, TFC_N will always be equal to TFC_{N-1} we can also state as follows:

$$MC_N = TFC_N + TVC_N - TFC_{N-1} - TVC_{N-1}$$

$$= TVC_N - TVC_{N-1}$$

This proves that MC is affected only by TVC and not by TFC.

Check Your Progress 3

1) Indicate the following statement as true (T) or false (F):

   i) Average fixed cost curve is a rectangular hyperbole ( )

   ii) Average variable cost curve is the reciprocal of the average variable factor productivity curve ( )

   iii) The average total cost curve has inverted U shape ( )

   iv) When the MC curve is below the AC curve, the latter rises ( )

2) What is average cost? What is the nature of the average total cost curve?

   ....

   ....

   ....

3) Define and distinguish between average cost and marginal cost.

   ....

   ....

   ....

4) Explain the relation between the average cost and the marginal cost. How is it possible that the marginal cost continues to rise while average cost declines?

   ....

   ....

   ....

5) The following table gives information on total cost, total fixed cost and total variable cost for a firm for different levels of output:

Output ⟶	0	1	2	3	4	5	6
TFC (Rs.)	120	120	120	120	120	120	120
TVC Rs.)	0	60	80	90	105	140	210
TC (Rs.)	120	180	200	210	225	260	330

Find (i) AFC (ii) AVC (iii) AC and (iv) MC.

....

....

....

....

10.6 LONG-RUN COST CURVES

In the long-run, all factors are variable. Due to the absence of fixed factors in the production function, all costs of production are variable in the long-run and therefore there is no need to distinguish between fixed and variable costs as is done in the short-run. In the long-run, to increase the level of production, all factors have to be increased and this results in an expansion of scale.

In the short-run, the production capacity of the firm depends upon the size of the plant. Generally, there are many options before a firm. According to the circumstance, it can choose any plant out of the large and small plants available to it. Let us suppose that a firm has three options and corresponding to them, the short-run average total cost (SATC) curves are as given in Fig. 10.11. We shall call the smallest plant as A, the medium size plant as B, and the large size plant as C. The short-run average total cost curves corresponding to these plants are designated $SATC_a$, $SATC_b$ and $SATC_c$.

The firm decides about the size of plant keeping the market considerations in view. If the demand is small, the firm will use plant A for purposes of production but in doing so it will have to incur a higher average total cost. If the firm has to produce OQ_2 quantity of output, it has two options open before it: firstly, it can employ plant A. The optimum level of output that can be produced with the help of this plant is itself OQ_2. Secondly, it can opt for plant B. If it does so, the capacity of plant B will not be fully utilised nevertheless per unit cost of production will be lower than the cost of production the firm will have to incur if it opts for producing OQ_2 amount of output with the help of plant A (even though OQ_2 is the optimum level of output that can be produced on plant A). This is due to the tendency of 'increasing returns to scale'. Not that plant C is larger in size than plant B yet, the curve $SATC_c$ is higher than the $SATC_b$ curve. If the firm opts for plant C in this case, the average total cost will increase due to the operation of 'diminishing returns to scale'.

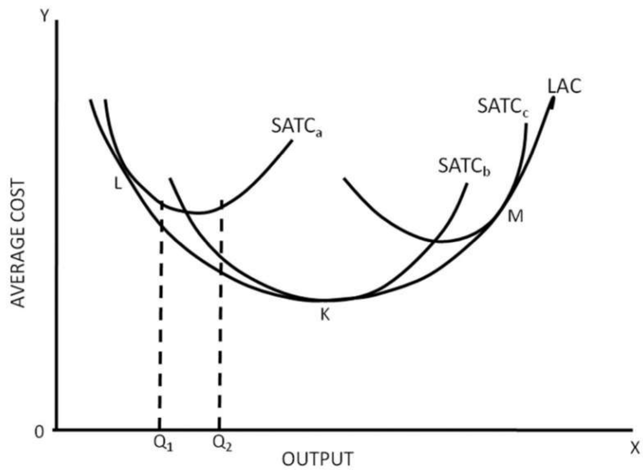

Fig. 10.11 : Long-run average cost curve envelopes short-run average total cost curves

Theoretically speaking, the long-run average cost (LAC) curve touches the short-run average total cost (SATC) curves on their minimum points. Geometrically this is possible only under those circumstances when the tendency of constant returns to scale prevails. It is due to the fact that initially increasing returns to scale and after some time diminishing returns to scale prevail in the production process that the LAC curve touches the lowest SATC curve at its minimum point. In the phase of increasing returns to scale when average total cost is falling, the LAC curve touches the SATC curves to the left of the minimum points of the SATC curves and in the phase of diminishing returns towards the right of minimum points of these curves. In Fig. 10.11, the curve LAC touches the $SATC_b$ curve at its minimum point K, the $SATC_a$ curve towards the left of its minimum point (at L) and the SAT C_c curve towards the right of its minimum point (at M).

In Economics, we say that the long-run average cost curve (LAC) 'envelopes' the short-run average total cost (SATC) curves.

10.6.1 Long Period Economic Efficiency

The behaviour of the firm which seems to be efficient in the short-run may be found to be inefficient in the long-run. To understand this let us consider Fig. 10.12. Let us suppose that the firm is producing OQ_1 quantity of output. If, due to an increase in demand, the firm wishes to increase output by Q_1Q_2, plant cannot be changed in the short-run and only variable factors will be increased. Thus, the firm will advance on the curve $SATC_1$. As a result, the efficiency of the variable resources will improve and per unit production cost will decline from BQ_1 to JQ_2. In the short-run the level of efficiency cannot improve further because this is the optimum level of production that can be achieved with the help of the plant available to the firm. However, in the long-run to produce the level of output OQ_2, the use of plant of such a small size is inefficient. If the firm uses a plant of a larger size, it will benefit from the increasing returns that would thus become available. As a result, the per unit cost will fall and come down to the level KQ_2. Though the full capacity of this plant will not be fully utilised, even then it would be more efficient as compared to the earlier plant.

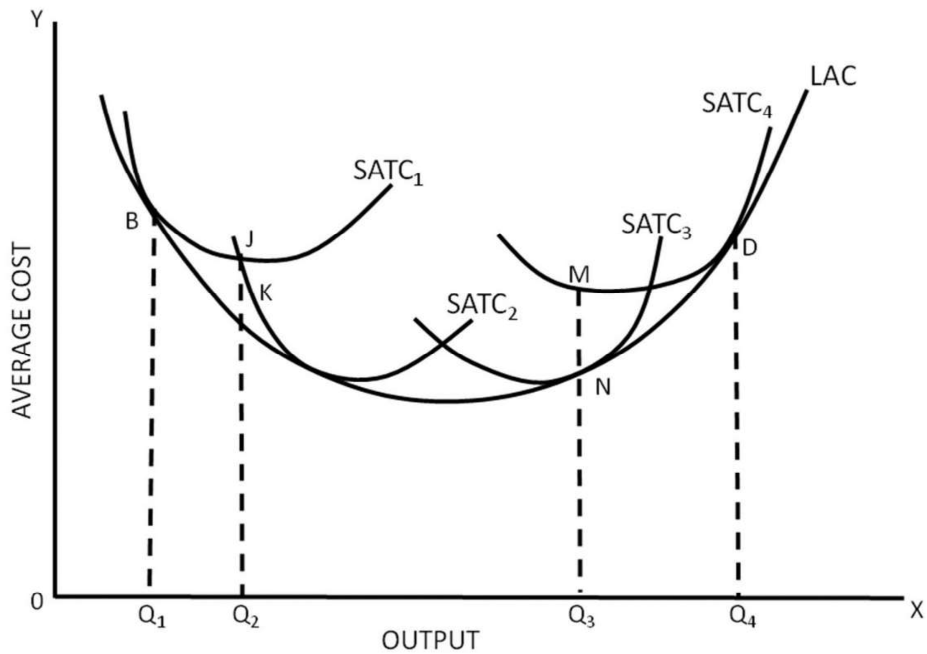

Fig. 10.12: Explanation of long-run economic efficiency

In a similar way when an expansion in scale leads to diseconomies or diminishing returns to scale emerge, it will be in the interest of the firm to reduce the level of production. If the firm is producing the output OQ_4 in Fig. 10.12, it will not be a right strategy from the point of view of maximising profits. The firm can cut down production by Q_3Q_4 in the short-run and this will enable it to reduce the average total cost from DQ_4 to MQ_3. This will result in optimum use of the plant. However, in the long-run, this position will not be satisfactory as the firm can reduce the average cost to the level NQ_3 by reducing the size of the plant. Since $NQ_3 < MQ_3$, the position which was optimum for the firm in the short-run becomes inefficient in the long-run. It is clear that when the firm uses plant of a relatively small size, it produces output much larger than is technologically optimum yet the cost remains low because it becomes possible to reduce the diseconomies of the large plant.

10.6.2 The Long-Run Average Cost Curve

We have explained in detail above that the short-run average total cost curve is U-shaped. Let us now discuss the shape of long-run average cost curve. There is general agreement that the long-run average cost falls initially due to economies of scale. But whether it falls to a certain point and then becomes constant or rises again, cannot be conclusively said.

In traditional analysis, the long-run average cost (LAC) curve is assumed to be U-shaped (as in Fig. 10.12). The shape of the long-run average cost curve is based on the assumption that ultimately the tendency of diminishing returns operates in the production process. If this belief of the economists is correct that every producer wishes to maximise profits and conditions of production are perfectly competitive, then it is true that the LAC curve must ultimately rise to the right.

10.6.3 Long-Run Marginal Cost Curve

After having understood the meaning of short-run marginal cost, it is not difficult to understand what long-run marginal cost is. Long-run marginal cost designates the change in total cost consequent upon a small change in total output when the firm has ample time to accomplish the output changes by making the appropriate adjustments in the quantities of all resources used, including those that constitute its plant. As can be seen, this definition of long-run marginal cost is practically the same as the definition of short-run marginal cost given by us earlier. The only difference between the two is that whereas in the short-run the existing plant will continue to be used for affecting an increase in output, in the long-run the plant itself will be changed.

As far as the relationship between the long-run marginal cost curve and long-run average cost curve is concerned, it is precisely the same as exists between the short-run marginal cost curve and the short-run average total cost curve. This would be clear from a mere glance at Fig. 10.13.

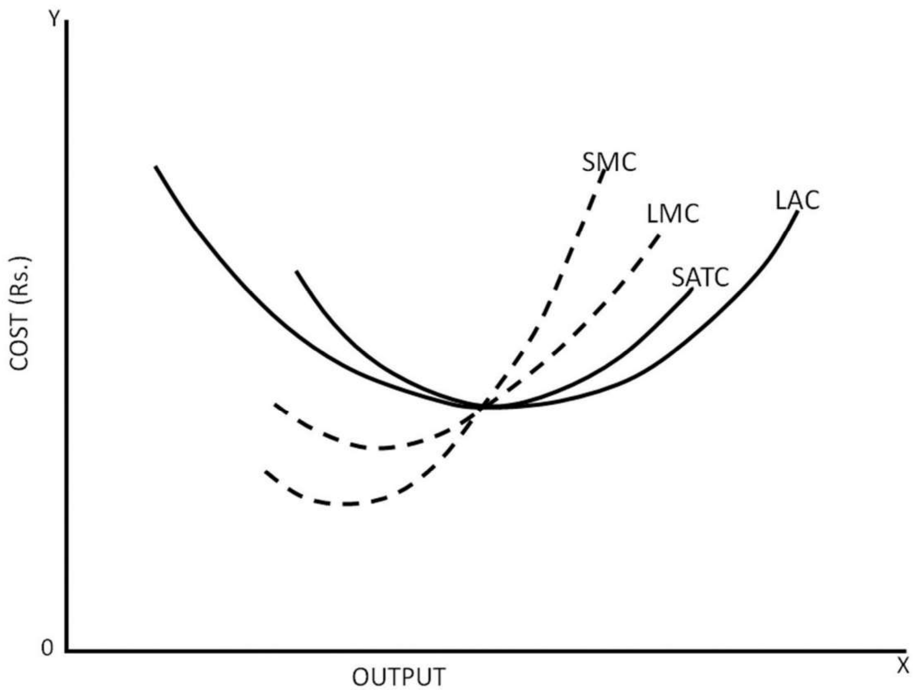

Fig. 10.13: Long-run marginal and average cost curves

10.6.4 Relationship between Long-Run Marginal Cost and Short-Run Marginal Cost

When to produce a certain given level of output, a firm sets up the most efficient plant, its short-run marginal cost (SMC) becomes equal to its long-run marginal cost (LMC). Let us explain this with the help of Fig. 10.14. In this figure, the given quantity of output is OQ_1. This output can be produced at lowest unit cost with the help of plant A. The short-run average cost curve of the firm when it produces with the help of plant A is given by SAC. Short-run average cost curves corresponding to other plants have not been drawn in Fig. 10.14. It is clear from the figure that at OQ_1 level of output, SMC and LMC are equal. However, we must see why they should be equal.

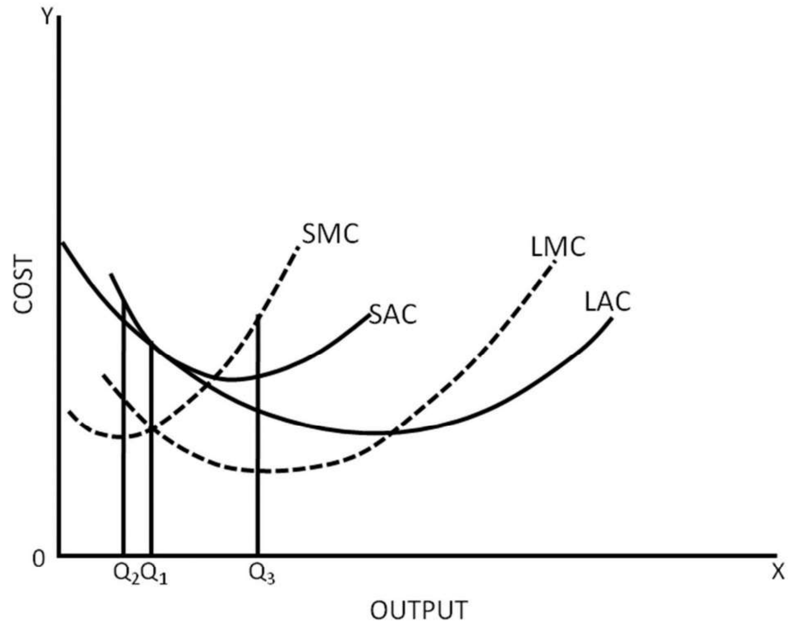

Fig. 10.14: Equality of SMC and LMC on use of an optimum size plant

To find out why SMC and LMC must be equal at the level of output OQ_1, let us consider the implications of a small change in the output by a small amount. For instance, let us take the level of output OQ_2. At this output level, short-run average cost will be greater than long-run average cost (SAC > LAC). In other words, short-run total cost is greater than long-run total cost (STC > LTC). When output rises from the level OQ_2 to the level OQ_1 the short run total cost becomes equal to the long-run total cost. If the level of output is raised to OQ_3 then since SAC is greater than LAC at this output, STC will also be greater than LTC. In other words, when output level is raised beyond OQ_1, we find that SMC exceeds LMC. Actually as we move from OQ_2 to OQ_1, we find that rate of decline in SMC is declining. In fact, beyond OQ_1, it stands rising. On the other hand, LMC keeps falling over the entire range. Therefore, between OQ_1 and OQ_3 SAC is rising and LAC is falling.

On practical considerations, the equality of short-run marginal cost and the long-run marginal cost is very significant for a firm. If the firm has to increase the level of output only by a very small amount whether it continues to employ the existing plant and changes only the quantity of the variable resources or makes a small change in the size of the plant, the results are the same. Therefore, from the point of view of the firm, both the methods are equally correct.

Check Your Progress 4

1) Indicate the following statements as True (T) or False (F):

   i) There is no need to distinguish between fixed costs and variable costs in the long-run. ( )

   ii) Long-run average cost curve envelopes the short-run average total cost curves. ( )

   iii) Long-run marginal cost curve cuts the long-run average cost curve from below at the latter's lowest point. ( )

2) Discuss the nature of the long-run average cost curve.

....

....

....

3) Discuss the concept of long period economic efficiency.

....

....

....

4) What is the relationship between long-run marginal cost curve and long-run average cost curve.

....

....

....

5) Discuss the relationship between long-run marginal cost and short-run marginal cost.

....

....

....

10.7 LET US SUM UP

In this unit, we start with a discussion of the various concepts of cost like private cost, social cost, and economic cost and accounting cost. This is followed by a discussion of short-run and long-run cost functions. We then proceed to define the distinction between fixed cost and variable cost. We note that total fixed cost curve is a straight line while the total variable cost curve and the total cost curve rise upwards to the right. We then turn to a discussion of short-run cost curves .We note that the nature of the average fixed cost curve is that of a rectangular hyperbola. When average variable cost curve is added to the average fixed cost curve, we get the average cost curve. This is followed by a discussion of the marginal cost and the nature of the marginal cost curve. The marginal cost curve cuts the average cost curve from below at the latter's minimum point.

10.8 REFERENCES

1) Robert S. Pindyck, Daniel L. Rubinfeld and Prem L. Mehta, *Microeconomics* (Pearson Education, Seventh Edition, 2010). Chapter 6, Section 6.1, 6.2, 6.3 and 6.4.

2) Dominick Salvatore, *Principles of Microeconomics* (Oxford University Presss, Fifth Edition, 2010). Chapter 8, Section 8.1, 8.2, 8.3, 8.4 and 8.5

3) A. Kountsoyiannis, *Modern Microeconomics* (The Macmillion Press Ltd., Second edition, 1982), Chapter 4.

4) John P. Gould and Edward P. Lazear, *Microeconomic Theory* (All India Traveller Bookseller, Sixth edition, 1996). Chapter 8.

5) Ahuja H.L., *Advanced Economc Theory* (S.Chand & Company Ltd., New Delhi 2001), Chapter 20 Page 396-439.

10.9 ANSWERS OR HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) (T) ; ii) (T) ; iii) (F) ; iv) (T) ; v) (T) ; vi) (F) .

2) See Sub-section 10.2.2 of Section 10.2.

3) See Sub-section 10.2.1 of Section 10.2.

4) See Sub-section 10.2.4 of Section 10.2.

5) See Sub-section 10.2.5 of Section 10.2.

Check Your Progress 2

1) (T) ; ii) (T) ; iii) (T) ; iv) (F)

2) See Section 10.3

3) See Sub-section 10.4.1 and 10.4.2 of Section 10.4

4) See Sub-section 10.4.3, 10.4.4 and 10.4.5 of Section 10.4

Check Your Progress 3

1) (T) ; ii) (T) ; iii) (F) ; iv) (F)

2) See Sub-section 10.5.3 of Section 10.5

3) See Sub-sections 10.5 .3 and 10.5.4 of Section 10.5

4) See Sub-section 10.5.5 of Section 10.5

5) (I)AFC = $\frac{TFC}{Q}$ (ii) AVC = $\frac{TVC}{Q}$ (iii) AC =$\frac{TC}{Q}$ (iv) MC = $\frac{\Delta(TC)}{\Delta Q}$

Check Your Progress 4

1) (i) T (ii) T (iii) T

2) See section 10.6

3) See Sub-section 10.6.1 of Section 10.6

4) See Sub-sections 10.6.2 and 10.6.3 of Section 10.6

5) See Sub-section 10.6.4 of Section 10.6

GLOSSARY

Average Product	:	Total product divided by the number of units of the input used is average product.
Accounting Cost	:	Accounting cost refers to actual expenses of the firm plus depreciation charges for capital equipment.
Barter	:	Exchange of goods/services against other goods/services.
Budget Line	:	The Budget Line, also called as Budget Constraint shows all the combinations of two commodities that a consumer can afford at given market prices and within the particular income level.
Comforts	:	Goods which are used for increasing our productive capacity and for making our lives more comfortable.
Consumption	:	Using up of Utility of goods in the satisfaction of a want.
Change in Demand	:	Shift of the entire demand of curve.
Change in Quantity Demanded	:	Movement on a demand curve itself caused by a changes in the price of the commodity in question.
Contraction in Supply	:	The decrease in quantity supplied because of a fall in the price of the commodity.
Curvilinear Supply Curve	:	The supply curve which is not a straight line.
Cardinal Utility	:	The **Cardinal Utility** approach is propounded by neo-classical economists, who believe that utility is measurable, and the customer can express his satisfaction in cardinal or quantitative numbers, such as 1, 2, 3 and so on.
Consumer Equilibrium	:	The point at which a consumer reaches optimum utility, or satisfaction, from the goods and services purchased, given the constraints of income and prices.
Constant Returns to Scale	:	Constant returns to scale implies that when all inputs are increased in a given proportion, output increases in the same proportion.
Complementary Commodity	:	It is the commodity whose demand is directly related to the demand of the commodity in question.
Demand	:	The amount of goods which the buyers are ready to buy, per period of time, at a given price per unit.

Dependent Variable	:	A variable which changes only with the change in the independent variable.
Decrease in Supply	:	The decrease in quantity supplied at a given price of the commodity.
Diminishing Returns to Scale	:	Diminishing returns to scale refers to the case when output grows proportionally less than input.
Economic Laws	:	Statements of tendencies. They depict the standardised or generalised response of economic units to different forces and stimuli.
Exchange Value	:	The price which an item commands in the market.
Elasticity of Demand	:	It quantifies the strength relationship between the quantity demanded of commodity and the price of the commodity or income of the consumer or price of another commodity which is related to the commodity in question.
Elasticity of Supply	:	The responsiveness of quantity supplied to a given percentage change in the price of the commodity.
Extension in Supply	:	The rise in quantity supplied due to a rise in the price of the commodity.
External Economies	:	When a firm enters production, it obtains a number of economies for which the firm's own strategies/policies are not responsible. These are economies external to the firm.
External Diseconomies	:	When the scale of operations is expanded, many such diseconomies accrue that have no particular ill-effect on the firm itself but their burden falls on the other firms. These are known as external diseconomies.
Economic Cost	:	Economic cost refers to cost to a firm utilising economic resources in production including opportunity cost.
Explicit Cost	:	Explicit costs arise from transaction between the firm and other parties in which the former purchases inputs or services for carrying out production.
Flow Variable	:	A variable which can be measured only with reference to a period of time.
Goods	:	Items which have a utility or can be used for the production of other goods or services.
Giffen Good	:	A good where higher price causes an increase in demand (reversing the usual law of demand). The increase in demand is due to the income effect of the higher price outweighing the substitution effect.

Historical Cost	:	Historical cost is the cost that was actually incurred at the time of purchase of an asset.
Inductive Reasoning	:	The technique of analysis in which factual information is used to discover the behaviour pattern of different economic units in response to various forces and stimuli.
Inferior Commodity (Good)	:	A commodity in which there is an inverse relationship between the income of the consumer and quantity demanded of a commodity.
Income Elasticity of Demand	:	It is the responsiveness of demand to a given proportional change in the income of the consumer.
Inequalities of Income	:	The distribution of income among different income groups of an economy.
Increase in Supply	:	The rise in quantity supplied at a given price of the commodity.
Income Effect	:	A change in the demand of a good or service, induced by a change in the consumers' real income. Any increase or decrease in price correspondingly decreases or increases consumers' real income which, in turn, causes a lower or higher demand for the same or some other good or service.
Isocost Line	:	An isocost line represents various combinations of inputs that may be purchased for a given amount of expenditure.
Isoquant	:	An isoquant is the of all the combination of two factrors of production that yield the same level of output.
Increasing Returns to Scale	:	Increasing returns to scale refer to the case when output grows proportionally more than inputs.
Internal Economies	:	Those economies that accrue to a firm on expansion of its own size are known as internal economies.
Internal Diseconomies	:	When the scale of production is continuously expanded, a point is reached where the increase in production becomes less than proportionate to the increase in the factors of production. As this point, internal diseconomies set in.
Implicit Cost	:	Implicit costs are the costs associated with the use of firm's own resources. Since these resources will bring returns if employed elsewhere, their imputed values constitute the implicit costs.
Incremental Cost	:	An incremental cost is the increase in total costs resulting from an increase in production or other activity

Luxuries : Goods which are meant for status or social standing.

Law of Supply : It shows the direct relationship between the price of a commodity and its quantity supplied, other factors influencing supply (except price of the commodity) remaining constant.

Law of Diminishing Returns : As more units of an input are used per unit of time with fixed amounts of another input, the marginal product of the variable input declines after a point.

Linear Homogeneous Production Function : When output increases in the same proportion in which inputs are increased, the production function is linear homogeneous. For example, if labour and capital are increased λ by times and, as a result, output also increases by λ times, the production function is linear homogeneous.

Merit Goods : The goods whose consumption is believed to be desirable for the benefit of the society and the consuming individuals.

Macroeconomics : Branch of economic analysis that focuses on the workings of the whole economy or large sectors of it.

Margin : The value of the variable under consideration related to the last unit of an item.

Marginal Utility : The additional or extra satisfaction yielded from consuming one additional unit of a commodity.

Microeconomics : Branch of economic analysis that focuses on individual economic units or their small groups and micro-variables like individual prices of individual commodities, etc.

Money Exchange : Sale of goods/services against money.

Monopolist : A producer who controls the whole supply of a commodity.

Marginal Product : Marginal product of an input is defined as the change in total output due to a unit change in the amount of an input while quantities of other inputs are held constant.

Marginal Rate of Technical Subsitution ($MRTS_{L,K}$) : Marginal rate of technical substitution of factor L for factor K ($MRTS_{L,K}$) is the quantity of K that is to be reduced on increasing the quantity of L by one unit for keeping the output level unchanged.

Necessities : Goods which are used for satisfying basic of existence.

Normative Economics	:	That part of economic analysis which is concerned with what ought to be, and the way it can be achieved by changing the existing situation.
Ordinal Utility	:	The Ordinal Utility approach is based on the fact that the utility of a commodity cannot be measured in absolute quantity. However, it will be possible for a consumer to tell subjectively whether the commodity gives more or less or equal satisfaction when compared to another.
Optimality	:	The point where maximum possible output is being achieved given the use of different factors of production.
Private Goods	:	Goods whose availability can be restricted to selected users. It is divisible in that sense.
Production Possibility Curve	:	A graphic representation of the combinations of maximum amounts of goods X and Y which can be produced with the given productive resources of the economy and under certain other simplifying assumptions.
Public Goods	:	Goods or services whose availability cannot be restricted to selected users only. The benefits of the goods are indivisible and people cannot be excluded.
Positive Economics	:	That part of economic reasoning which covers what is, without going into its desirability or otherwise, and without suggesting ways for changing the existing state of affairs.
Price Effect	:	The impact that a change in its price has on the consumer demand for a product or service in the market. The price effect can also refer to the impact that an event has on something's price. The price effect is a resultant effect of the substitution effect and the income effect.
Point of Inflexion	:	The point where total product stops increasing at an increasing rate and begins increasing at a decreasing rate is called the point of inflexion.
Production Function	:	The technical law which expresses the relationship between factor inputs and output is termed production function.
Rectangular Hyperbola	:	It is a curve in which every rectangle drawn with one corner on the curve has the same area.
Ridge Lines	:	The lines forming the boundaries of the economic region of production are known as the ridge lines.

Replacement Cost : Replacement cost is the cost that will have to be incurred now to replace that asset (i.e., the replacement cost is the current cost of the new asset of the same type).

Stock Variable : A variable which can be measured only with reference to a point of time.

Supply : The quantity of goods which the sellers are ready to sell, per unit of time, at a given price per unit.

Substitution Effect : It shows how with a change in the price of a commodity, prices of other commodities remaining unchanged, a consumer substitutes one commodity for the other.

Substitute Commodity : It is the commodity whose demand is inversely related to the demand of the commodity in question.

Supply Schedule : A table having two columns, one showing different prices of the commodity and the other showing quantities supplied during a given period at each of these prices.

Supply Curve : A curve showing the relationship between price of a commodity and its quantity supplied during a given period, other factors influencing supply remaining unchanged.

Sub-optimality : It is a point where optimality has not been achieved, i.e. output is less than the possible maximum given the use of the resources.

Sunk Cost : Sunk cost is a cost that has already been incurred and can’t be recovered.

Technology : The method employed to produce a commodity or service.

Total Utility : The total satisfaction derived from all the units of an item.

Use Value : Utility of goods.

Utility : The want satisfying capacity of goods. It is the service or satisfaction an item yields to the consumer.

SOME USEFUL BOOKS

1) Kautsoyiannis, A. (1979), *Modern Micro Economics*, London: Macmillan.

2) Lipsey, RG (1979), *An Introduction to Positive Economics*, English Language Book Society.

3) Pindyck, Robert S. and Daniel Rubinfield, and Prem L. Mehta (2006), *Micro Economics*, An imprint of Pearson Education.

4) Case, Karl E. and Ray C. Fair (2015), *Principles of Economics*, Pearson Education, New Delhi.

5) Stiglitz, J.E. and Carl E. Walsh (2014), *Economics*, viva Books, New Delhi.