UNIT 7 PRODUCTION WITH TWO AND MORE VARIABLE INPUTS

Structure

7.0 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;

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

- find out the level at which output will be maximised subject to a given cost;
- for a given level of output, find the point on the isoquant where cost will be minimised;
- describe the nature of optimal expansion path both in long run and short run;
- state to concept of returns to scale; and
- discuss the concept of economies and diseconomies of the scale.

7.1 INTRODUCTION

How do firms combine inputs such as capital, labour and raw materials to produce goods and services in a way that minimises the cost of production is an important issue in the principles of microeconomics. Firms can turn inputs into outputs in a variety of ways using various combinations of labour, capital and materials. Broadly there can be three ways:

1) by making change in one input or factor of production.
2) by making change in two factors of production.
3) by making change in more than two or more inputs /factor of production.

The nature and characteristics of production function of a firm under the assumption that firm makes variation in one input has been discussed in previous unit. Here we would like to discuss the nature, forms and characteristics of production function if firm decides to make variation in two or more inputs.

Let us begin to recapitulate the concept of production function.

7.2 PRODUCTION FUNCTION: THE CONCEPT

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 thus 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.

Take the case of a production process which uses two variable inputs say, labour (L) and capital (K). We can write the production function of this case as

$$\mathbf{Q = F\ (L, K)}$$

This equation relates the quantity of output Q to the quantities of the two inputs, labour and capital. A popular production function of such a case in economics is Cobb Douglas production function which is given as

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

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

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

$$\text{i.e. } \beta= 1 - \alpha$$

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

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

7.3 PRODUCTION FUNCITON WITH TWO VARIABLE INPUTS

The behaviour of the production function of a firm which makes use of two variable inputs or factors of production is analysed by using the concept of isoquants or iso product curves. Hence, let us understand the concept of isoquants.

7.3.1 Definition of Isoquants

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

Let us understand the concept of an isoquant with the help of an example. Suppose 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 7.1.

Table 7.1: Isoquant Table showing combinations of Labour and Capital producing 100 Units of X

Process	Units of Labour	Units of Capital
1	1	10
2	2	7
3	3	5
4	4	4
5	6	3
6	9	2

From Table 7.1, it is clear that all the six 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 five more processes available to the firm and in each of them factor intensities differ. The sixth 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. 7.1.

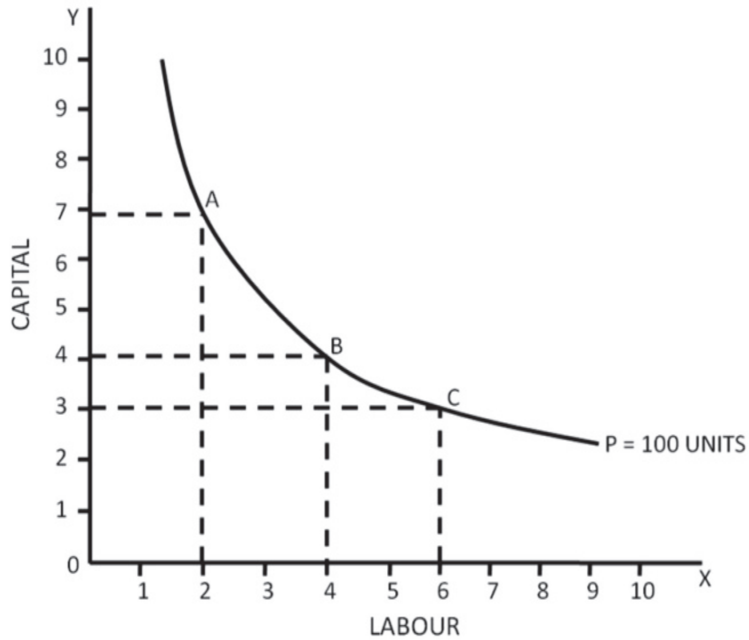

Fig. 7.1: This figure shows that at point A, B and C same level of output (=100 units) is obtained by using different combinations of labour and capital. Curve p is known as isoquant

7.3.2 Types of Isoquants

Depending upon the degree of substitutability of the factors, Isoquants can assume three shapes categorised as:

1) Convex isoquant

2) Linear isoquant

3) Input-output isoquant

1) **Convex Isoquants:** This isoquant take the shape of curve sloping downward from left to right as shown in Fig. 7.1. The explanation for assumption of this shape has been given in next section.

2) **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 shown in Fig. 7.2. In Fig. 7.2 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.

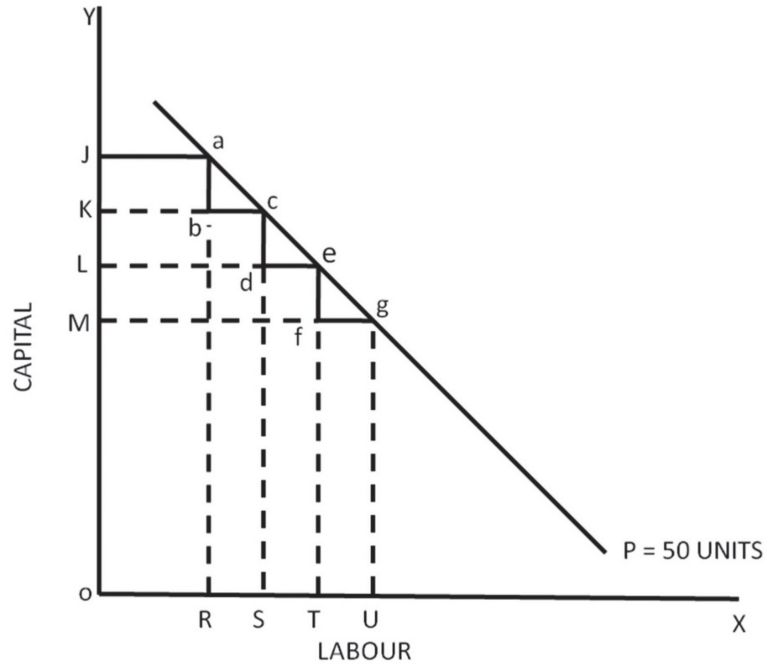

Fig. 7.2: In the case of perfect substitutability of factors of production, the isoquant becomes a straight line and is, therefore, known as linear isoquant

3) **Input-Output Isoquant:** When factors of production are not substitutes but complementary, technical coefficients are fixed. This means 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. 7.3. 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. In Fig. 7.3, the slope of isoquant P_1 and P_2 indicates the capital-labour ratio has to be maintained for ensuring efficiency in production.

Isoquant 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. 7.4.

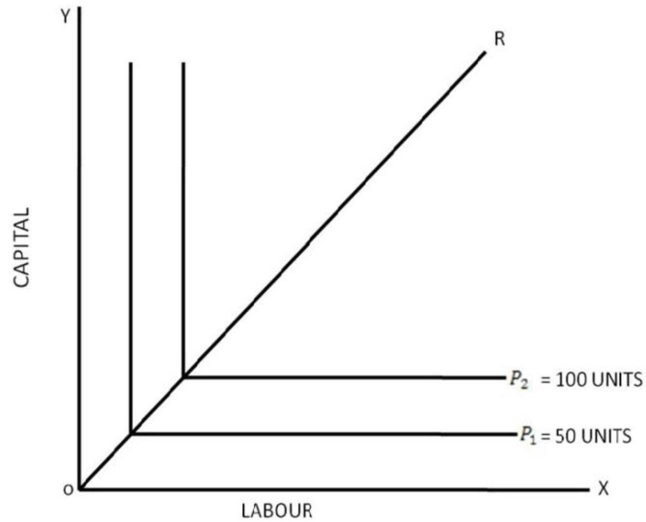

Fig. 7.3: If factors of production can be used only in a fixed proportion, the isoquant is 'L' shaped and is known as an input-output isoquant

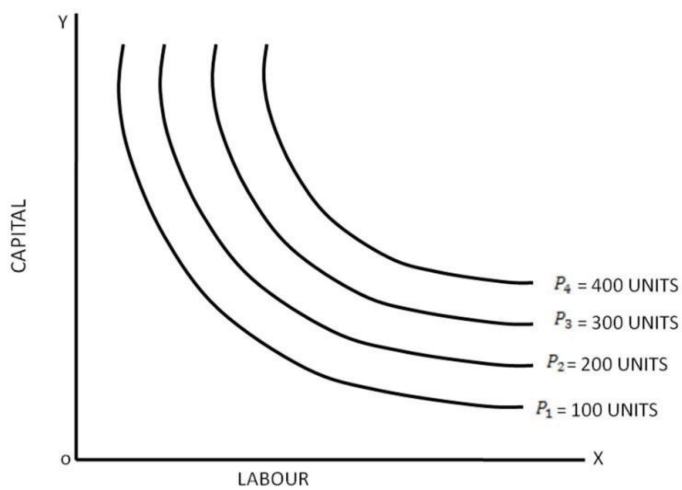

Fig. 7.4: When a number of isoquants are depicted together, we get an isoquant map

In Fig. 7.4, 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.

7.3.3 Assumptions of lsoquants

Isoquant analysis is normally based on the following assumptions:

1) 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.
2) 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).

7.3.4 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. 7.5, such an isoquant is shown. AB and CD segments of this isoquant are positively sloped.

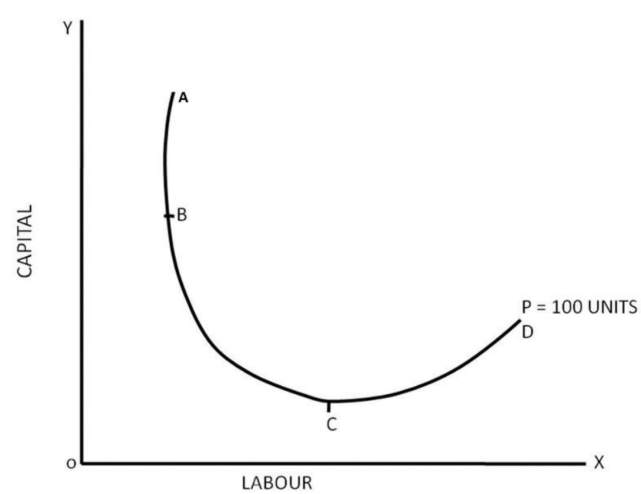

Fig. 7.5: Isoquant having positively sloped segments

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. 7.6. 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.

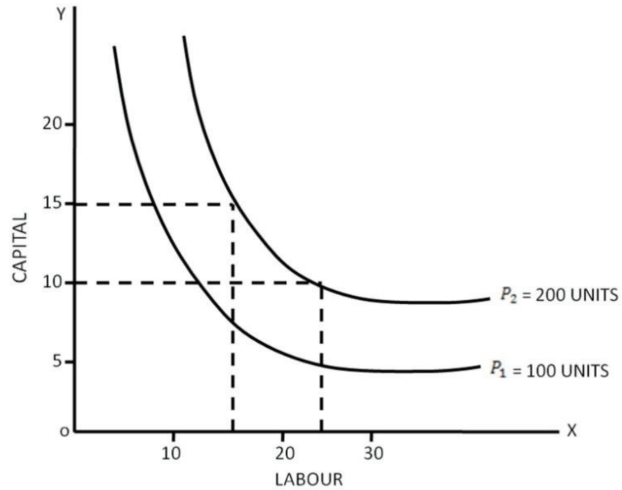

Fig. 7.6: Two isoquants representing different output levels. A higher isoquant depicts a higher amount of output

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. 7.7) 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.

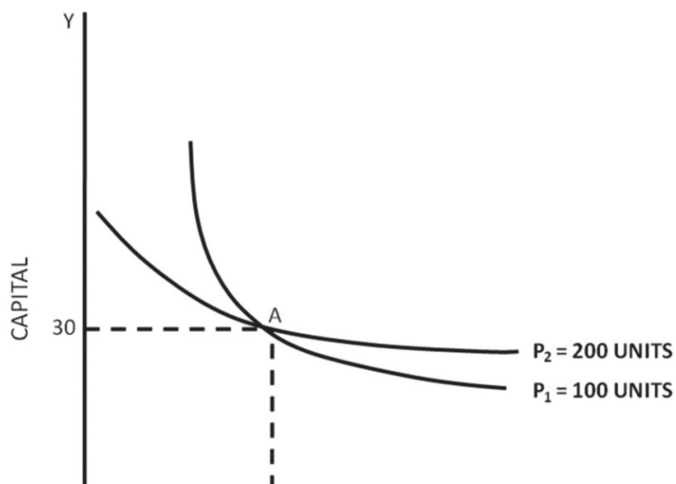

Fig. 7.7: No two isoquants intersect each other because each isoquant depicts a different level of output

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. 7.8. 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.

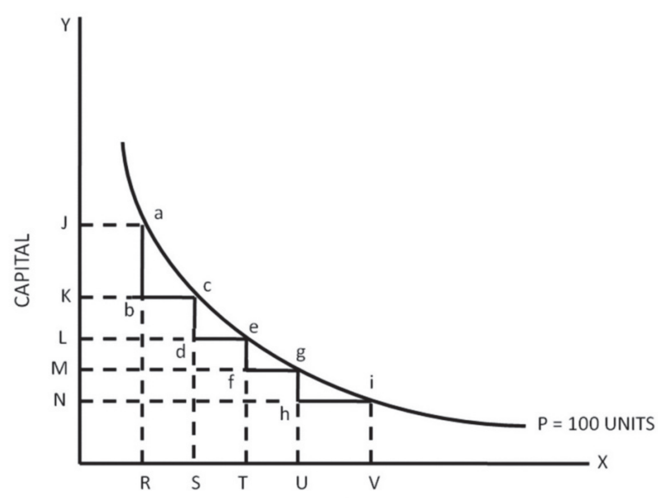

Fig. 7.8: An isoquant is covex from below because the marginal rates of technical substitution tends to fall

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. 7.2. 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. 7.3. 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) Draw the possible shapes which the isoquants may assume depending on the degree of substitutability.

   ....

   ....

   ....

7.4 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. 7.9, 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. 7.9 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.

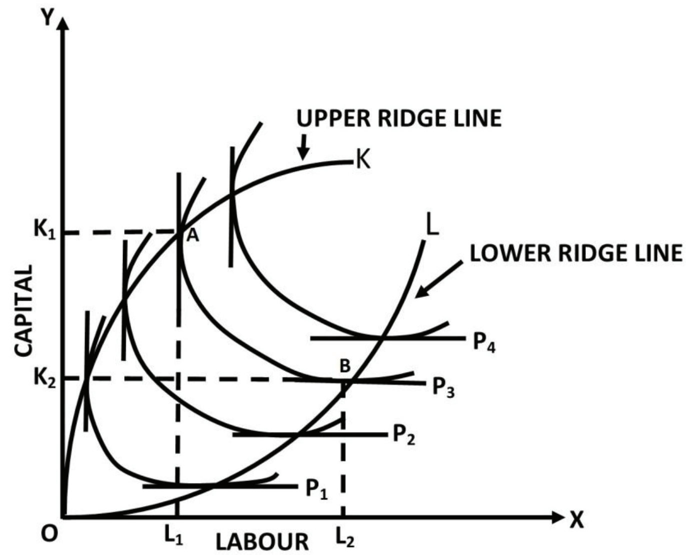

Fig. 7.9: Area enclosed within the upper side line OK and the lower side lint OL indicates the economic region of production

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 results in a fall in 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. 7.9, 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 result in smaller output. Therefore, the marginal product of capital is zero at point A. 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 will reduce rather than 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. 7.9, 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. 7.9 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.

7.5 THE OPTIMAL 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.

7.5.1 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. 7.10, 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,

$$\text{Slope} = \frac{\Delta K}{\Delta L} = \frac{K}{L} = \frac{\text{Production budget}}{r} / \frac{\text{Production budget}}{w} = \frac{w}{r}$$

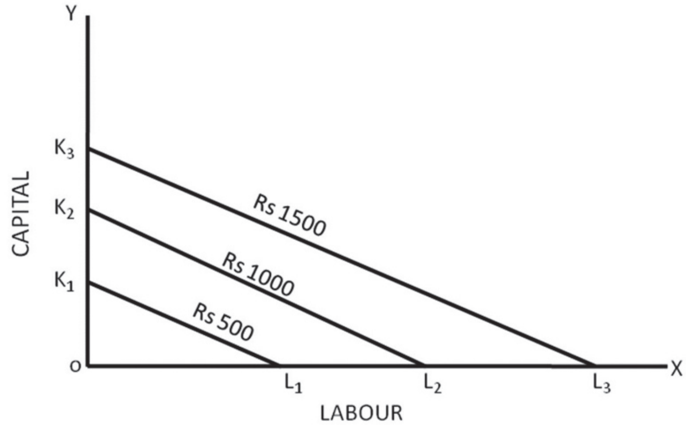

Fig. 7.10: Isocost Lines- A higher cost line indicates a higher cost

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. 7.10, 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.

7.5.2 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. 7.11 represents 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 labour as it enables the producer to reach the highest level of production possible given the cost conditions.

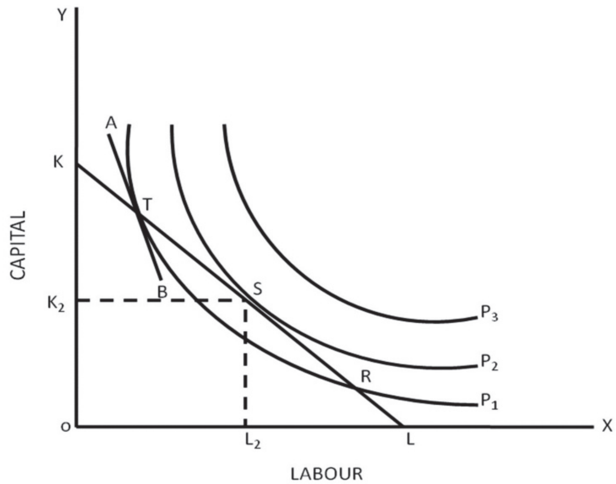

Fig. 7.11: With the given cost line KL, the highest isoquant that a producer can reach is P_2. Point S on this isoquant, therefore, indicates producer's equilibrium

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,

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

7.5.3 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. 7.12, 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.

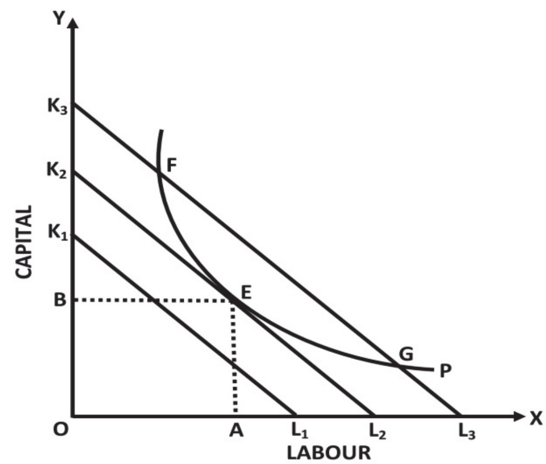

Fig. 7.12: To obtain a level of production indicated by isoquant P, the minimum cost that must be incurred is given by point E on the isocost line K_2L_2. Therefore, point E indicates the point of producer's equilibrium

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.**

7.6 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.

7.6.1 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. 7.13. 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.

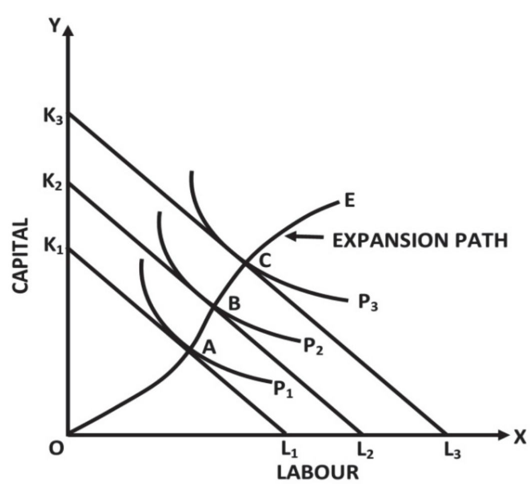

Fig. 7.13: Expansion path in the case of non-linear production function

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.

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. 7.14. 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.

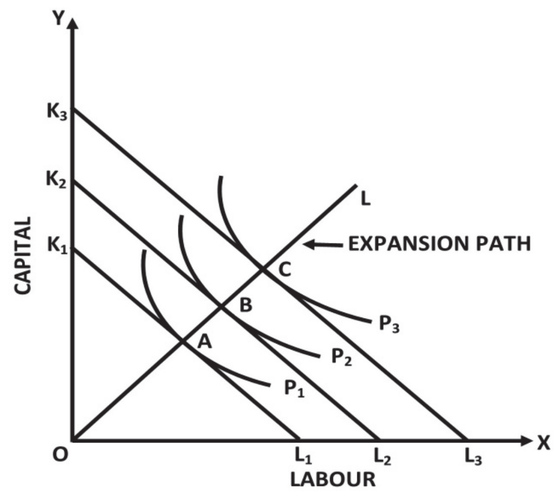

Fig. 7.14: Expansion path in the case of linear homogeneous production function is a straight line

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.

7.6.2 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. 7.15, 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. 7.15. 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.

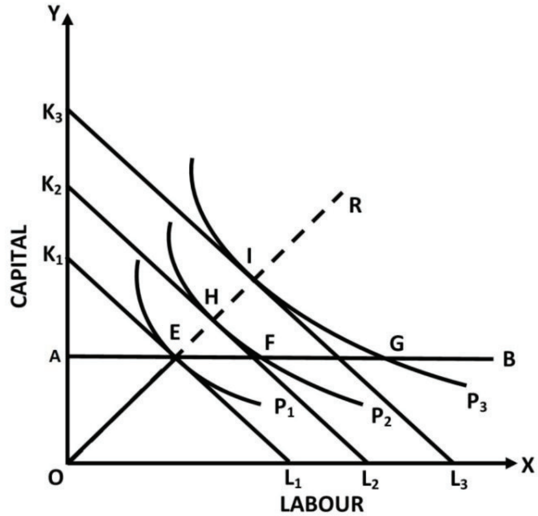

Fig. 7.15: Expansion path in the short run in the case of linear homogeneous production function

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?

   ....

   ....

   ....

7.7 PRODUCITON FUNCTION WITH SEVERAL VARIBALE INPUTS

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.**

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, i.e. use of all the factors is changed in same proportion.

7.7.1 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.

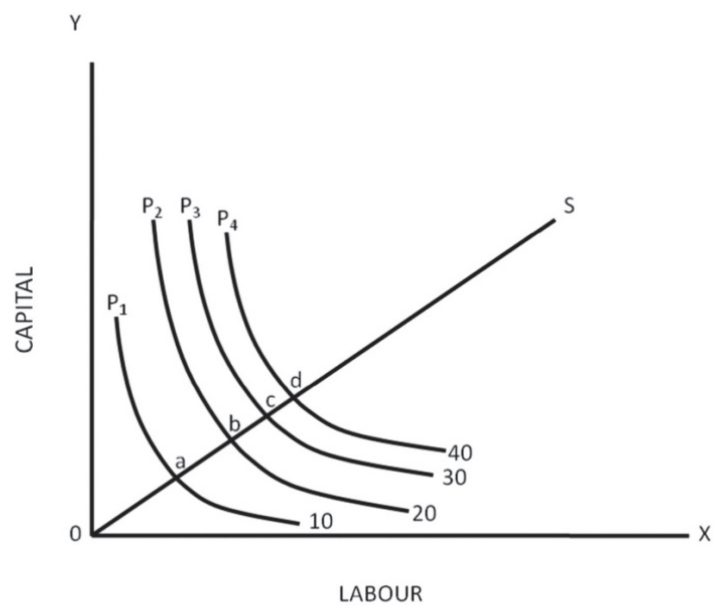

Fig. 7.16: Increasing Returns to scale output increases in a greater proportion than the increase in the factors of production

There are main factors which account for increasing returns to scale are given below:

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.

7.7.2 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. 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. 7.17.

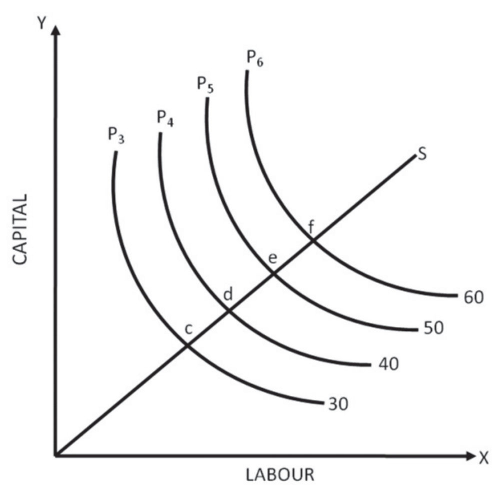

Fig. 7.17: Constant Returns to Scale-output increases in the same proportion in which inputs are increased

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.

Further if the factors of production are perfectly divisible, the production function must exhibit constant returns to scale.

7.7.3 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. 7.18.

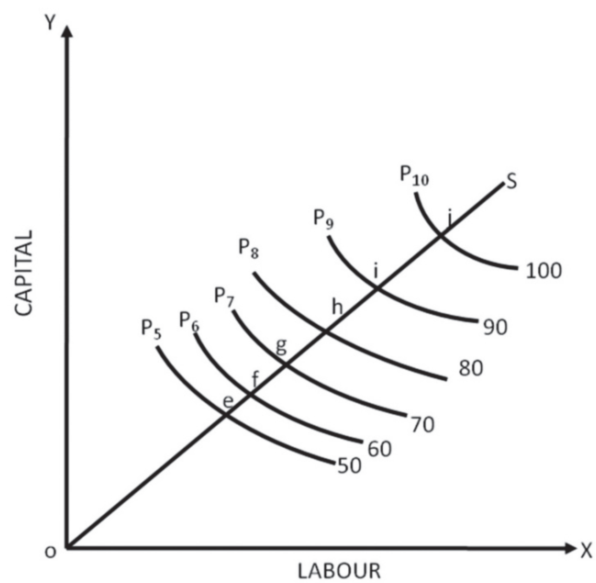

Fig. 7.18: Diminishing Returns to Scale – output increases proportionally less than inputs

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 of different factors of production tends 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.

7.8 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.

7.8.1 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. These economics that determine the nature of the long-run average cost curve are listed below:

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: (1) production economies, (ii) selling or marketing economies, (iii) managerial economies, and (iv) economies in transport and storage.

2) **Pecuniary Internal Economies:** 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 on 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.

7.8.2 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 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.

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.

7.8.3 External Economies of Scale

External economies were discussed first of all by Alfred Marshall. According to him, when a firm enters production, it obtains a number of economies 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.

7.8.4 External Diseconomies of Scale

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.

Check Your Progress 3

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

   ....

   ....

   ....

7.9 LET US SUM UP

The unit begins with the concept of production function which referes 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. When the ratio between the factors is kept constant and several variable inputs are used this give rise to three possibilities– increasing returns to scale, constant returns to scale and diminishing returns to scale. Economics of scale are decided into two parts- internal economies of scale and external economies of scale. Similarly diseconomies of scale are both internal and external.

7.10 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.

7.11 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 7.3.1 and 7.3.4

3) See Sub-section 7.3.2

Check Your Progress 2

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

2) See-section 7.5.3

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

Check Your Progress 3

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

2) See Sub-section 7.7.1 and answer

3) See Sub-section 7.8.3 and 7.8.4