Econ 2021 – Microeconomics-I

Chapter three

Theory of Production

Bekele M.

Contents

1. Production function
   - The Short run Production Function
   - The Long run Production Function
2. Laws of production
   - The law of variable proportions
   - Laws of returns to scale
3. Choice of optimal combination of factors of production

THEORY OF PRODUCTION

Production Function

- A production function shows the technical relationship between factor inputs and output.
  - It describes the amount of output expected from different combination of input usage.
- It can be expressed in tabular or graphic form or by a mathematical formula.

Factors of production

- A method of production (process, activity) is a combination of factor inputs required for the production of one unit of output. Usually a commodity may be produced by various methods of production.

THEORY OF PRODUCTION

For example,

- A unit of commodity x may be produced by the following processes:

	Process P1	Process P2	Process P3
Labour units	2	3	1
Capital units	3	2	4

- Factors of production could be fixed or variable. The difference between fixed and variable factors relates to the time horizon involved.
- In economics, there are two main horizons;

  The short run & **The long run** production function.

THEORY OF PRODUCTION

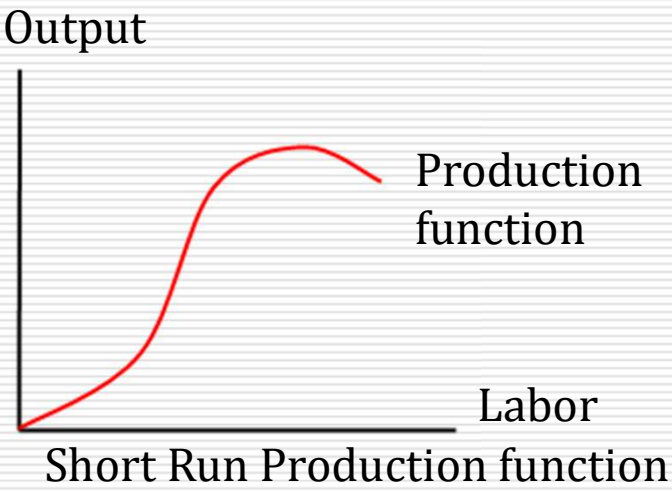
The Short run Production Function

- The short run is a relatively short period of time in which the quantity of some factors of production such as equipment and buildings cannot be varied. Such factors are called fixed factors.
- Factors of production whose quantity can be varied in the short run are called variable factors.
- The long run, on the other hand, is a relatively long period which allows the variation of all factors of production including plants and equipment.
- In this section, we will focus on a production process in which there are only two factors of production; labour and capital, where capital is the fixed factor and labour is the variable factor.

THEORY OF PRODUCTION

- The production function in this case is defined by $y = f(L)_{\bar{k}}$

As more and more units of a variable factor (L) are combined with fixed factors (in our case K), we may initially obtain increasingly larger additions to output, but we eventually obtain smaller increments in output



This economic phenomenon is referred to as the principle of diminishing returns (also called the law of variable proportions)

The bar on K indicates that capital is constant and variation in output depends on variation in labour L

THEORY OF PRODUCTION



THEORY OF PRODUCTION

Average and Marginal Product Curves

- Given the production function which transforms a set of inputs into output, the technical relationship between inputs and output could further be analysed by deriving marginal and average product curves.
- The behaviour of total product curve (schedule) has an interesting implication to the marginal cost and average cost curves (schedules).
- Total product (TP) is the total amount that is produced during a given period of time.
  - If the inputs of all but one factor are held constant, total product will change as the quantity of the variable factor used changes.

THEORY OF PRODUCTION

- Suppose capital is fixed at 5 units. By applying varying quantity of labour, the firm can produce different levels of output

Qty of capital (*K*)	Qty of Labor (*L*)	Total product (*TP*)	Average product (*AP*)	Marginal product (*MP*)
5	0	0	-	-
5	1	15	15.0	15
5	2	34	17.0	19
5	3	48	16.0	14
5	4	60	15.0	12
5	5	62	12.4	2
5	6	60	10	-2

Table 3.1: Total Product, Average Product and Marginal Product

THEORY OF PRODUCTION

- Capital being held fixed at 5 units, as the quantity of labour increases, the level of output (TP) increases.
- Average product (AP) increases at first and then declines. The same is the true of marginal product.
- Average product (AP) is the total product divided by the number of units of the variable factor used to produce it.
- If we let the number of units of labour be denoted by L, the average product can be written as

$$\text{AP}\ \frac{TP}{L}$$

- As more of the variable factor (labour) is used, AP first rises and then falls.

THEORY OF PRODUCTION

- The level of output at which AP reaches maximum is called the point of diminishing average productivity. Up to that point, average productivity is increasing; beyond that point, average productivity is decreasing.
- Marginal Product (MP) – is the change in total product resulting from the use of one unit more of a variable factor.
- As the first unit of labour is an employed, total product increase by 15 units. This increase in TP is the marginal product of the first unit of labour.
- The second unit of labour adds 19 more units of output to TP; therefore, its MP is 19 units.

THEORY OF PRODUCTION

- The fifth unit of labour's MP is just 2 units of output, while the last unit of labour's contribution to output is negative.
  - MP= $\frac{dTP}{dL}$, where d(TP) is the change in TP, and d(L) is the change in L.
- At initial stages, MP increases as additional variable factors (labour) are employed. Then reaches maximum and declines. After certain range, it becomes negative with employment of additional variable factors.
- The level of output at which, marginal product reaches its maximum level is called point of diminishing marginal productivity.

THEORY OF PRODUCTION

Example

- Suppose the production function that a firm faces is given as

$$Q = f(L)_K = 8L^2 - \frac{2}{3}L^3$$

- This production function shows the maximum output that can be produced from various levels of labour employment, therefore represents TP. The AP is, then, derived as

$$AP = \frac{TP}{L} = \frac{8L^2 - \frac{2}{3}L^3}{L} = 8L - \frac{2}{3}L^2$$

- Given the AP function, the point of diminishing average productivity occurs where the slope AP is zero.

THEORY OF PRODUCTION

- The slope of

$$AP = \frac{dAC}{dL} = 8 - \frac{4}{3}L$$

$$\frac{4}{3}L = 8$$

$$L = 6$$

- Similarly, MP is derived as

$$\mathrm{M}P = \frac{dTP}{dL} = \frac{d(8L^2 - \frac{2}{3}L^3)}{dL} = 16L - 2L^2$$

THEORY OF PRODUCTION

- Given the MP function, the point of diminishing marginal productivity occurs where the slope of MP is zero.
- The slope of MP

$$= \frac{dMP}{dL} = 16 - 4L$$

$$4L = 16$$

$$L = 4$$

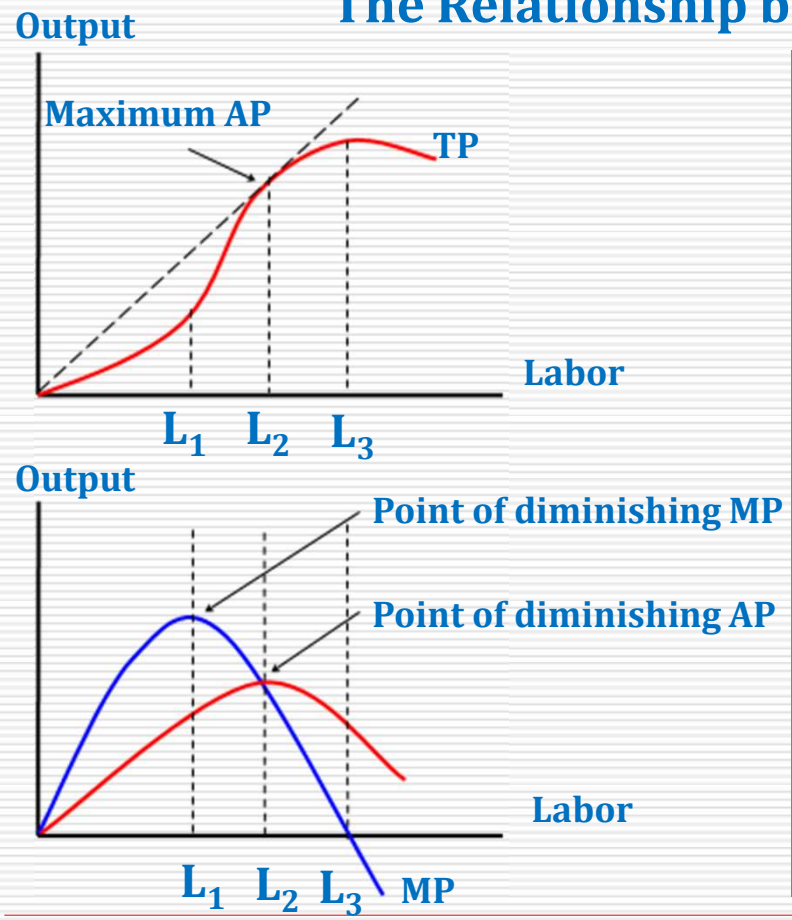
THEORY OF PRODUCTION

Relationship between MP and AP: Graphic approach

- To clearly understand this relationship between AP and MP consider, for example, a class of 10 students.
- Suppose the average height of the students is 170 cm. If now we bring in a tall student with a height of 192 cm, the average height of the class will increase to 172 cm.
- This happens simply because the height of the new student (which we may call the marginal height) is greater than average height of the class.
- On the contrary, if we bring in a short student with height just 159 cm, the average height of the class will fall to 169 cm, because the height of the new student (marginal height) is less than the average height of the class.

THEORY OF PRODUCTION

The Relationship between AP and MP



- For employment level between 0 and L1 units, TP increases at an increasing rate (slope of the TP becomes steeper and steeper), this means that MP and AP are also increasing.
- At L1, MP reaches its maximum. Between L1 and L2 units of labour, TP increases at a decreasing rate and MP declines, but is above AP. As a result, AP increases.
- At L2, AP reaches maximum and is cut by MP. Between L2 and L3, TP still increases at a decreasing rate, but MP is below AP. This pulls the AP down and AP falls.
- At L3, MP is zero and TP reaches maximum. Beyond L3, MP is negative, therefore, TP declines.

THEORY OF PRODUCTION

The Relationship between MP and AP: Algebraic Approach

- Suppose the production function is given by $Q = -cL + bL^2 - aL^3$
- The marginal product and average product functions are, then defined as

$$MP = \frac{dQ}{dL} = -c + 2bL - 3aL^2$$

$$AP = \frac{Q}{L} = -c + bL - aL^2$$

- The maximum point of MP occurs where its slope is zero.

$$2b - 6aL = 0 \qquad L = \frac{b}{3a}$$

THEORY OF PRODUCTION

- The maximum point of AP, on the other hand, occurs where its slope is zero.

$$b - 2aL = 0 \qquad L = \frac{b}{2a}$$

- Since , $\frac{b}{3a} < \frac{b}{2a}$ the maximum point of MP lies to the left of the maximum point of AP.
- Let's now try to show that MP cuts AP at its maximum point. If this is true, the point of intersection of MP and AP must coincide with the maximum point of AP.

THEORY OF PRODUCTION

- At the point of intersection $MP = AP$ This means:

$$-c+2bL-3aL^2 = -c+bL-aL^2$$

$$3aL^2 - aL^2 + bL - 2bL = 0$$

$$2aL^2 - bL = 0$$

$$L(2aL - b) = 0$$

$$L = 0 \text{ or } 2aL = b$$

$$L = 0 \text{ or } L = \frac{b}{2a}$$

- With the production function defined as , $Q = f(L)_{\bar{K}}$ however, we are interested with positive level of labour employment. Therefore $L = \frac{b}{2a}$ The result verifies that MP cuts AP at its maximum point.

THEORY OF PRODUCTION

Example

- Consider the production function $Q = 8L^2 - \frac{2}{3}L^3$ The MP and AP functions were found to be

  $MP = 16L - 2L^2$ or $AP = 8L - \frac{2}{3}L^2$

- In the previous example, it was found that the maximum point of AP occurs at L = 6. At the point of intersection of AP and MP

$$8L - \frac{2}{3}L^2 = 16L - 2L^2$$

$$2L^2 - \frac{2}{3}L^2 - 16L + 8L = 0 \qquad \frac{4}{3}L^2 - 8L = 0$$

$$L(\frac{4}{3}L - 8) = 0 \qquad L = 0 \text{ or } L = 6$$

THEORY OF PRODUCTION

- Since output varies with variation in L, we are interested with positive level of labour employment.
  - Therefore, MP crosses AP at L = 6.
- For employment level less than L = 6, MP is greater than AP, therefore AP will rise. Consider for example L = 5.

$$MP = 16(5) - 2(5)^2 \qquad MP = 30$$

$$AP = 8(5) - \frac{2}{3}(5)^2 = 23.33$$

- For employment level greater than L = 6, MP is less than AP, therefore AP will fall. For example consider L = 7.

$$MP = 16(7) - 2(7)^2 \qquad MP = 14 \qquad AP = 8(7) - \frac{2}{3}(7)^2 = 23.33$$

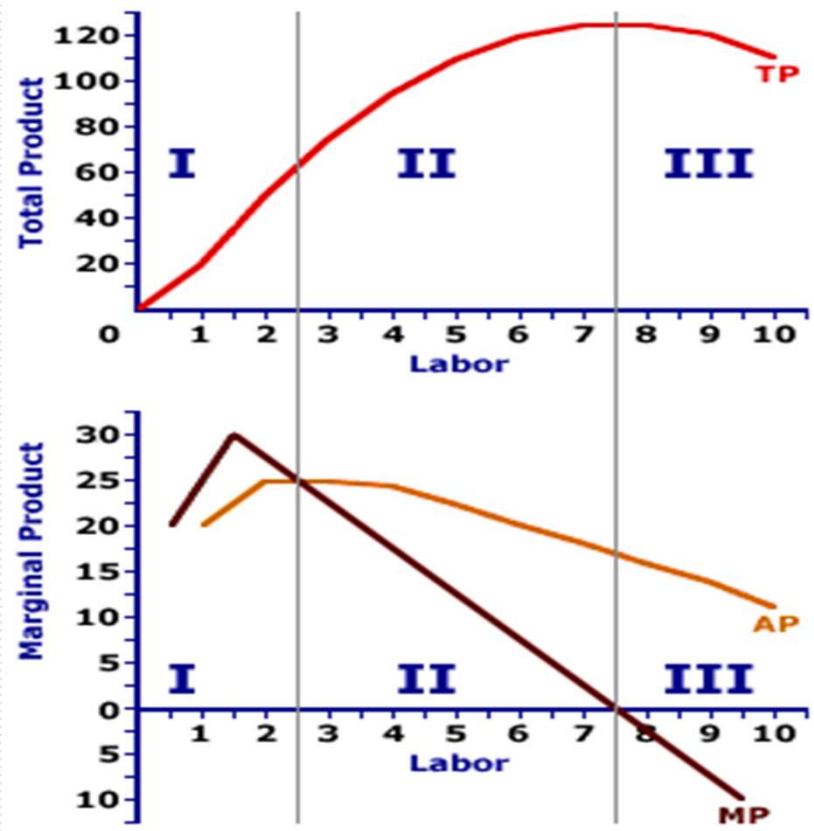
THEORY OF PRODUCTION

The Stages of Production

- We may split the whole production functions into three segments (or stages).
  - **Stage I** represents a range over which output grows at an increasing rate causing the MP to rise, therefore, there is no rationale for a firm to stop its employment over this range because an additional unit of labour produces more than the previous units.
  - **Stage I I** Starts where AP of labour inputs begins to decline. Out put still continues to increase but at a decreasing rate and in fact reach maximum.
  - **Stage III** is one on which output decreases, implying that employment of additional units reduces output, i.e. MP is negative.

THEORY OF PRODUCTION

The Stages of Production



- The basic theory of production usually concentrates on the range of output over which the MP of a variable factor (labour) decreases, but is positive; i.e.,
- the range of diminishing (but non-negative) productivity of the factor. See Stage II.
- Formally, the efficient stage of production is defined by the condition:

$$MP_L = \frac{\partial Q}{\partial L} > 0$$

$$\frac{\partial MP_L}{\partial L} = \frac{\partial^2 Q}{\partial L^2} < 0$$

MP of labour should be positive & slope o f MP is negative

THEORY OF PRODUCTION

The Long run Production Function

The Production Function

- Suppose that a production process involves just two factors labour (L) and capital (K).
- Production function is an equation, table, or graph showing the amount of output of a product that a firm can produce per period of time for each set of inputs.
- Both inputs and outputs are measured in physical rather than in monetary units.
- Technology is assumed to remain constant during the period of the analysis. Technology summarizes the feasible means (knowhow) of converting raw inputs, in to an output.

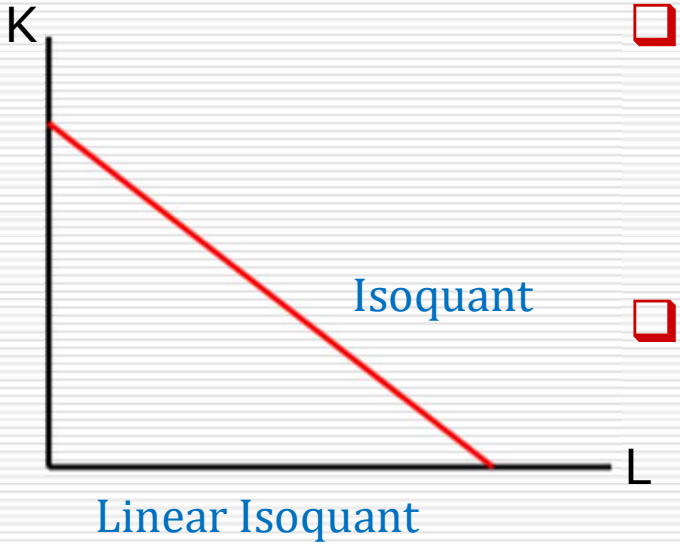
THEORY OF PRODUCTION

An isoquant

- **An isoquant** is the locus of all technically efficient methods (all combinations of factors of production) for producing a given level of output;
- i.e. it shows the different combinations of L and K that can be used to produce a given level of output.
- **The production isoquant** may assume various shapes depending on the **degree of substitutability** of factors.

THEORY OF PRODUCTION

- **Linear Isoquant-** assumes perfect substitutability of factors of production: a given commodity may be produced using only capital, or only labour, or by an infinite combination of K and L.
- This isoquant implicitly assumes that factors are perfectly divisible (See the figure below).



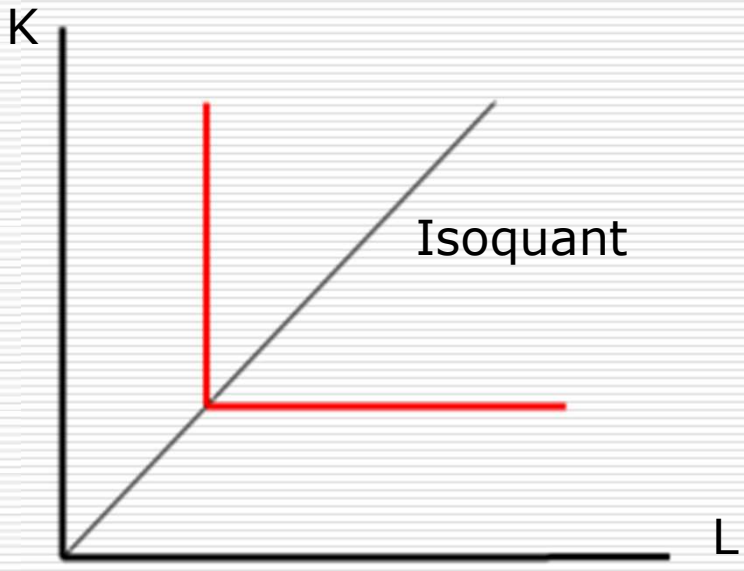
Linear Isoquant

- The assumption of perfect divisibility is, however, unrealistic. Most factors of production are not perfectly divisible.
- Tractor cannot be, for example, considered divisible. There is no such thing as 0.5 unit of tractor, 0.21 unit of tractor, etc.

THEORY OF PRODUCTION

- **Input-output Isoquant-** assumes strict complementarily (zero substitutability) between factors of production.
- There is only one efficient method of production for any given level of output, which is defined at the corner of the L-shaped isoquant.
- Though, operation anywhere on the vertical as well as the horizontal segments is possible, it involves larger quantity of one of the factors and not less of the other. This implies, therefore, irrational behaviour.

THEORY OF PRODUCTION

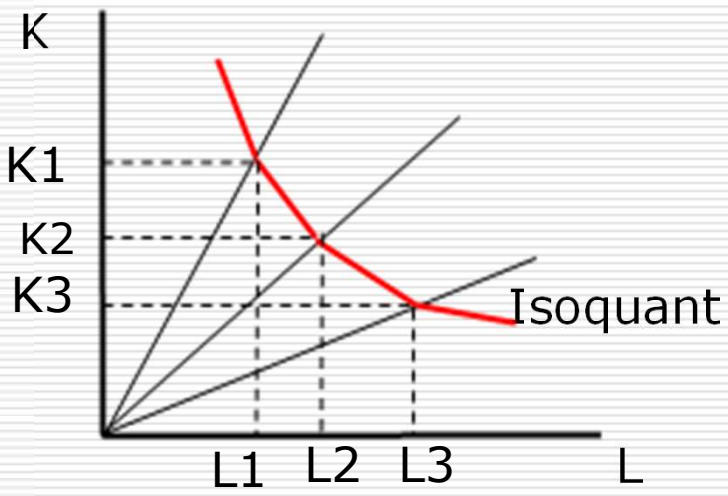


Input-Output Isoquant

- Along the vertical segment of the input-output isoquant (Figure 4.8 above), a firm uses more and more of capital but not less of labour to produce the same level of output.
- Along the horizontal segment, on the other hand, the firm uses more and more units of labour but not less of capital.

- It is not rational, therefore, for a firm to operate along the vertical or horizontal segments of the L-shaped isoquant.

THEORY OF PRODUCTION

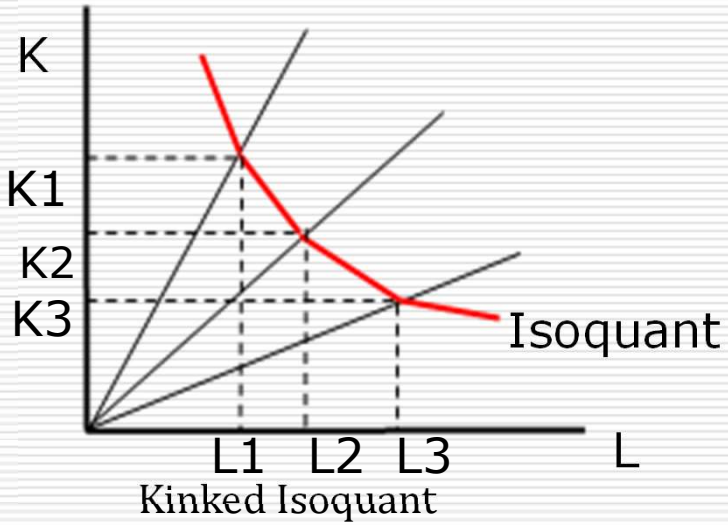


Kinked Isoquant

- **Kinked Isoquant-** assumes limited substitutability between factors of production L and K.
- There are a few processes for producing any one commodity. Substitutability of factors is possible only at the kinks.

- This isoquant assumes that factors are not perfectly divisible. Among the four widely known types of isoquants this one is the most realistic. In the real world, factors substitute each other only to a limited extent.

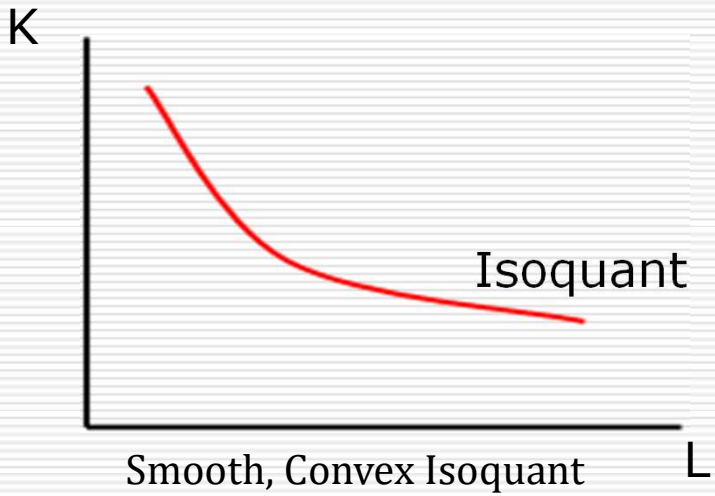
THEORY OF PRODUCTION



Kinked Isoquant

- Factor combinations (L1, K1) and (L2, K2) yield the same level of output since both are located on the same isoquant. This implies the effect of fall in capital from K1 to K2 could be offset by rise in labour from L1 to L2.
- Substitution between these two points is, however, impossible. This point is in harmony with the fact that some factors are not perfectly divisible. Continuous substitution of inputs requires that the factors involved be perfectly divisible.

THEORY OF PRODUCTION



Smooth, Convex Isoquant

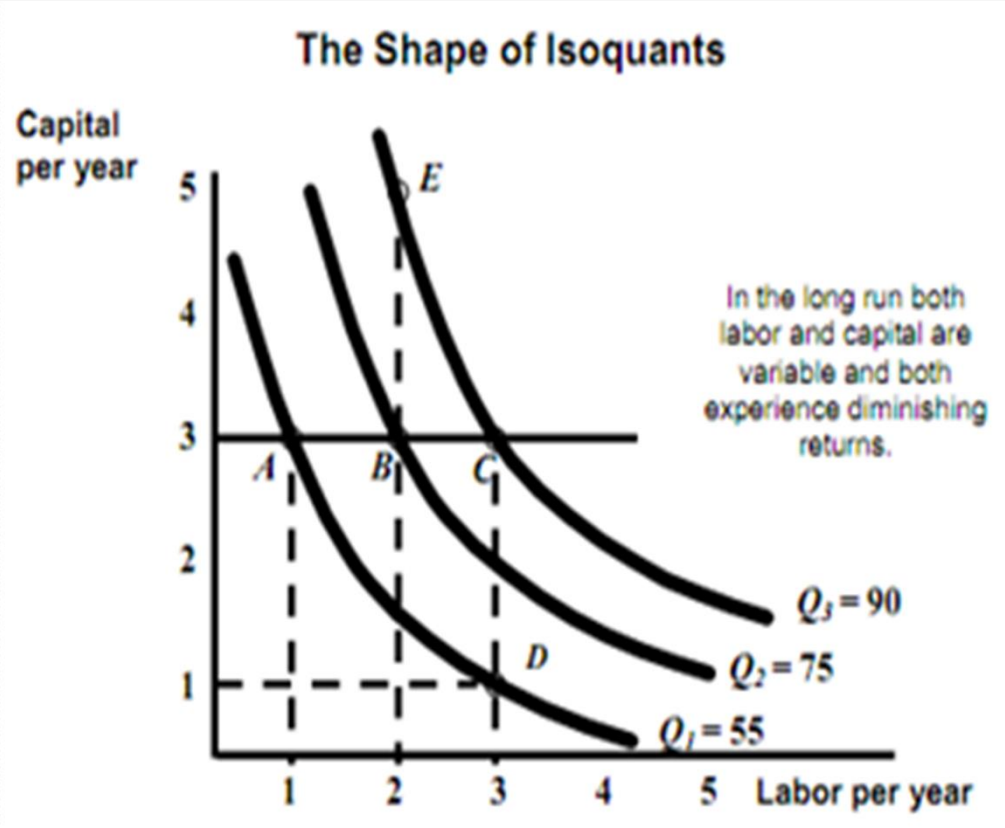
- Smooth, Convex Isoquant – assumes continuous substitutability of K and L only over a certain range, beyond which factors cannot substitute each other.
- The isoquant appears as a smooth curve convex to the origin. Substitution of one factor by the other is possible only on the convex segment of the isoquant.
- The smooth, convex isoquant is a limiting case of the kinked isoquant with infinite number of kinks.
- Traditional theory of production mostly adopted the continuous isoquants because; they are mathematically simple to handle by simple rules of calculus.

THEORY OF PRODUCTION

Production with two variable inputs

	Labor Input				
Capital Input	**1**	**2**	**3**	**4**	**5**
1	**20**	**40**	**55**	**65**	**75**
2	**40**	**60**	**75**	**85**	**90**
3	**55**	**75**	**90**	**100**	**105**
4	**65**	**85**	**100**	**110**	**115**
5	**75**	**90**	**105**	**115**	**120**

THEORY OF PRODUCTION



- Production isoquants show the various combinations of inputs necessary for the firm to produce a given output.
- A set of isoquants, or isoquant map, describes the firm's production function.
- Output increases as one moves from isoquant Q1 (55 units per year) to isoquant Q2 (75 units per year) and to isoquant Q3 (90 units per year).

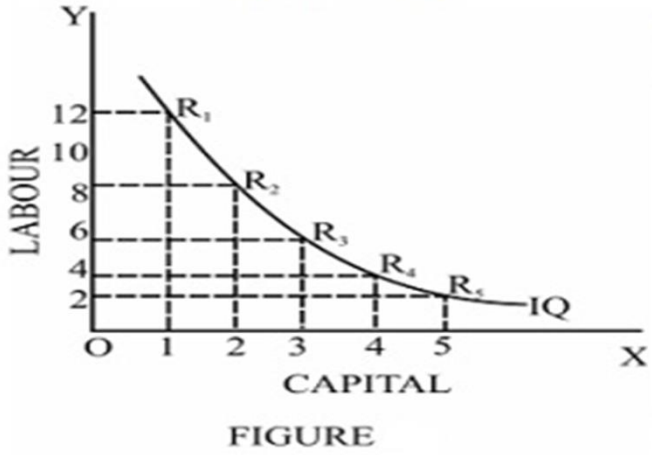
THEORY OF PRODUCTION

Characteristics of Isoquants

- Given a set of isoquants, the higher to the right an isoquant, the higher the level of output it represents.
  - Conversely, the lower an isoquant the lower the level of output it represents. This is because higher isoquants are associated with higher quantity of factors while lower isoquants are associated with lower quantity of factors.
- Isoquants do not intersect each other. The point at which two isoquants cross each other implies two different levels of output, which is not possible.

THEORY OF PRODUCTION

Isoquant Curve



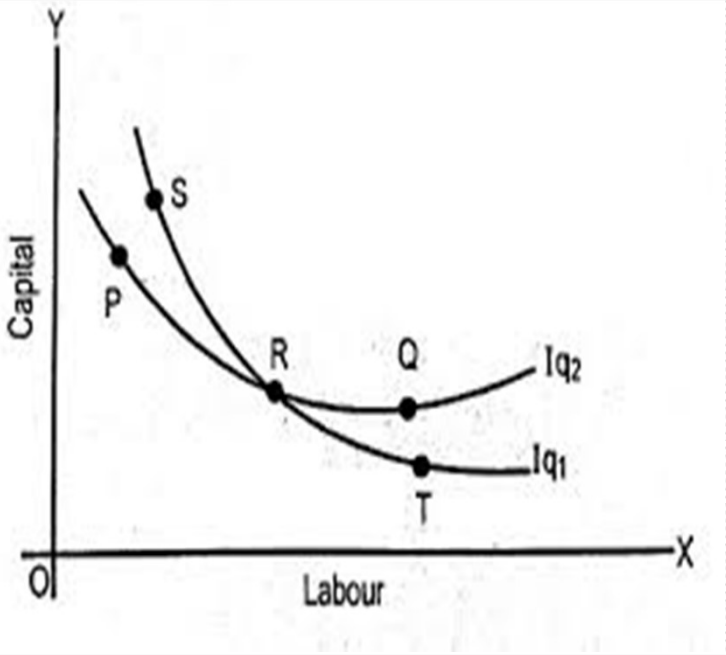
FIGURE

Table : Production with two variable inputs

Combination	Units of capital	Units of Labour	Output in units
A	1	12	1000
B	2	8	1000
C	3	5	1000
D	4	3	1000
E	5	2	1000

THEORY OF PRODUCTION

- Iisoquants Iq1 and Iq2 cross each other at point R.
  - Points P, R and Q are located on the same isoquant, therefore, represent the same level of output.
  - Similarly points S, R and T are located on the same isoquant, therefore represent the same level of output.
  - Point S is, however, located to the right of point P; and must represent a higher level of output.



By rule of transitivity, since S represents higher level of output than P and the same level of output as R, therefore, R must represent higher level of output than P. Yet, this is not true.

THEORY OF PRODUCTION

Laws of Production

- The laws of production describe the technically possible ways of increasing the level of production. Output may increase in various ways.
- Output can be increased by changing all factors of production. Clearly this is possible only in the long run. Thus the laws of returns to scale refer to the long-run analysis of production.
- In the short run output may be increased by using more of the variable factor(s), while capital (and possibly other factors as well) are kept constant.

THEORY OF PRODUCTION

- The marginal product of the variable factor(s) will decline eventually as more and more quantities of this factor are combined with the other constant factors.
- The expansion of output with one factor (at least) constant is described by the law of (eventually) diminishing returns of the variable factor, which is often referred to as the law of variable proportions.
  - *We will first examine the long-run laws of returns of scale.*

Laws of Returns to Scale: Long-Run Analysis of Production

- In the long run expansion of output may be achieved by varying all factors. In the long run all factors are variable.

THEORY OF PRODUCTION

- The laws of returns to scale refer to the effects of scale relationships.
- In the long run output may be increased by changing all factors by the same proportion, or by different proportions.
- Traditional theory of production concentrates on the first case, that is, the study of output as all inputs change by the same proportion.
- The term 'returns to scale' refers to the changes in output as all factors change by the same proportion.‘
- Suppose we start from an initial level of inputs and output

$$X_0 = f(L, K)$$

THEORY OF PRODUCTION

- And we increase all the factors by the same proportion k. We will clearly obtain a new level of output X*, higher than the original level X 0 ,

$$X^* = f(kL, kK)$$

- If X* increases by the same proportion k as the inputs, we say that there are **constant returns to scale**.
- If X* increases less than proportionally with the increase in the factors, we have **decreasing returns to scale**.
- If X* increases more than proportionally with the increase in the factors, we have **increasing returns to scale.**

THEORY OF PRODUCTION

Returns to scale and homogeneity of the production function

- Suppose we increase both factors of the function

$$X_0 = f(L, K)$$

- by the same proportion k, and we observe the resulting new level of output X*

$$X^* = f(kL, kK)$$

- If k can be factored out (that is, may be taken out of the brackets as a common factor), then the new level of output X* can be expressed as a function of k (to any power v) and the initial level of output

$$X^* = k^v f(L, K) \text{ or } X^* = k^v X0$$

And the production function is called homogeneous

THEORY OF PRODUCTION

- If **k cannot be factored out**, the production function is **non-homogeneous.** Thus:
  - **A homogeneous function** is a function such that if **each of the inputs** is **multiplied by k**, then k can be completely factored out of the function.
  - **The power v of k** is called **the degree of homogeneity** of the function and is a measure of **the returns to scale**:
- If v = 1 we have **constant returns to scale.** This production function is sometimes called linear homogeneous.
- If v < 1 we have **decreasing returns to scale.**
- If v > 1 we have **increasing returns to scale.**

THEORY OF PRODUCTION

- **Returns to scale** are measured mathematically by the coefficients of the production function.
- For example, in a Cobb-Douglas function

$$X = b_0 L^{b1} K^{b2}$$

- The returns to scale are measured by the sum (b1 + b2) = v.

THEORY OF PRODUCTION

Graphical presentation of the returns to scale for a homogeneous production function

- The returns to scale may be shown graphically by the distance (on an isocline) between successive 'multiple-level-of-output' isoquants,
  - that is, **isoquants** that show **levels of output** which are multiples of some base level of output, e.g., X, 2X, 3X, etc.

THEORY OF PRODUCTION

Constant returns to scale.

- Along any isocline the distance between successive multiple isoquants is constant. Doubling the factor inputs achieves double the level of the initial output; trebling inputs achieves treble output, and so on

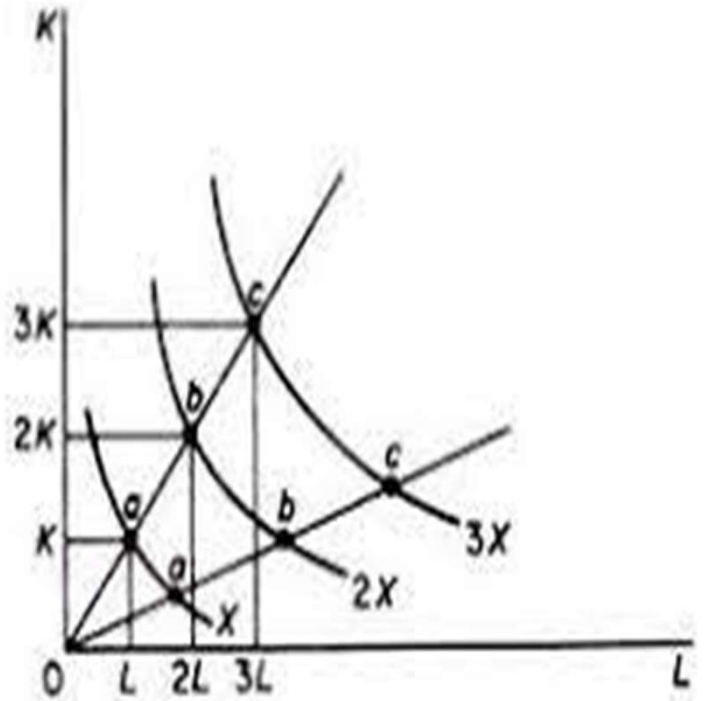


Figure 3.18 Constant returns to scale: $Oa = ab = bc$

THEORY OF PRODUCTION

Decreasing returns to scale.

- The distance between consecutive multiple-isoquants increases. By doubling the inputs, output increases by less than twice its original level. In figure 3.19 the point a', defined by 2K and 2L, lies on an isoquant below the one showing2X.

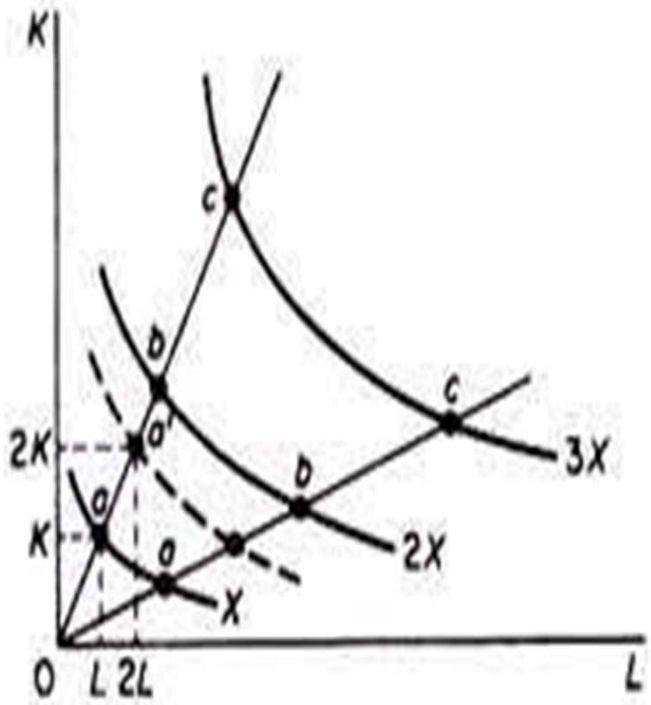


Figure 3.19 Decreasing returns to scale: $0a < ab < bc$

THEORY OF PRODUCTION

Increasing returns to scale.

- The distance between consecutive multiple isoquants decreases.
- By doubling the inputs, output is more than doubled.
- In figure 3.20 doubling K and L leads to point b' which lies on an isoquant above the one denoting 2X.

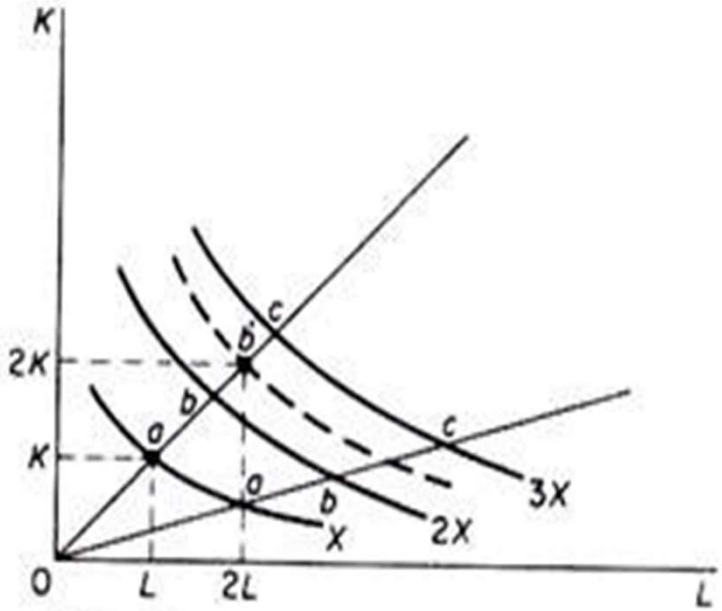


Figure 3.20 Increasing returns to scale: $0a > ab > bc$

THEORY OF PRODUCTION

The Law of Variable Proportions: Short-run Analysis of Production

- We said above that if one factor is variable while the other(s) is kept constant, the product line will be a straight line parallel to the axis of the variable factor
- In general if one of the factors of production (usually capital K) is fixed, the marginal product of the variable factor (labour) will diminish after a certain range of production.
- We said that the traditional theory of production concentrates on the ranges of output over which the marginal products of the factors are positive but diminishing.

THEORY OF PRODUCTION

- The ranges of increasing returns (to a factor) and the range of negative productivity are not equilibrium ranges of output.
  - If the production function is homogeneous with constant or decreasing returns to scale everywhere on the production surface, the productivity of the variable factor will necessarily be diminishing.
  - If, however, the production function exhibits increasing returns to scale, the diminishing returns arising from the decreasing marginal product of the variable factor (labour) may be offset, if the returns to scale are considerable.
  - This, however, is rare. In general the productivity of a single-variable factor (ceteris paribus) is diminishing.
  - Let us examine the law of variable proportions or the law of diminishing productivity (returns) in some detail.

THEORY OF PRODUCTION

- If the production function is homogeneous with constant returns to scale everywhere, the returns to a single-variable factor will be diminishing.
- This is implied by the negative slope and the convexity of the isoquants. With constant returns to scale everywhere on the production surface,

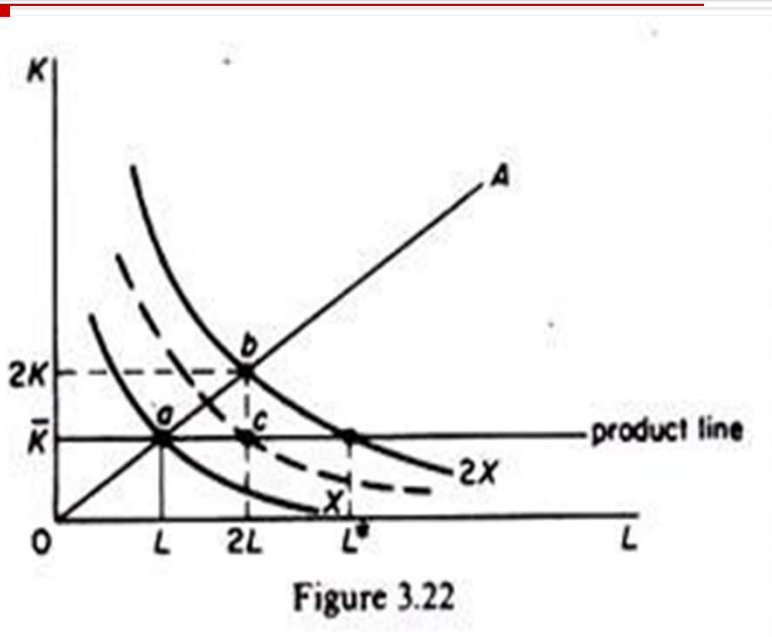


Figure 3.22

doubling both factors (2K, 2L) leads to a doubling of output. In figure 3.22 point b on the isocline OA lies on the isoquant 2X. However, if we keep K constant (at the level K) and we double only the amount of L, we reach point c, which clearly lies on a lower isoquant than 2X

THEORY OF PRODUCTION

- In figure 3.23 we see that with 2L and 2K output reaches the level d which is on a lower isoquant than 2X.
- If we double only labour while keeping capital constant, output reaches the level c, which lies on a still lower isoquant.

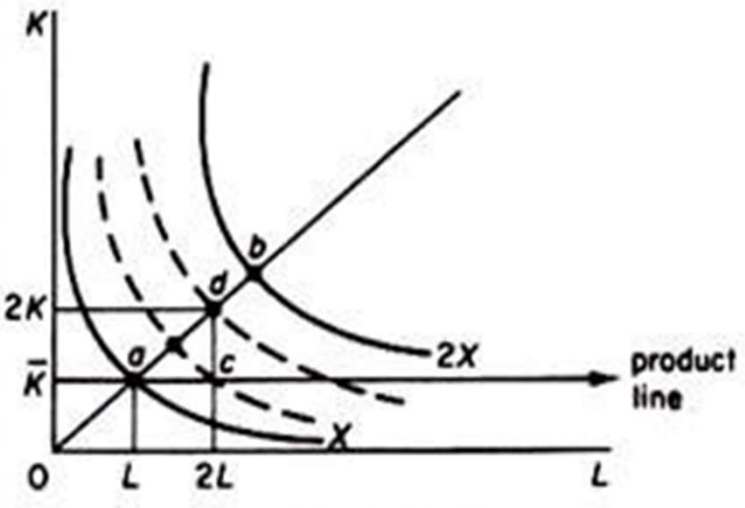


Figure 3.23

THEORY OF PRODUCTION

- If the production function shows increasing returns to scale, the returns to the single- variable factor L will in general be diminishing (figure 3.24), unless the positive returns to scale are so strong as to offset the diminishing marginal productivity of the single- variable factor.
- Figure 3.25 shows the rare case of strong returns to scale which offset the diminishing productivity of L.

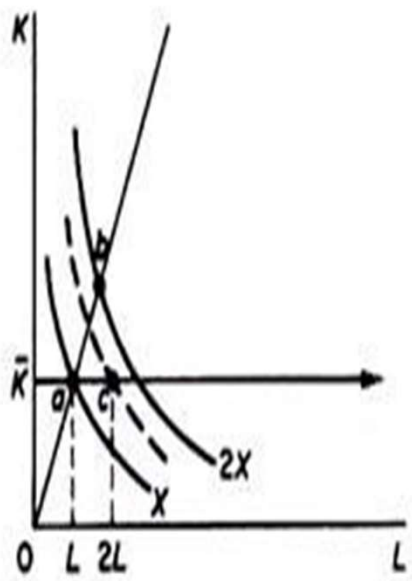


Figure 3.24

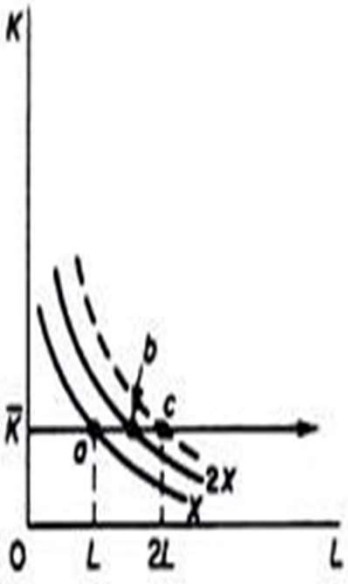


Figure 3.25

THEORY OF PRODUCTION

Equilibrium of the Firm: Choice of Optimal Combination of Factors of Production

- it is assumed that the firm can choose the optimal combination of factors, that it can employ any amount of any factor in order to maximise its profits.
  - This assumption is valid if the firm is new, or if the firm is in the long-run. However, an existing firm may be coerced, due to pressure of demand, to expand its output in the short-run, when at least one factor, usually capital, is constant.
- we make the following assumptions:

I. The goal of the firm is profit maximisation-that is, the maximisation of the difference II = R - C where

THEORY OF PRODUCTION

II. The price of output is given, P".

III. The prices of factors are given:

- w is the given wage rate
- r' is the given price of capital services (rental price of machinery).

Single Decision of the Firm

- The problem facing the firm is that of a constrained profit maximisation, which may take one of the following forms:
- (a) Maximise profit II, subject to a cost constraint. In this case total cost and prices are given (C, w, r, P "), and the problem may be stated as follows

Max ii=R-C II= P"X- C

THEORY OF PRODUCTION

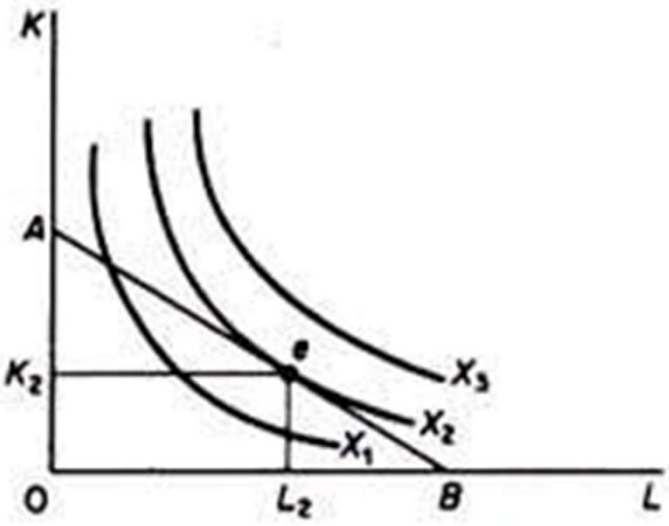


Figure 3.34

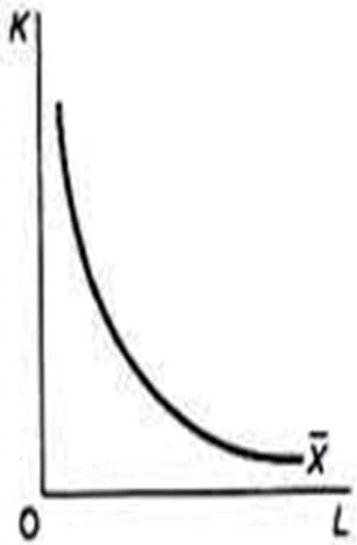


Figure 3.36

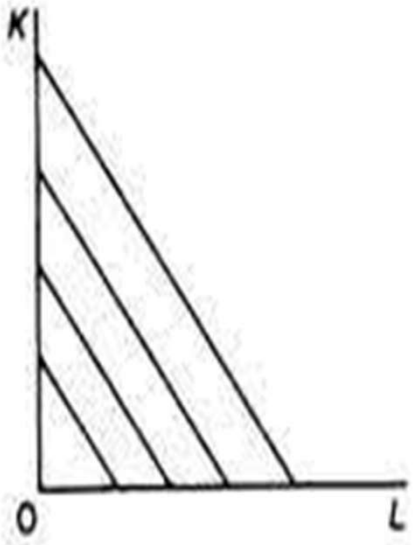


Figure 3.37

THEORY OF PRODUCTION

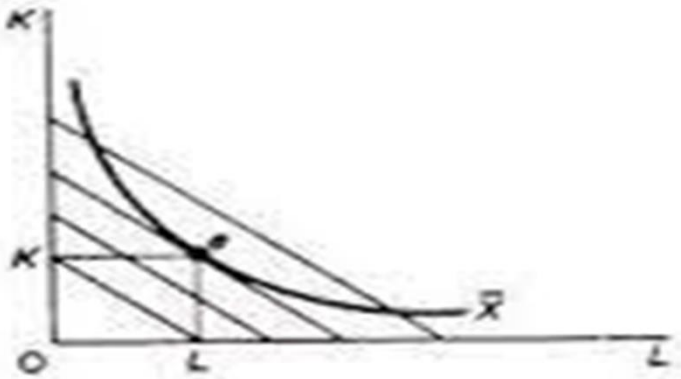


Figure 3.38

Formally:
Minimise $C = f(X) = wL + rK$

subject to $\bar{X} = f(L, K)$

Rewrite the constraint in the form

$\bar{X} - f(L, K) = 0$