SELF LEARNING MATERIAL

M.A ECONOMICS

COURSE : ECO - 101

(1st Semester)

MICROECONOMIC THEORY

BLOCK - I & II

Directorate of Distance Education
DIBRUGARH UNIVERSITY
DIBRUGARH - 786 004

ECONOMICS

COURSE : ECO - 101

MICROECONOMIC THEORY

Contributors :

Prof. J. K. Gogoi
P. P. Buragohain
J. Borgohain
Department of Economics
Dibrugarh University

Editor :

Prof. K. C. Borah
Department of Economics
Dibrugarh University

Published on behalf of the Directorate of Distance Education, Dibrugarh University by the Director, DDE, D.U. and printed at Maliyata Offset Press, Mirza, Kamrup (Assam)

MICROECONOMIC THEORY

BLOCK - I

BASIC CONCEPTS OF MICROECONOMICS AND THE THEORY OF DEMAND

BLOCK - II

THEORY OF PRODUCTION AND COSTS

UNIT 1: BASIC CONCEPTS

Structure

1.0 Objectives

The objective of this unit is to introduce the learner with some basic concepts of economics. After going through the unit you will be able to:

1. Learn types of microeconomic analysis and importance of models in economics.
2. Concept and Importance of Firm and Industry.
3. Criteria for Classification of Firms into Industries and markets and
4. Explain the concept of partial and general equilibrium.

1.1 Introduction

In Block-I we shall discuss some of the basic concepts of Microeconomics and the Theory of Demand. The terms 'static', 'dynamic' and 'comparative static' have become very popular in modern economic analysis. Any type of theory that is 'dynamic' enjoys a reputation for being more realistic and more complete. Although it has been well recognized that 'static', 'dynamic' and 'comparative static' are three distinct approaches to the study of economics, yet the distinction between them could not be drawn very clearly and controversy persists. Despite that it is essential to learn the preliminary ideas of these basic concepts. Similarly, 'Economic Model' which is a simplified representation of a real situation has gained popularity among the economists will also be discussed.

In this Block we shall also discuss the concepts and importance of firm and industry, criteria for the classification of firms and industries as well as criteria for the classification of markets which will help the students for better understanding of various microeconomic theories at a later stage. Various 'equilibrium' analyses in economics can be divided into two categories, viz., partial equilibrium and general equilibrium, which are also included in this Block.

1.2 The Basic Concepts

There are three techniques of microeconomic analysis - micro static, micro dynamics and comparative micro static. Now let us explain the essential features of each of these techniques of economic analysis.

1.2.1 Micro Static:

Simple micro static analysis studies a set of microeconomic variables and the interrelations between them when they are in equilibrium at a given point of time. It cannot explain how the equilibrium has been brought about. It also implies that the state of equilibrium holds constant through time. The concept of simple micro static can be explained with the help of the following diagram where the equilibrium price of commodity X has been determined through the interplay of the demand and supply forces.

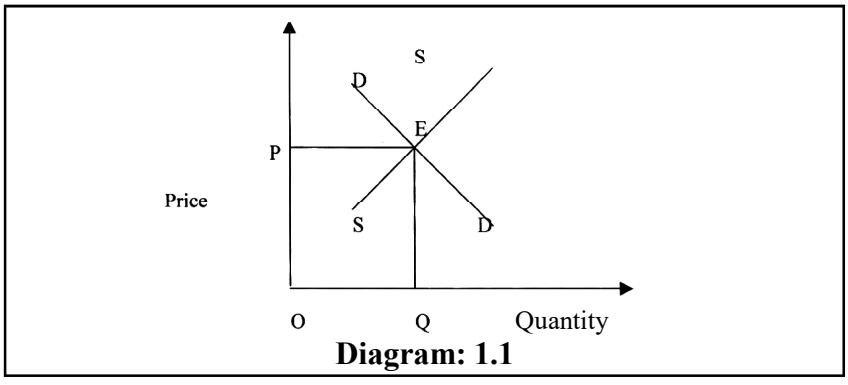


Diagram: 1.1

In the above Diagram 1.1, equilibrium is reached at point E where both the demand and the supply curve intersect with each other indicating the fact that the total demand for commodity X (OQ amount) is equal to total supply of the commodity X (OQ amount) at a particular price, viz., OP. Until and unless the demand and supply functions are remaining the same, there will be no change in the equilibrium condition. Thus the equilibrium price of the commodity will be established at OP price and the total demand for and supply of commodity X will be equal to OQ.

1.2.2 Comparative Static

In the simple micro static we assume that the demand and the supply functions do not change through time. However, when either or the two or both the functions shift then the market will, after the process of adjustment has worked itself out, reach a new state of equilibrium giving a new set of price and the amounts demanded and supplied. Comparative micro static analysis studies the relationship between the positions of ex-ante and ex-post equilibrium as shown in the Diagram 1.2.

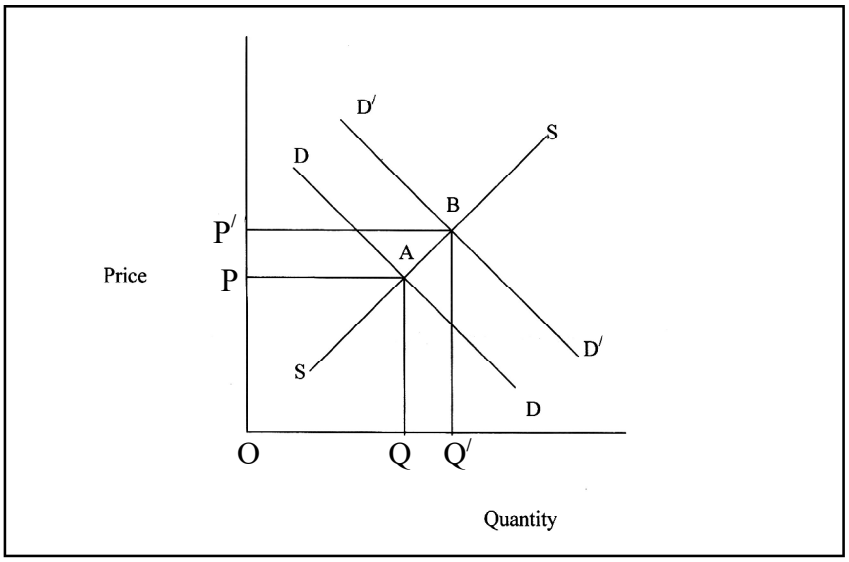


Diagram: 1.2

In the above diagram it is seen that originally at point A both the demand curve and the supply curve intersect and thereby setting equilibrium price at OP and output OQ. Now due to shift of demand curve from DD position to D/D/, which intersects with the supply curve SS at point B and thereby setting the new price and output respectively at OP/ and OQ/. The analysis, which compares the A and B states of equilibrium, is called the comparative analysis.

1.2.3 Micro Dynamic :

The micro dynamics is concerned with the study of the market in a state of disequilibria or with the path of the movement from one equilibrium position to another equilibrium position. It studies the happenings in the market during the period of transition from one equilibrium to the other. For example, when the demand function shifts upward, the original state of equilibrium is disturbed and the commodity market is plunged in to a sea of disturbances before attaining the state of new equilibrium. It is explained with the help of the following diagram:

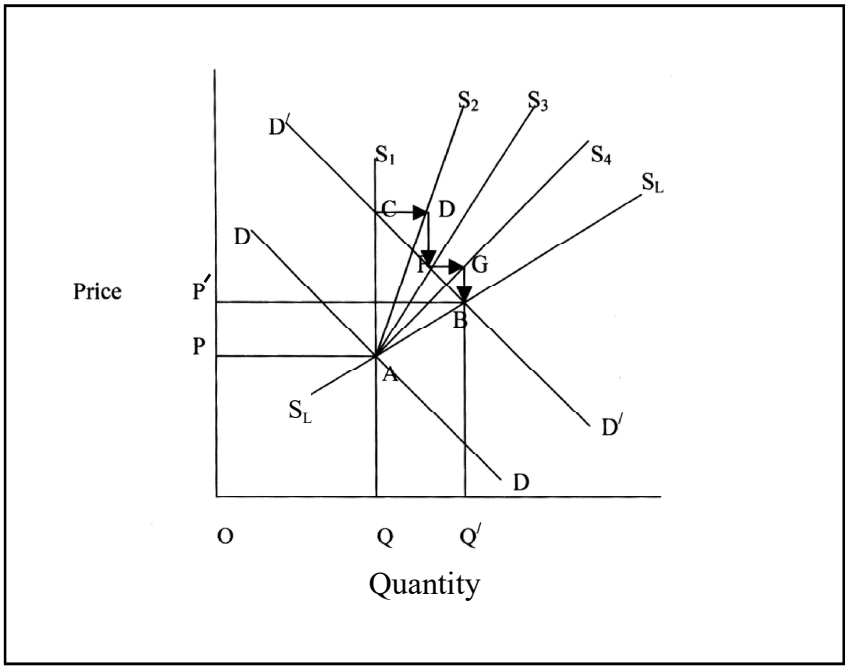


Diagram: 1.3

In the above diagram 1.3, S_LS_L is the long run supply curve and DD is the demand curve. Initially the equilibrium is at point A where both the demand and supply curve intersect each other and the equilibrium price is OP. Now suppose the demand curve shifts from DD to $D'D'$ and as a result of this immediately price shoots up from OP to OP'. Since the supply in the short run is inelastic the seller will enjoy a pure surplus of PP' per unit of commodity. Consequently, they endeavor to increase their output with the given scale of plant by working longer hour. This shifts the supply curve from S_1 to S_2 the process of expansion of output will be continued until the long run equilibrium price is established at OP' where both the demand curve $D'D'$ and the long run supply curve S_LS_L intersect with each other. Thus the micro dynamic is concerned with the study of the path of movement from one equilibrium to other equilibrium over a period of time.

1.3 Economic Models:

In economics, models are built to express the relationship between different economic variables. An economic model shows the relationship among economic variables in a precise manner. In other words, the main objective of building an economic model is to present the complex reality in to a simplified form.

An economic model is a simplified version of reality. It includes the main features of the real situation, which it represents. A model implies abstraction from reality, which is achieved by a set of meaningful and consistent assumptions, which aims at the simplification of the phenomenon or behavioural pattern that the model is designed to study. The degree of abstraction from reality depends on the purpose for which the model is constructed. A model can be built in several ways such as: in the form of diagram, flow charts, statistical tables, or a system of equations. Among these the most popular and widely used one is building mathematical model.

Model building usually consists of two important stages. The first stage is to develop the structure of the model. In the second stage, actual strength of the relationship being postulated is estimated using econometrics

For a long time the economists have been using economic models for describing, analyzing and predicting various economic concepts and events. An economic model constructed by incorporating two or more variables: (a) describes the relationship between the variables, (b) depicts the economic outcome of their relationship and (c) predicts the effects of changes in the variables on the economic outcome.

Economic models are used for several purposes. The following are the some of the important uses of the economic models:

(i) to work out the theoretical implications of economic analysis,

(ii) to test economic theories,

(iii) to make economic forecast,

(iv) to plan for the economic development of the country,

(v) to examine the alternate economic policies,

(vi) to stimulate alternative scenarios of economy, and

(vii) to replay economic history.

Now the question is about the validity of an economic model. The validity of a model may be judged on the basis of certain criteria. Its predictive power, the consistency and realism of its

assumption, the extent of information it provides, its generality (that is, the range of cases to which it applies) and its simplicity. But there is no general agreement regarding the most important criterion. To Milton Friedman, the most important criterion of the validity of the model is its predictive power, while to Paul Samuelson, realism of assumptions and power of the model in explaining the behaviour of the economic agents, producers or consumers, is the most important attribute of a model. But most of the economists are of the view that the most important attribute of a model depends on its purpose to which it is put. Predictive power of the model is important when the objective of the model is forecasting. On the other hand, if purpose of the model is to explain why a system behave as it does, realism of assumptions and explanatory power is the most important criterion for the validity of a model. However, an ideal model should satisfy both the criteria.

1.4 Concepts and importance of industry :

The concept of an industry is important for economic analysis. It is also of essential to the businessman, government, to those involved in the collection and processing of economic data and to all researchers. In economics concept of industry is very important because: (i) it reduces the complex relationships of all firms of an economy to a manageable dimensions. (ii) the concept of industry makes it possible to derive a set of general rules from which we can predict the behaviour of the competing members of the group that constitute the industry, (iii) the concept of industry provides the framework for the analysis of the effects of entry on the behaviour of the firm and on the equilibrium price and output.

1.5 Criteria for the Classification of Firms into Industries:

Two criteria are used for the definition of an industry: the market criterion and technology criterion. According to the market criterion, firms are grouped into an industry if the products are close substitutes, while according to the second criterion firms are grouped in an industry on the basis of similarity of processes and/or of raw materials being used. Which classification is more meaningful depends on the market structure and on the purpose for which the classification is chosen.

1.6 Criteria for the Classification of Markets:

Three criteria have been used for the classification of the markets. These are: (a) the product substitutability criterion, (b) the interdependence criterion and (c) the entry criterion. The first criterion takes in to account the existence and closeness of substitutes. The second criterion takes in to account the reactions of the competitors, while the third criterion takes in to account the ease of entry in various markets.

The substitutability of product can be measured with the help of cross price elasticity of demand. The cross price elasticity of demand measures the degree of responsiveness of quantity demanded for a product to a change in the price of related products. The cross price elasticity of demand for commodities produced by any two firm can be written as below:

$$e_{p.ji} = \frac{dq_j}{dp_i} . \frac{p_i}{q_j}$$

This measures the degree to which the sales of the jth firm are affected by changes in the price charged by the ith firm in the industry. The cross price elasticity of demand is positive for the substitute goods and negative for the complementary goods. In a perfectly competitive market, goods are perfectly substitute and hence the cross price elasticity of demand is infinity, while in monopoly form of market products are not substitutable as there is no rival firm and hence its elasticity is zero. On the other hand under monopolistic competition, sellers sell differentiated but closely substituted products, so the cross price elasticity of demand is greater than zero but less than infinity.

The second criterion is the interdependence criterion, which is closely related to the number of firms in the industry and the degree of product differentiation. It can be measured by the quantity cross elasticity of demand. Quantity cross elasticity of demand for products for any two firms can be expressed as:

$$e_{q.ji} = \frac{dp_j}{dq_i} . \frac{q_i}{p_j}$$

This measures the proportionate change in price of the jth firm resulting from an infinitesimal small change in the quantity produced by the ith firm. Under perfect competition, since the products are homogeneous and so the price output decision of the firms are not affected by the price output decision of the other firms. In case of monopoly also, there exists no interdependences between

the firms while despite the existence of the closely related products in monopolistic competition, each firm can take their own decision. However, under oligopoly each firm takes into account the actions and reactions of rival sellers.

The third criterion for the classification of markets is the entry criterion, which can be explained by the Bain's concept of barriers to entry. The condition of entry as per Bain is given by:

$$E = \frac{P_a - P_c}{P_c}$$

Where E stands for condition of entry, P_c is the price under perfect competition and P_a is the actual price charged by the firm. It is obvious that if the value of E is reduced to zero (i.e., there is no barrier to entry in to the market), the market is said to be either a perfectly competitive one or a monopolistically competitive one. Under monopoly, entry is blockaded.

The market classification which emerges from the application of the above three criteria is shown in the table below:

Classification of Markets

Type of market	Substitutability-of-product criterion $e_{p.ji} = \frac{dq_j}{dp_i} \cdot \frac{p_i}{q_j}$	Interdependence-of-sellers criterion $e_{q.ji} = \frac{dp_j}{dq_i} \cdot \frac{q_i}{p_j}$	Ease-of-entry criterion $E = \frac{P_a - P_c}{P_c}$
Perfect Competition	$\to \infty$	$\to 0$	$\to 0$
Monopolistic Competition	$0 < e_{p.ji} < \infty$	$\to 0$	$\to 0$
Pure Oligopoly	$\to \infty$	$0 < e_{q.ji} < \infty$	E >0
Heterogeneous Oligopoly	$0 < e_{p.ji} < \infty$	$0 < e_{q.ji} < \infty$	E >0
Monopoly	$\to 0$	$\to 0$	Blockaded entry

1.7 Equilibrium: Partial and General

Generally the term equilibrium may be of two kinds: the partial equilibrium and the general equilibrium. Let us discuss both the concept below:

1.7.1 Partial Equilibrium Analysis:

Partial equilibrium analysis is the study of equilibrium position of an individual, a firm, an industry or group of industries.

It is a market process for the determination of product prices and factor prices in which one or two variables are discussed, keeping all other constant. This analysis is based on the "ceteris paribus" assumption, i.e., other things remaining the same. Therefore the analysis is called the "ceteris paribus" analysis.

In short, partial equilibrium analysis is concerned with two types of economic problems. Firstly, it focuses the behaviour of individual segments like consumers, firms and market separately, rather than the whole economy. Secondly, it studies first the components of the system, and then tackles the broader question of how the parts fit together within the over-all economy.

1.7.2 General Equilibrium Analysis:

General equilibrium analysis also known as the macroeconomics is an extensive study of a number of economic variables, their interrelations and interdependences, for understanding the working of the economy as a whole. It brings together the cause and effect sequences of changes in the prices and quantities of commodities and services in relation to the entire economy. An economy can be in general equilibrium only if all consumers, all firms, all industries and all factor services are in equilibrium simultaneously and they are interlinked through commodity and prices. According to Stigler, "The theory of general equilibrium is the theory of interrelationship among all parts of the economy." General equilibrium analysis has two variants: one is the Walrasian model and other is the input-output analysis developed by W.W. Leontief.

To explain the partial and general equilibrium, let us take the example of product market and money market equilibrium.

Partial equilibrium analysis consider equilibrium in only one of the two segments of the economy, viz, the goods segments and money segment, without specifying whether the other segments of the economy will be in equilibrium or not.

The general equilibrium analysis, on the other hand explains equilibrium in both the markets (1) product market and (2) money market .The following two conditions must be fulfilled if the economy is to be in equilibrium:

(i) saving must be equal to investment

(ii) the demand for money is in equilibrium with the supply of money.

Now we are going to discuss the graphical representation of the equilibrium in product market and money market separately.

Product Market (Derivation of IS curve):

The product market is in equilibrium when saving and investment are equal or aggregate demand for goods equal aggregate supply .The investment demand can be represented as a decreasing function of the rate of interest. Suppose the total money supply is Rs. 125 crore and the investment function is given by:

$I = 125 - 25i$ (1)

When the rater of interest is 0,1,2,3 and 4 then the investment will be as follows:

When i = 0, I = 125 - 25X0 = 125 - 0 = 125 crore

When i = 1, I = 125 - 25X1 = 100 Crore

When i = 2, I = 125 - 25X 2 = 75 Crore

When i = 3, I = 125 - 25X3 = 50 Crore

When i = 4, I = 125 - 25X4 = 25 Crore

We plot this function in the first quadrant of the diagram - 1. Again, according to the classical economists, saving equal to the investment, that is, I = S ..... (2)

So, if I= 100, S= 100

I= 75, S=75

I= 50, S= 50

I= 25, S= 25

The saving investment equilibrium condition is represented in quadrant II of the diagram. This curve is a straight line rising from the origin at 45^0 angle indicating the equality between saving and investment.

The saving-schedule is plotted with respect to the level of income in quadrant III of the diagram 1. Let the saving function be:

$S = - 50 + .5Y$ ..... (3)

If Y = 50, S = -50 + .5X50 = -25

If Y = 100, S = -50 + .5X 100 = 0

If Y = 150, S = -50 +.5X 150 = 25

If Y = 200, S = -50 + .5X 200 = 50

If Y = 250, S = -50 + .5X 250 = 75

If Y = 300, S = -50 + .5X 300 = 100

Thus we have seen that there is a direct relationship between income and saving. If the level of income increases, saving also increases.

In our diagram the rate of interest is measured on the vertical axes and the level of investment is on the horizontal axes.

Starting with an interest rate of 3%, we note in quadrant I, that investment of Rs. 50 crore will be undertaken. Moving to quadrant Ii, we observe that saving must also be Rs 50 crore. In quadrant III, saving function indicate that Rs. 50 crore will be saved at an income of Rs. 200 crore. Finally, in quadrant IV, we obtain a point of product market equilibrium, i.e., when the rate of interest is 3 % the level of income that will make investment equal to saving is Rs. 200 cr.

If the interest rate is raised to 4%, the level of investment decreases to Rs. 25 cr. If the interest rate falls to 2%, investment will rise to Rs. 75 crore and the equilibrium income level is Rs. 250 crore. If this procedure of selecting arbitrary rate of interest and then finding the level of incomethat is consistent with each interest rate continued, a curve called IS curve will be traced out in quadrant IV. This graph is a simple graphical relationship of the product market equilibrium condition and it shows the level of intended investment and saving at different interest rates.

Product Market Equilibrium
(Derivation of IS Curve)

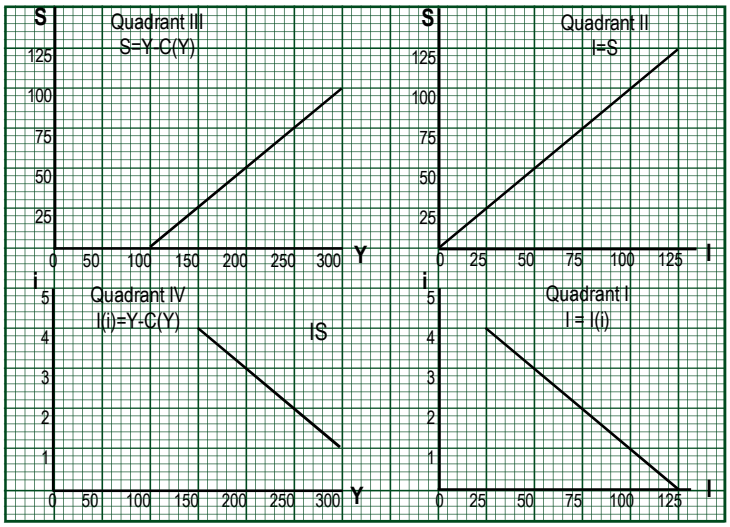


Diagram-1.4

The Money Market (Derivation of LM Curve):

Equilibrium in the money market implies equality between the demand for and supply of money. If the demand for money is greater than the supply of money, the rate of interest has a tendency to increase under the pressure of increased demand for bonds and the increased preference for the speculative motive and vice versa.

The same procedure as used in the product market may be applied here to find out the monetary equilibrium. Let us suppose the speculative demand for money is a function of :

$$M_2 = 100 - 25i \qquad (1)$$

If the rate of interest is respectively 0%, 1%, 2%, 3% and 4% then M_2 is respectively 100 crore, 75 crore, 50 crore, 25 crore and 0 crore.

In the quadrant I of diagram 2, the speculative demand for money is plotted against the rate of interest.

The total money supply (M_S) includes transaction and precautionary demand for money, i.e., M_1. let the total supply of money in the economy is Rs. 125 crore. In quadrant II, we show how the given money supply is split between transaction and precautionary demand for money and the speculative demand for money.

$$M_S = M_1 + M_2 \qquad .....(2)$$

$$\text{Or, } M_1 = M_S - M_2 \qquad .....(3)$$

Thus, when M_2 is respectively 0, 25, 50, 75 and 100, M1 will be respectively 125, 100, 75, 50 and 25 crore.

In quadrant III, the transaction demand is assumed to be proportional to the level of income in a ratio 1:2, and is plotted. The transaction demand for money is:

$$M_1 = 0.5\ Y \qquad .....(4)$$

When $Y = 50$, $M_1 = 0.5 \times 50 = 25$ crore.

Similarly when Y is respectively 100, 150, 200 and 250, M_1 is respectively 50, 75, 100 and 125 crore.

Finally, in quadrant IV, the interest rate that is consistent with monetary equilibrium is plotted against the level of income.

Beginning with an interest rate of 3% we have seen in quadrant I, the wealth holders desire to hold Rs. 25 crore of idle cash and deposits for speculative motive. In quadrant II, we observe that Rs. 100 crore will be released for transaction purposes. But in quadrant III, indicates that Rs. 100 crore of transaction money is consistent with a level of income Rs. 200 crore. Then if we move to quadrant IV, we see that the level of income that yields monetary equilibrium with a money supply of Rs. 125 crore and an interest rate of 3% is Rs. 200 crore.

Again, if the rate of interest is 2% the cash holder holds Rs. 50 crore for speculative motive realize Rs. 75 crore for other motives which is consistent with a level of income of Rs. 150 crore in quadrant III. Accordingly, we note in quadrant IV that the level of income that will yield monetary equilibrium at a interest rate of 2% is Rs. 150 crore.

If this procedure of selecting arbitrary rates of interest and finding the level of income that is consistent with monetary equilibrium at each interest rate is continued, a curve called the "LM" curve will be traced out in quadrant IV. This curve is a diagrammatical representation of monetary equilibrium condition and specifies the level of income which, for different rates of interest makes the demand for money equal to the supply of money. Since higher rates of interest are associated with a lower demand for idle balances, a given money supply can support a larger volume of transactions, and the LM curve therefore has a positive slope.

Money Market Equilibrium
(Derivation of LM Curve)

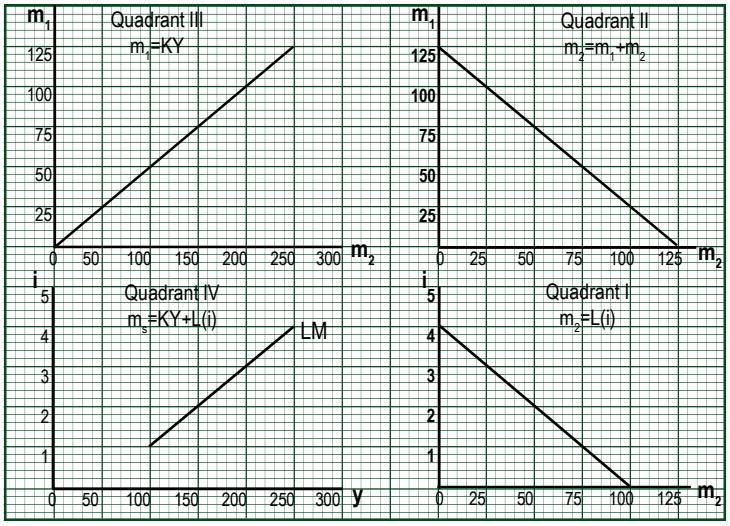


Diagram-1.5

Although the LM curve schedule suggests that several rates of interest are consistent with monetary equilibrium and the IS schedule suggests that several rates are consistent with the product market equilibrium, there is only one rate of interest and level of income that is consistent with the both. In diagram 3, the IS and LM functions are super imposed. The intersection of the curves at an income level Rs. 200 crore with an interest rate 3%. This is the point of equilibrium where general equilibrium sets in. once general

equilibrium prevails; there is no tendency for any magnitude to change.

Diagram 1.4 represents the product market equilibrium and diagram 1.5 represents money market equilibrium and both shows partial equilibrium analysis. The diagram 1.6 is a representation of general equilibrium analysis.

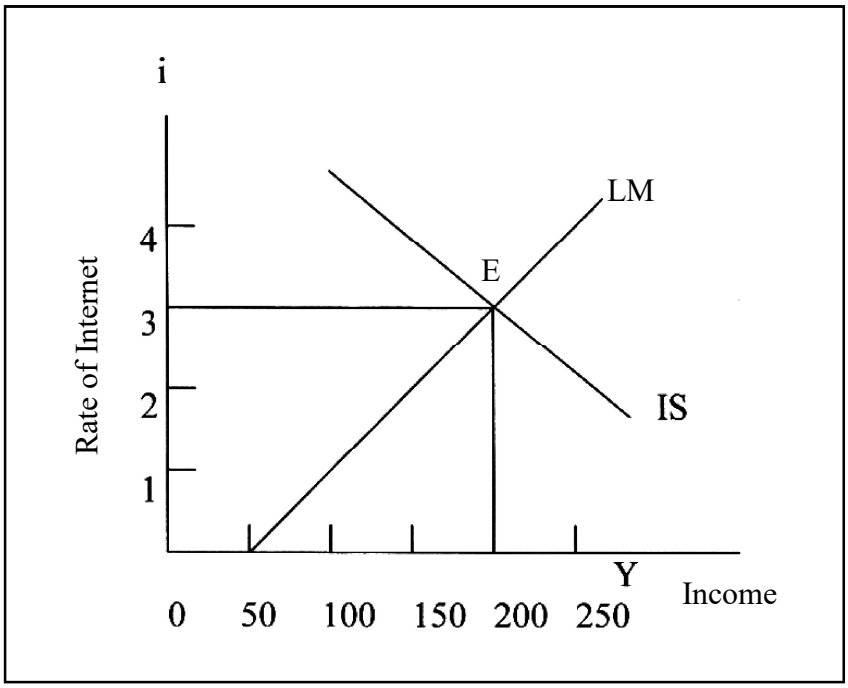


Diagram-1.6

Check Your Progress

1. Define: (a) micro static

.....

.....

.....

.....

.....

(b) Micro comparative static

.....

.....

.....

.....

.....

(c) Micro dynamic

.....

.....

.....

.....

.....

2. Fill up the blanks

(a) Economic models are simplified version of..... ……..……

(b) …….. criteria are used for the definition of an industry.

(c) The cross price elasticity of demand measures the..... ……..of quantity of demanded for a product to a change in the price of related products.

3. What are the criteria generally used for classification of markets?

.....

.....

.....

.....

.....

.....

.....

.....

.....

4. Distinguish between partial equilibrium and general equilibrium analysis.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

1.8 Let Us Sum Up

In this unit we discussed different types of microeconomics. Then we learn what is an economic model and what is the role of a model in economic analysis. We also learn from the unit about the criteria for classification of firms in to industries and markets. We also have an idea the importance of the concept of firms and industries. And lastly, we have learnt about partial and general equilibrium analysis.

1.9 Key Words

Micro Static: micro static analysis studies a set of microeconomic variables and the interrelations between them when they are in equilibrium at a given point of time.

Micro Comparative Static: Comparative micro static analysis studies the relationship between the positions of ex-ante and ex-post equilibrium

Micro Dynamic: Micro dynamics is concerned with the study of the market in a state of disequilibria or with the path of the movement from one equilibrium position to another equilibrium position.

Model: A model implies abstraction from reality, which is achieved by a set of meaningful and consistent assumptions, which aims at the simplification of the phenomenon or behavioural pattern that the model is designed to study.

Partial or particular equilibrium analysis: Partial or particular equilibrium analysis is the study of equilibrium position of an individual, a firm, an industry ar group of industries. It is a market process for the determination

General equilibrium analysis: General equilibrium analysis is an extensive study of a number of economic variables, their interrelations and interdependences, for understanding the working of the economy as a whole.

Suggested Readings:

1. Modern Microeconomics, A. Koutsoyiannis, Macmillan
2. Advanced Economic Theory, H.L Ahuja, S. Chand
3. Microeconomics : Theory and Application, Salvatore, Oxford.

UNIT 2: Theory of Demand:

Structure

2.0 Objectives:

The basic objective of this unit is to relate consumer's behaviour to demand analysis. As soon as you finish reading the unit it will be clear to you the following concepts:

(1) the concept of demand

(2) the cardinal utility and ordinal utility
(3) revealed preference theory
(4) consumers surplus
(5) consumer's choices under risky situations and
(6) recent development in demand analysis.

2.1 Introduction

The theory of consumer demand is an important concept of microeconomics. It is concerned with study of the exact nature of relationship between the given changes in any one or more of the demand determining variables and the induced change in the amount demanded of any given quantity. On the other hand utility analysis has also played an important role in the economic theory. How the consumer can attain maximum utility or what is the behaviour of a consumer to a change in the price of the goods, other things remaining the same, these are the goal of such analysis. This unit discusses the consumer's behaviour.

2.2 The Theory of Demand:

The theory of demand, as stated by Alfred Marshall, establishes a qualitative and functional relationship between price of a commodity and quantity demanded of it. It states that other things remaining same, a fall in the price of a commodity increases the quantity of the commodity demanded and a rise in the price causes a fall in its quantity demanded. Marshall puts the law as such: "The greater the amount to be sold, the smaller will be the price at which it is offered in order that it may find purchasers, or in other words, the amount demanded increases with a fall in prices and diminishes with a rise in price."

The law of demand can be explained by drawing demand curve of a commodity. The demand curve is a graphical representation of the demand schedule. On the other hand demand schedule shows the relationship between quantity demanded of a commodity and its price. Let us take the following hypothetical demand schedule to explain the demand law.

Hypothetical Market Demand Schedule

Price (in Rs.)	Quantity per period
5	10
4	15
3	20

The above table shows that the quantity demanded of a commodity is small at a high price and large at a low price. In other words the table explains the demand law. Now to explain the law we draw the following diagram :

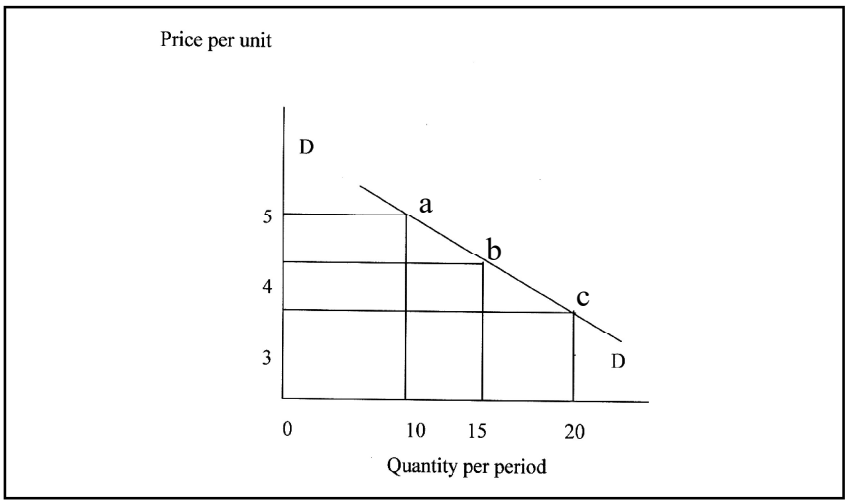


Diagram 2.1

In the above diagram, we measure on the horizontal axis the amount of quantity demanded and on the vertical axis, the price per unit. We consider three price quantity combination as indicated by the three points: a, b and c. By combining these three points we get a downward slopping curve and this is the demand curve that explains the statement of the demand curve. In the diagram we have seen that at price 5 per unit, the quantity demanded is 10 units. When the price is lowered to Rs 4 and Rs. 3 respectively, the amount of quantity demanded has been increased to 15 and 20 units respectively. Thus the downward slopping demand curve reflects that people wants to buy more at low price and vice versa. On the other hand a downward slopping demand curve means that the price change and quantity demanded moves in the opposite direction.

Exceptions to the theory of Demand :

There are some special circumstances where the law of demand does not hold good. For example:

(i) If the concerned good is a giffen good, the rational consumer will go on decreasing his consumption of the as the price of the good falls. This is because the nature of the giffen good is such that the consumer go on reducing consumption as the price of the good falls and vice versa.

(ii) Sometimes, the consumer used to judge the quality of goods by its price. Under such cases, when the price of the goods increases the consumer believe that the quality of the product has also improved

simultaneously. In fact, the consumer was being misguided. Such an effect is known as the Veblen effect.

(iii) Sometimes it may happen that when the price of a good is on a rise the consumer expects it to rise further. In such a case he may purchase more and more units of a good as its price goes on increasing.

2.3 Elasticities of Demand: Price, Income and Cross

The law of demand explains the relationship between price change and the quantity demanded. The elasticity of demand on the other hand quantifies such changes and gives a precise measure how consumer responds to the price change. Simply, elasticity of demand indicates the magnitude of change in consumer's demand for goods as a result of change in the price of the commodity. In general, elasticity of demand refers to the degree of responsiveness of quantity demanded of a commodity to a change in one of the variables affecting demand (i.e., to a change in any one of the demand determinants). On the basis of the demand influencing variables, elasticity of demand can be measured in three different ways and thus elasticity of demand can be of three types, viz., price elasticity of demand, income elasticity of demand and cross elasticity of demand.

2.3.1 Price Elasticity of Demand:

Price elasticity of demand is the most commonly used concept of elasticity of demand. It refers to proportionate change in quantity demanded of a commodity to a proportionate change in price of that particular commodity. Mathematically,

$$e_p = \frac{\frac{\partial q}{q}}{\frac{\partial p}{p}} = \frac{\partial q}{\partial p} x \frac{p}{q}$$

where e_p is the coefficient of price elasticity of demand, δ represents any change, q for quantity demanded and p for price.

Price elasticity of demand may be of unity, greater than unity, less than unity, zero or infinite. Price elasticity of demand is said to be unity when the change in demand is exactly equal to the change in price. For example, if 30% change in price causes a 30% change in the demand for the commodity, it will be a case of unit elasticity of demand.

When the change in demand is more than the change in the price of the commodity, then price elasticity of demand is greater than unity. Thus any change in the demand for a commodity above 30% change in price is an example of price elasticity of greater than one. It is also known as the relative elastic demand.

On the other hand price elasticity of demand is less than unity when say 30% change in price of a commodity causes a less than proportionate change (say 20%) in the quantity demanded. It is also known as the relatively inelastic demand.

Zero elasticity of demand is one when wherever the change in price, there is absolutely no change in demand. Price elasticity of demand is perfectly inelastic in this case. A 30% rise or fall in price leads to no change in the amount demanded.

When infinitesimal rise in the price of a commodity causes the quantity demanded to fall to zero; and an infinitesimal fall in price causes an infinite increase in the quantity demanded then the price elasticity of demand of the commodity is said to be infinite.

2.3.2 Income Elasticity of Demand:

The measure of responsiveness of demand to changes in income is called the income elasticity of demand. The income elasticity of demand for commodity X may be defined as the ratio of proportionate change in the quantity demanded to a given proportionate change in the income of consumers. Here the concept is based on the assumption that the prices of all the goods are given and that it is only the consumer's income, which changes. The formula for calculating income elasticity of demand is similar to that of the price elasticity of demand except that we substitute the given change in consumer's money income in place of the given change in the price of a commodity.

$$e_y = \frac{\frac{\partial q}{q}}{\frac{\partial y}{y}} = \frac{\partial q}{\partial y} x \frac{y}{q}$$

where e_y is the elasticity of income, δ represents any change, y is the income and q is the quantity demanded.

2.3.3 Cross Elasticity of Demand:

The cross elasticity of demand refers to the ratio of proportionate change in the quantity of a given commodity to a

given proportionate change in the price of some other related commodity. It provides a convenient basis for studying relationship between two different commodities. The formula for measuring coefficient of cross elasticity of demand is:

$$e_c = \frac{\frac{\partial q_x}{q_x}}{\frac{\partial p_y}{p_y}} = \frac{\partial q_x}{\partial p_y} x \frac{p_y}{q_x}$$

cross elasticity of demand may be positive or negative depending upon the relationship between the two commodities. If the two commodities are substitutable cross elasticity between them will be positive, while if it is complementary, the cross elasticity of demand is negative. Thus in case of substitute products any rise in the price of one commodity will cause an increase demand for the other. The higher the coefficient of cross elasticity of demand, more substitutable the commodity is. If the commodities are perfect substitutes, their elasticity is infinite. On the other hand in case of the complementary goods, a rise in the price of one commodity will cause a fall in the demand for the other commodity.

2.4 Elasticity of Supply:

Like elasticity of demand, elasticity of supply measures the degree of responsiveness to any change in the price of the commodity. Price elasticity of supply can be defined as the degree of responsiveness of the quantity supplied of commodity in response to a small percentage change in its own price. The co-efficient of price elasticity of supply can be expressed as:

$$E_s = \frac{\text{\% change in the quantity supplied of good X}}{\text{\% change in price of good X}}$$

This can also be expressed as:

$$e_s = \frac{\frac{\partial q}{q}}{\frac{\partial p}{p}} = \frac{\partial q}{\partial p} x \frac{p}{q}$$

where, e_s is the elasticity of supply, p for price and q for the quantity supplied.

The price elasticity of supply is always positive as supply increases with an increase in the price of that particular good (as indicated by the upward slopping supply curve).

Check Your Progress : I

1. Define (a) Elasticity of demand

.....

.....

.....

.....

.....

(b) Price elasticity of demand

.....

.....

.....

.....

.....

(c) Income elasticity of demand

.....

.....

.....

.....

.....

(d) Cross elasticity of demand

.....

.....

.....

.....

.....

2.5 Consumer Choice Concerning Utility

Under this section we will discuss about the cardinal utility and ordinal utility concept.

2.5.1 Cardinal Utility Theory:

There are two approaches to the concept of utility, viz, the cardinal utility and ordinal utility analysis. The first concept was given by Alfred Marshall and to that utility can be measured not only in terms of theory and practice but also in terms of money. Economists distinguish between cardinal utility and ordinal utility.

When cardinal utility is used, the magnitude of utility differences is treated as an ethically or behaviorally significant quantity. On the other hand, ordinal utility captures only ranking and not strength of preferences. An important example of a cardinal utility is the probability of achieving some target.

Utility functions of both sorts assign real numbers (utils) to members of a choice set. For example, suppose a cup of coffee has utility of 120 utils, a cup of tea has a utility of 80 utils, and a cup of water has a utility of 40 utils. When speaking of cardinal utility, it could be concluded that the cup of coffee is exactly the same amount better as a cup of tea as the cup of tea is better than the cup of water.

It is tempting when dealing with cardinal utility to aggregate utilities across persons. The argument against this is that interpersonal comparisons of utility are suspect because there is no good way to interpret how different people value consumption bundles.

When ordinal utilities are used, differences in utils are treated as ethically or behaviorally meaningless: the utility values assigned encode a full behavioral ordering between members of a choice set, but nothing about *strength of preferences*. In the above example, it would only be possible to say that coffee is preferred to tea to water, but no more.

Neoclassical economics has largely retreated from using cardinal utility functions as the basic objects of economic analysis, in favor of considering agent preferences over choice sets. As will be seen in subsequent sections, however, preference relations can often be rationalized as utility functions satisfying a variety of useful properties.

Cardinal utility functions are unique. Ordinal utility functions are equivalent up to monotone transformations, while cardinal utilities are equivalent up to positive linear transformations.

2.5.2 Ordinal Utility Analysis - Indifference Curve Analysis:

Indifference curve analysis is the analysis of consumer demand based on the notion of ordinal utility. The indifference curve analysis is developed as an alternative to the cardinal utility analysis was due to the pioneering works of utilitarian F. Y. Edgeworth (1881), G.B. Antoneli (1886) and Irving Fisher (1892). Later on it was J.R. Hicks and R.G.D. Allen (1934) in their article "A Reconsideration of the Theory of Value" presented a brief but clear and understandable idea to the indifference curve analysis. A detailed presentation of indifference curve analysis is found in the pioneering work of Hicks', "Value and Capital" published in 1939.

Definition of Indifference Curve :

The indifference curve analysis is based on ordinal utility analysis. It explains the consumer behavior in terms of his preferences or rankings for different combinations of two goods, say X and Y. An indifference curve is drawn from the indifference schedule of the consumer. An indifference curve shows the locus of points representing pairs of quantities between which the individual is indifferent. In fact, an indifference curve is an iso-utility curve showing equal satisfaction at various points of it. A single indifference curve represents only one level of satisfaction. An indifference curve further away from the origin represents the higher level of satisfaction. The concept of indifference curve can be explained with the help of the following imaginary indifference schedule representing various combinations of good X and Y.

Indifference Schedule

Combinations	X	Y
1	15	0
2	11	1
3	8	2
4	6	3
5	5	4

In the above schedule, the consumer obtains as much total satisfaction from the consumption of 11 units of X and 1 unit of Y and other combinations. One feels indifferent whether he gets first combination (15X + 0Y), the 2nd combination (11X + 1 Y), the third combination (8X + 2Y), the fourth combination (6X + 3Y) or the fifth combination (5X + 4Y). The total satisfaction is same in all the combinations. Now plotting the various combinations in a diagram and then by combining we will get a curve called as indifference curve.

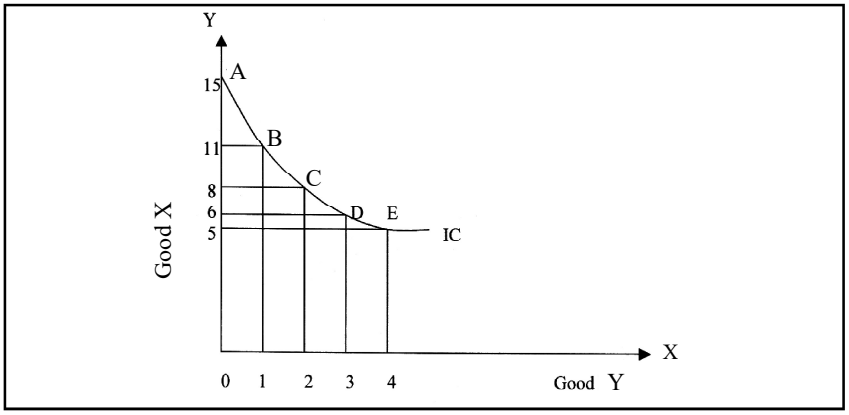


Diagram: 2.2

In the above diagram, the indifference curve IC is the locus of the points A, B, C, D and E, showing the combination of the two

goods X and Y between which the consumer is indifferent. The indifference curve possesses certain properties which are mentioned below:

(i) Indifference curve slopes downwards from left to the right. This is because when consumer decides to have more units of a good than the other, he will have to reduce the number of units of the other to stay in the same indifference curve, i.e., if the level of satisfaction is to remain same.

(ii) The second important property of indifference curve is that no two indifference curves intersect each other. Since we assume that each indifference curve represents a particular level of consumer satisfaction it will necessarily different from other indifference curve representing different level of consumer satisfaction. If two indifference curves intersect, it will mean that one satisfaction is at the same time greater or less than as equal to the other.

(iii) The third important property of indifference curve is that they are convex to the origin. This property follows from the principle of the diminishing marginal rate of substitution. The greater the number a good one acquires it mean that smaller will be the number of the unit of the other that will be substituted to maintain the given level of satisfaction.

(iv) Another important property of indifference curve is that they need not be parallel to each other. Firstly they are not based on the cardinal measurement of utility. Secondly the marginal rate of substitution between two goods may not be same in all the indifference curves. From this it follows that the indifference curves can be drawn either parallel to each other or otherwise.

(v) Lastly every consumer has a series of indifference curves for any pair of goods and each indifference curve situated to the right and away from the origin of the axes indicates progressively higher level of satisfaction. We assume that this preference pattern of the indifference curve is independent of the consumer's income.

Indifference Map :

All the indifference curves in the commodity space together constitute the indifference map of the consumer. It is the complete set of indifference curves. An indifference map located in the

positive quadrant of a graph indicates the consumer's preferences among all the combinations of goods and services. The farther away from an indifference curve from the origin, it means that more the combinations of goods along that curve is being preferred by the consumer.

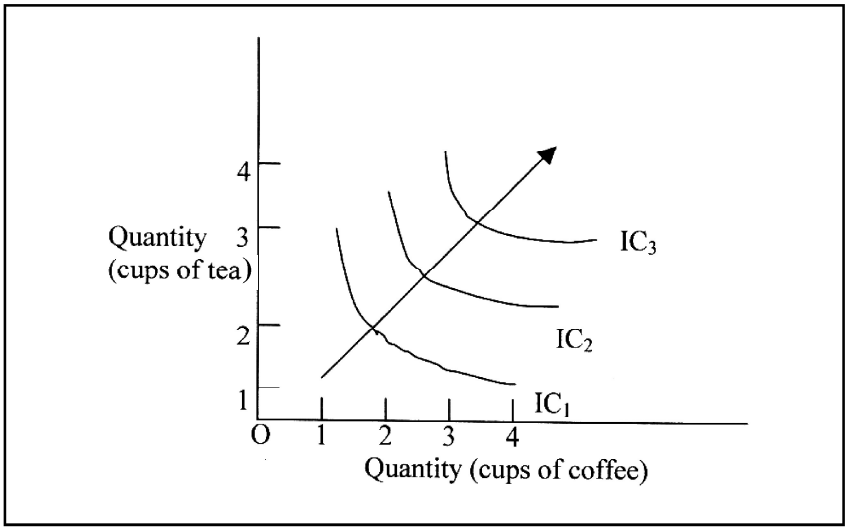


Diagram: 2.3

From the above diagram it is quite clear that IC_2 is preferred to IC_1 and IC_3 is preferred over IC_2. This indicates that more is preferred to less, that is, any point in the higher indifference curve is always better than any point on a lower indifference curve.

2.6 Price Effect: Hick's version

Price effect means the changes in the purchases of the good as a result of the fall in the price of good X, while keeping his money income, tastes and preferences of other good Y remaining the same. As a result of the price effect, the equilibrium position of the consumer would move to go equilibrium position at a higher indifference curve and would buy more the goods whose price is falling unless it is a giffen good. This price effect can be sprit up into two district forces- substitution effect and income effect.

There are two approaches for decomposing price effect is to parts: a substitution effect and an income effect. They are Hicksian approach and Slutsky approach. Now we are going further, Hicksian approach uses two methods of splitting the price effect, namely, 1.compensating variation in income and 2.equivalent variation in income. But, now we are devoted only with the decomposing price effect by 1.Hicks compensating variation is income method and 2. Stulsky method.

2.6.1 Income effect:

Income efect refers to the change in the purchases of a good caused by a given change in the money income of the consumer, other things remaining constant. If the income of the consumer increases his budget line will shift upward to the right, parallel to the original buget line. On the contrary, a fall in the income will shift the budget line inward to the left. But it is to be noted here that the budget lines are parallel to each other as the relative price remain unchanged. This is explained with the help of the following diagram:

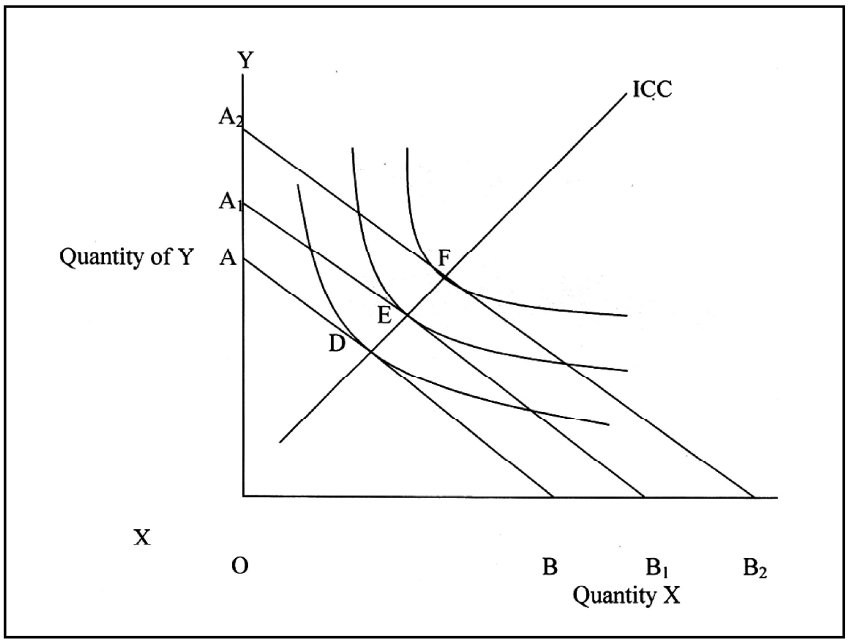


Diagram: 2.4

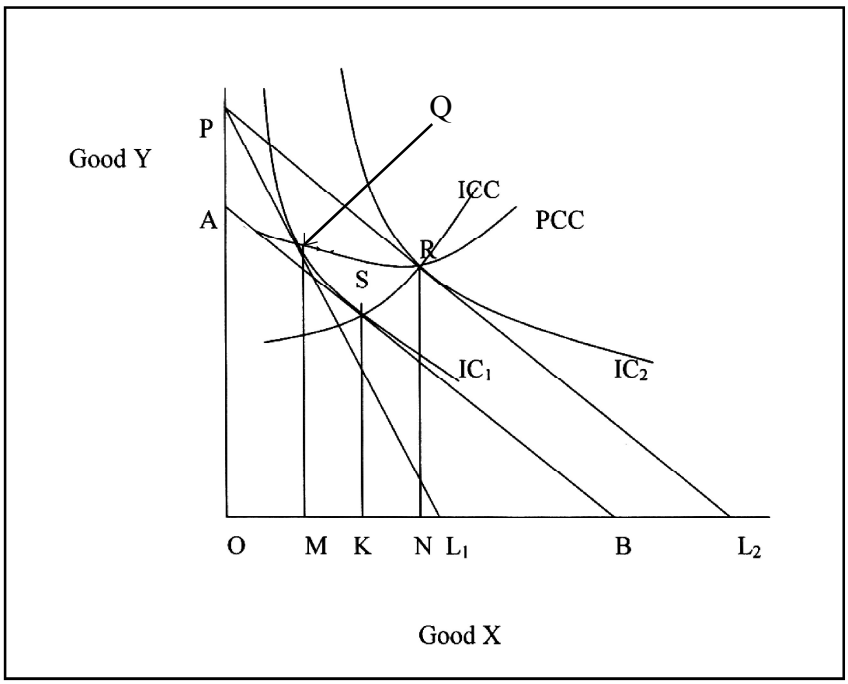
In the above diagram it is seen that as the income of the consumer has gone up, the price has also gone up. The price line AB, A_1B_1 and A_2B_2 indicate an increse in the money income of the consumer that enables him to move to the higher indifference curve and purchase larger quantity of the both goods. With higher price lines, the consumer can go over to higher indifference curve. By connecting the tangency points between the price lines and indifference curve we will get the income consumption line that shows the relationship between change in money income and the consumption of the two commodities, given the prices of the commodities remain constant. The income consumption curve indicates the income effect of a given change in the money income of the consumer. The income consumption curve may have any shape but it must not intersect one indifference curve more than once. However in case of inferior goods the income consumption curve is backward bending.

2.6.2 Substitution Effect:

The substitution effect relates to the change in the quantity demand resulting from a change in the price of a good due to the substitution of a relatively cheaper good for a dearer one, while keeping the price of the other good and real income and tastes of the consumer constant. That is , substituion effect means the change in the quantity of a good purchased which is due to the change in relative prices, real income remaining constant. When price of a good say X, falls, the real income of the consumer would increase. In order to find the change in the quantity of X purchased which is attritutable only to the change in relative price of X, the consumer's money income must be reduced by an amount so as to cancel out the gain in real income that results from price decrease. There are two approaches to the measurement of substituion effect: the Hick's approach and the Slutsky's approach. Hick's has explained the substituion effect independent of the income effect through compensating variation in income. The amount by which the money income is reduced so that the consumer should be neither better off nor worse off than before is called the compensating variation in income.

2.6.3 Decomposing price effect: compensating variation in income:

This method of decomposing price effect by compensating variation we adjust the income of the consumer so as to effect the change in satisfaction and bring the consumer back to his original indifference curve, i.e. his initial level of satisfaction which he was obtaining before the change in price occurred. For instance, the price of a commodity falls and consumer move to a new equilibrium position at a higher indifference curve which increases his satisfaction. To offset this gain in satisfaction resulting from a fall in price of the good all must take away from consumer enough income to force him to come back to original indifference curve. This required reduction in income to cancel out the gain in satisfaction or welfare occurred by reduction in price of a good is called the compensating variation in income. This is so called because it compensates for gain in satisfaction resulting from reduction in price. How price effect can be split up into income effect and substitution effect can be illustrated with the help of the following diagram:



When the price of good X falls and as a result budget line shift to PL_2 the real income of the consumer rises, i.e., he can buy more of both goods with his given money income. That is, price reduction enlarges the scope for setting up of the two goods. With the new budget line PL_2, the consumer is in equilibrium at point R on the indifference curve IC_2 and thus gains in satisfaction as a result of fall in price of good X. Now if his money income is reduced by the compensating variation in income so that he is forced to come back to the original indifference curve, as before. Now he would buy more of X because it becomes cheap.

In the above figure it is seen that as the price of the good X falls, the budget line of the consumer shifts from PL_1 to PL_2. But with the reduction in the money income of the consumer by compensating variation in income, the budget line shifts to AB parallel to the PL_2 so that it touches the IC_1 where he was before the fall in price of X. Since the price line AB has got the same slope as PL_2, it represents the changed relative prices which means that X becomes cheaper than before. Now as X becomes cheaper than before, the consumer in order to satisfy his satisfaction will substitute more of Y for X. thus when the money income of the consumer is reduced by the compensating variation in income (equal to PA or L_2B), the consumer moves along the IC_1 and is buying MK more of X in place of Y. This movement from Q to S represents the substitution effect since it occurs due to change in relative price alone, real income remaining constant. Now, if the amount of money which is taken away from the consumer is now given back to him, he would move from S to R on higher indifference curve IC_2. This movement from S on a lower indifference curve to R on a higher

indifference curve is the result of income effect. Thus, the movement from Q to R due to price effect can be regarded as having been taken place into two steps: 1st from Q to S due to substitution effect and second from S to R as a result of income effect. It is thus manifest that price effect is the combined result of a substitution effect and an income effect.

Thus, it is clear from the above diagram, it is clear that

MN= Price Effect

MK=Substitution Effect

KN=Income Effect

Or, MN=MK+KN

Or, Price Effect= Substitution Effect+ Income Effect

2.7 Slutsky's Approach:

Slutsky theorem relates to the decomposition of price effect in to income effect and substitution effect by taking into account the apparent real income of the consumer constant. Slutsky uses cost difference method to decompose price effect into its two component parts: the income effect and price effect. The price effect indicates the changes in the purchases of a good say X as a result of change in price of that good, keeping money income, tastes and preferences for other good Y constant. Income effect refers to changes in quantity demanded of a good X as a consequence of change in income while prices of good Y held constant. On the other hand substitution effect means changes in the consumption of goods as a result of the changes in the relative prices of the good alone, keeping the real income constant. Here consumer substitutes good X for relatively cheaper good Y. In the following section, we are trying to discus how price effect can be split in to substitution effect and income effect by using Slutsky's cost difference method.

In this version, when the price of a good changes, consumer's real income increases, the income of the consumer is changed by the amount equal to the change in its purchasing power which is as a result of change in price. The amount equal changes his purchasing power to the change in price multiplied by the number of units of the good, which the individual used to buy at old price. In other words, in Slutsky's approach income is reduced or increased by the amount, which leaves the consumer to be just able to purchase the same combination of goods, if he so desires, which he was having at old price. That is, income is changed by the difference between the cost of the amount good X purchased at old price and the cost of the amount of good X purchased at new price. Income is then said to be changed by cost difference. In Slutsky's substitution effect income is reduced or increased by the cost differences.

Now we are in a position to decompose the price effect in to income effect and substitution effect in light of the above discussion. For this purpose we draw the following diagram:

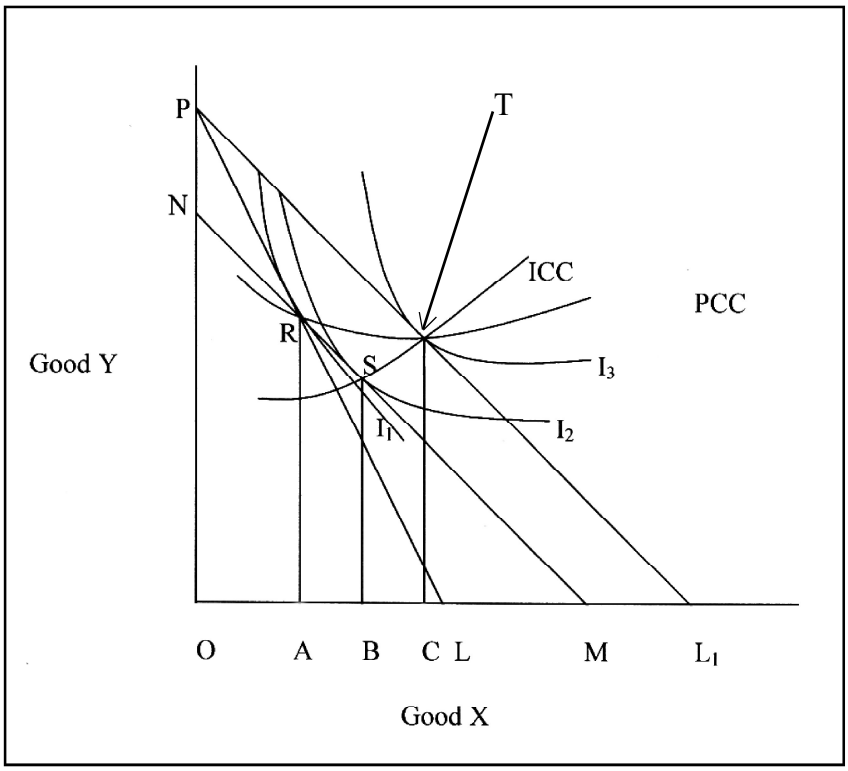


Diagram: 2.6

In the above diagram good X is taken on the horizontal axis and good Y is taken on the vertical axis. Initially, the consumer is in equilibrium at point R where I_1 is tangential to the budget line PL_1. Now, suppose the price of good X fall and consequently the budget line PL shifts to PL_1. Here the consumer is in equilibrium at point T where the I_3curve is tangential to the budget line PL_1. The movement as a result of the fall in the price of good X from R to T is known as the price effect. Now to determine the substitution effect, we have to take away the increased income of the consumer so that he may be able to buy the original combination of goods if he desires. For this, we draw the line MN in such a way that it must be passed through the point R. With price line MN, which is parallel to PL_1, he can buy R if he wishes but in practice he cannot buy this, because X is now relatively cheaper than before. It will substitute Y for X. With budget line MN, the consumer is in equilibrium at S on the indifference curve I_2. This movement from R to S is known as the substitution effect. If now the money that is taken away from the consumer is decided to give back, then he will be able to purchase at point T on the indifference curve I_3. The movement from S to T

is known as the income effect. Thus the movement from R to T can be divided in to two steps: R to S as substitution effect and S to T as the income effect.

In brief,

RT= RS+ST

Price Effect = Substitution Effect + Income Effect.

2.7.1 Superiority of Slutsky over Hicksian Version:

The Slutsky theorem is a good approximation to keep real income constant and is superior to Hick's method. The Slutsky substitution effect provides the consumer greater satisfaction by bringing him on a higher indifference curve, while the Hicksian substitution effect bring him back to the initial level of satisfaction on the original indifference curve. Hick's substitution effect is weak because it is based on compensating variation in income. In the Slutsky method income can be calculated equal to the cost-difference directly by studying market phenomena and behaviour; whereas compensating variation in income is difficult to estimate. In the Slutsky method, the income and substitution effect can be calculated by observing market prices and quantities bought at those prices without any knowledge of indifference curve even. Thus, from practical point of view Slutsky's approach is superior to Hicksian approach.

2.8 Revealed Preference Theory:

"The revealed preference theory" which has been put forward by Prof Samuelson, is a behaviouristic ordinal utility explanation of consumer's demand. This theory analyses consumer's demand from his actual behaviour in the market in various price-income situations. The revealed preference theory is based on the following assumptions:

(i) Choice reveals preference.

(ii) It assumes strong-ordering preferences, under which relation of indifference between various alternative combinations is ruled out.

(iii) It assumes consistency of consumer behaviour

(iv) Tastes and preferences of the consumer do not change.

(v) The consumer prefers larger combination than smaller one.

On the basis of these assumptions, the revealed preference theorem states that when a consumer is observed to choose a combination A out of various alternative combinations open to him, then he reveals his preference for A over all other alternative combinations which he could have purchased. In other words, when a consumer chooses a combination A, it means he considers all other alternative combinations which he could have purchased to be inferior to A.

Graphically, in the following figure, given the income and prices of the two goods X and Y, LM is the price-income line of the consumer. The triangle OLM is the area of choice for the consumer which shows the various combinations of X and Y on the given price-income situation LM. In other words, the consumer can choose any combination between A and B on the line LM or Between C and D below this line. If he chooses A, it is revealed preferred to B. Combinations C and D are revealed inferior to A because they are below the price- income line LM. But the combination E is beyond the reach of the consumer being dearer for him because it lies above his price-income line LM. Therefore, A is revealed preferred to other combinations.

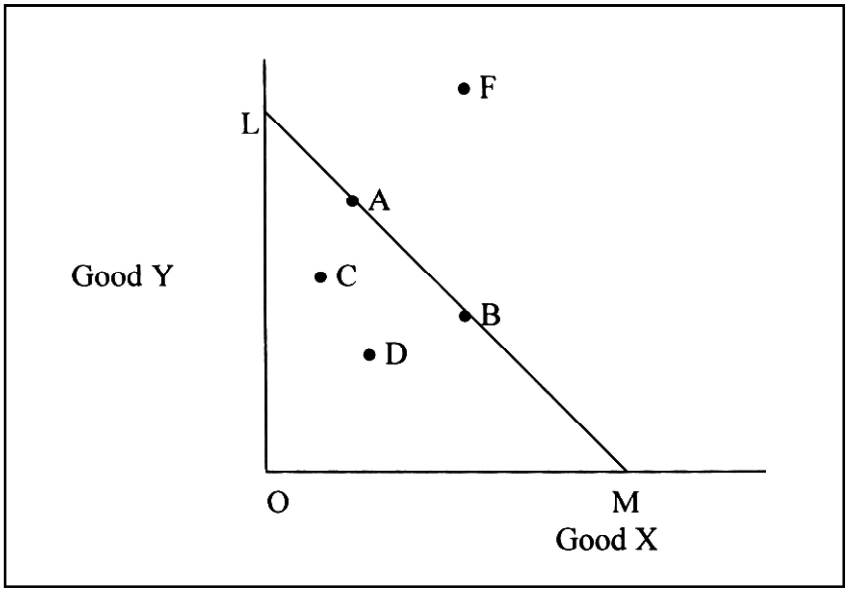


Diagram: 2.7

Derivation of Demand Theorem:

In his revealed preference theory, Samuelson has also derived the Marshallian law of demand what he calls the "Fundamental theorem of consumption theory". He states it in the following words- "Any good that is known always to increase in demand when money income alone rises must definitely shrink in demand when its price alone rises". This statement comprises two parts:

Firstly, it expresses that there is direct positive relationship between consumer's money income and the demand for the commodity, i.e., the income elasticity of demand is positive. Let us explain it with the help of the following diagram :

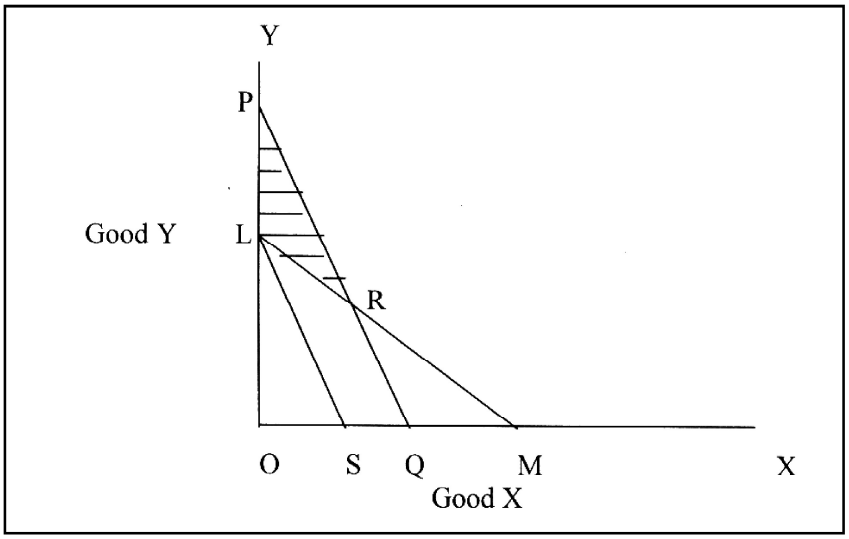


Diagram 2.8

Let us suppose the consumer spends his entire income on two goods X and Y. LM is the original price-income line on which the consumer reveals his preference at point R. Again suppose, the price of X rises so that the new price-income line is LS. Now, in order to compensate the consumer for the loss in his real income as a result of rise in price of X, let us give him LP amount of money income in terms of the good Y. as a result, PQ becomes the new price line and OPQ the new triangle of his choice. Since the consumer was revealing his preference at the point R on the original price-income line LM, therefore, all points lying below R on RM will be inconsistent with his behaviour. He cannot have more of good X when its price has risen. Thus he will either choose combination R or any other combinations in the shaded triangle LPR. But he will choose any combination on the segment RP above point R so that he will have less of X and more of Y. Again, if the extra money is taken back from the consumer, he will buy less of X to the left of R on the budget line LS. Since with the rise in the price of X, demand has shrunk, it proved that income elasticity is positive.

Secondly, demand theorem expresses the inverse relation between price and demand implying that the price elasticity of demand is negative. In figure, LM is the original price-income line on which the consumer reveals his preference at the point R. With the fall in price of X, the price-income line extends to the left as

LS. Suppose, an increase in the real income of the consumer as a result of the fall in the price of X is taken away from him in the form of LP quantity of Y. As a result, PQ becomes the new price-income line and OPQ the new triangle of his choice. Since the consumer is revealing his preference at point R on LM, therefore, all points lying on RL will be inconsistent with his behaviour. He can not have less of good X when its price has fallen. Thus he will either choose combination R or any other combination in the shaded area MRQ. But he can choose any combination in the segment RQ below point R so that with the fall in the price of X, the consumer buys more of X and less of Y. if the money taken from the consumer is returned to him, he will definitely buy more of X to the right of R on the budget line LS. Thus, it is proved that positive income elasticity means negative price elasticity of demand.

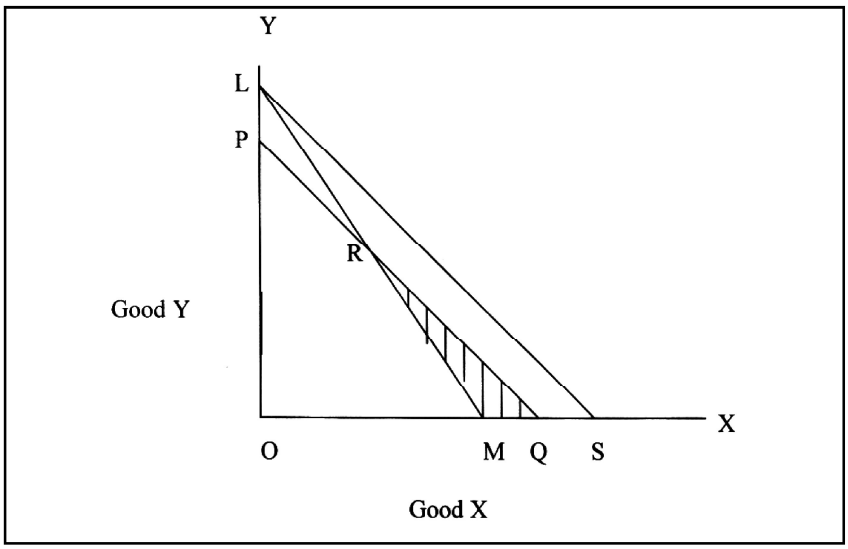


Diagram:2.9

2.9 Consumer's Surplus:

The concept of consumer surplus, which occupies an important place in the Marshallian system of welfare economic analysis, was originally indicated by the classical economists and later formulated and concretized by Javons and Dupuit in1844. Marshall improved and popularized the concept in 1879 in his 'Pure Theory of Domestic Values' and later perfected in his epoch making work "Principles of Economics".

According to Marshall, the satisfaction, which a consumer derives from the purchase of a commodity, exceeds the money value he pays for it. The amount of money spent on a commodity secures for the consumer a greater satisfaction than he could obtain by spending that money on any other thing (otherwise he would have

spent it on other things). Marshall defines consumer surplus thus: 'the excess of the price which he would be willing to pay rather than go without the thing, over that which he actually does pay, is the economic measure of this surplus satisfaction. It may be called consumer's surplus.' Marshall derives the economic measurement of this surplus from the demand curve as shown in the figure below:

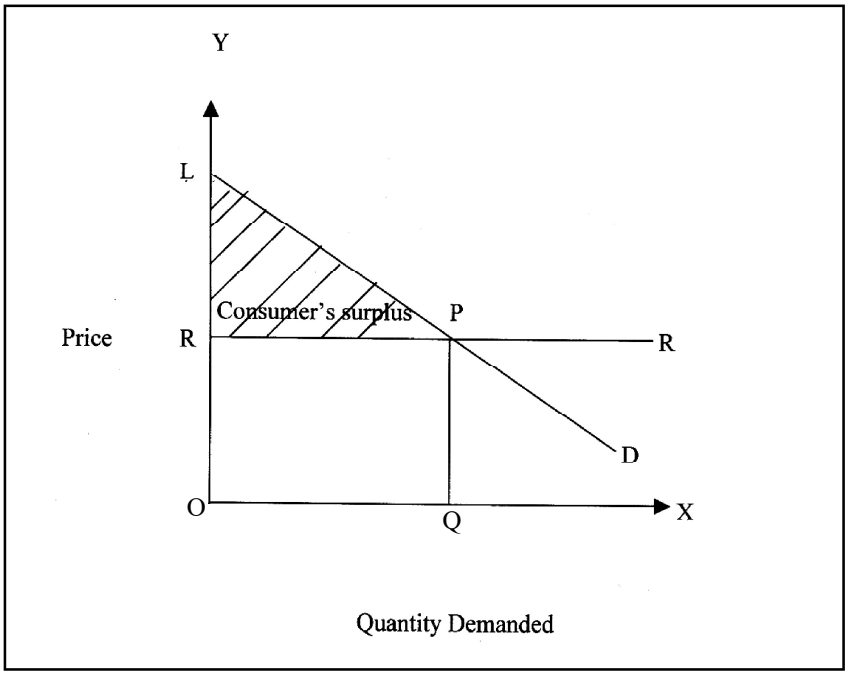


Diagram: 2.10

The demand curve LD shows the set of prices the consumer is willing to pay for the successive units of good, say tea; the demand curve is based on marginal utility. The horizontal straight line RR indicates the actual price in the market for the good on the assumption that the market price is same for all units. Given his demand curve LD, OQ quantity of the commodity will be purchased by consumer at OR price.

OQPL is the total amount of money the consumer is prepared to spend to secure OQ quantity of the commodity. OQPR is the actual amount of money the consumer actually spends on the purchase of OQ units of the commodity and hence LPR is the consumer's surplus.

The surplus secured by the consumer from the purchase of a good is shown to be equal to the excess of the maximum price, which he would be willing to pay-as reflected by his demand curve-rather than go without it over the price which he actually does pay. The concept of consumer's surplus has been criticized on the following grounds:

(i) The concept is not theoretically valid.

(ii) Even if it is theoretically valid, it cannot be measured in terms of money; and

(iii) It is of no practical significance.

2.10 Consumer's Choice Involving Risk:

One very important draw back of the indifference curve analysis is that it cannot explain the consumer behaviour in a risky or uncertain situation. The traditional utility theory also cannot explain the consumer behaviour under risky and uncertain conditions. But in reality, many of the goods and services involve risk or uncertainty, such as investment in stocks, shares, insurance and gambling. Neumann and Morgenstern in their book "Theory of Games and Economic Behaviour" studied the behaviour of the consumer in a risky situation. Later on their theory was refined by Friedman and Savage and by Markowitz.

2.10.1 Neumann-Morgenstern Method of Measuring Utility :

John Von Neumann and Oscar Morgenstern in their book "Theory of Games and Economic Behaviour" developed a method of cardinal measurement of expected utility from risky situations which are found in gambling, lottery tickets etc.,. To measure utility, they have constructed an index called as Neumann-Morgenstern utility index. The Neumann Morgenstern utility analysis is based on the following assumptions:

i. The individual behaves in a risky situation in order to maximize expected utility.

ii. Consumer's choice is transitive.

iii. The probability lies between 0 and 1 ($0<P<1$) such that individual is indifferent between prize A which is certain and the lottery tickets offering prizes C and B with probability P and 1-P respectively.

iv. If two lottery tickets offer the same prizes, the consumer will prefer the ticket with grater probability to win.

v. Uncertainty or risk does not possess utility or disutility of its own.

To quote Neumann and Morgenstern: "consider three events, C, A, B, for which the order of the individual's preferences

is the one stated. Let X be a real number between 0 and 1such that A is exactly equally desirable with the combined event consisting of a chance of probability $(1-\alpha)$ for B and the remaining chance of probability α for C. Then we suggest the use of α as a numerical estimate for the ratio of the preference of A over B to that of C over B." Thus according to the statement the utility of A is certain to occur is equal to the utility of the event B that has (1-P) probability of happening plus the utility of event C which a probability of P happening. Thus

$$U(A)= (1-P).\ U(B) + P.\ U(C) \qquad \text{.....(1)}$$

Which means that the utility of event A which is less than the utility of event C with the probability of happening as one and more than the utility of event B is equal to some value that lies between these two ends depending upon the probability of one or the other event happening. Now the problem is to find out that particular combination of probabilities of events C and B occurrence which would induce the user to desire A to a chance of securing C and B. Let us try to find out that combination of the probabilities with regard to which the consumer is indifferent between event A and events C and B. Let us assign some arbitrary value to C and B. Suppose we assign 3 units of utility to C and 1 unit of utility to B and $2/5$ value to P. Substituting the values in the above equation 1 we get:

$$U(A) = U(3) + U(1)$$
$$= 6/5 + 3/5 = 1.8$$

Proceeding this way, one can derive utility values for U(B) and U(C) and can construct a complete Neumann-Morgenstern utility index for all possible combinations starting from two arbitrary situations involving probabilities of risk. There are two serious shortcomings of the analysis: firstly, it is argued that in the usual consumer decisions risk does not play an important part. Secondly, the analysis fails to take account of the pleasure that gambling gives to certain people. Certain individual's derive pleasure in taking decisions in the face of formidable odds because the mere prospect of winning gives them unfathomable joy. This type of behaviuor is ruled out in the analysis.

2.10.2 The Friedman-Savage Hypothesis:

Milton Friedman and L.J. Savage have developed an

improved analysis of the consumer behaviour in a risky situation. The doctrine of utility maximization cannot be applied to analyze consumer behaviour in situations like gambling and horse race because in these cases economic expectations are negative. But people usually are quite interested in such type of activity although there is a fear in the mind of them to loss that may cause disutility to them. Friedman and Savage have presented an analytical framework to this type of consumer behaviour. Here they have rejected the basic assumption of marginal utility of money so that it may be possible to apply the rule of utility maximization in the study and analysis of choices involving risks and uncertainties like where no risk or uncertainty is involved. In a situation where there is no risks or uncertainties, the people can choose any one alternative that gives him the maximum yield with highest level of satisfaction. But when it involves both the analysis becomes quite difficult one. In such a situation, consumer's selection will depend on the utility attached with various probability distribution of income from different alternatives.

2.11 Recent Developments in Demand Analysis: the Pragmatic Approach

Many writers have questioned the usefulness of the various theories of consumer behaviour. There has been an increasing awareness that although the various approaches to utility are theoretically impressive, there is very little applied economists can use to explain the complex realities of the world. Therefore many writers have followed a pragmatic approach. They accepted the fundamental law of demand on trust and formulated demand functions directly on the basis of market data without regference to the theory of utiltity and the behaviour of the individual consumer behaviour. Demand is expressed as a multivariate function, and can be estimated with various econometric methods. However, a lots of difficulties are associated with such type of estimation. The economists in applied research usually use two types of demand functions. These are:

1. The constant elasticity demand function
2. Dynamic versions of demand function: Distributed lag models of demand.

Check Your Progress

1. What is meant by an indifference curve?

.....

.....

.....

.....

.....

.....

.....

.....

2. Mention the different properties of indifference curve.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

3. What is price effect?

.....

.....

.....

.....

.....

.....

.....

.....

4. State the reveal preference theory.

.....

.....

.....

.....

.....

5. Define consumer surplus.

.....

.....

.....

.....

.....

.....

.....

2.12 Let Us Sum Up

In this unit we discuss theory of consumer demand, various concepts of elasticity of demand, the cardinal and ordinal utility analysis, price effect (both Hicksian and Slutsky's version) and finally on the consumer behaviour under rssky situation and recent developments in the theory of demand.

2.13 Key Words

Elasticity of demand: Elasticity of demand indicates the magnitude of change in consumer's demand for goods as a result of change in the price of the commodity.

Price elasticity of demand: Price elasticity of demand refers to proportionate change in quantity demanded of a commodity to a proportionate change in price of that particular commodity.

Income elasticity of demand: Income elasticity of demand may be defined as the ratio of proportionate change in the quantity demanded to a given proportionate change in the income of consumers.

Cross elasticity of demand: Cross elasticity of demand refers to the ratio of proportionate change in the quantity of a given commodity to a given proportionate change in the price of some other related commodity.

Price elasticity of supply : Price elasticity of supply can be defined as the degree of responsiveness of the quantity supplied of

commodity in response to a small percentage change in its own price.

Indifference curve : An indifference curve shows the locus of points representing pairs of quantities between which the individual is indifferent.

Income effect : Income efect refers to the change in the purchases of a good caused by a given change in the money income of the consumer, other things remaining constant.

Substituion effect : The substitution effect relates to the change in the quantity demand resulting from a change in the price of a good due to the substitution of a relatively cheaper good for a dearer one, while keeping the price of the other good and real income and tastes of the consumer constant.

Price effect : Price effect refers to the way the consumer's purchase of a good changes when its price changes, other things remaing constant.

Suggested Readings

1. Modern Microeconomics, A. Koutsoyiannis, Macmillan
2. Advanced Economic Theory, H.L Ahuja, S. Chand
3. Microeconomics : Theory and Application, Salvatore, Oxford.

2.14 Answer or Hints to Check Your Progress

Please go through the relevant topics. There lies your answer.

BLOCK - II

THEORY OF PRODUCTION AND COSTS

UNIT 1 THEORY OF PRODUCTION

STRUCTURE

1.0 OBJECTIVES

The main aim of this unit is to acquaint you with some important aspects of the theory of production. However, before going into detailed discussions of the principal themes such as the laws of production and other complex derivations, we will explain certain basic concepts related to the theory of production. After studying this unit you should be able to:

- define certain basic concepts related to the theory of production and to draw the lines of differentiation wherever necessary
- understand the meaning and significance of the production function
- represent graphically the production function
- know about the range of output on which the basic theory of production concentrates

- understand how the laws of production describe the technically possible ways of expanding the level of output in the short-run as well as in the long run.
- describe the effect of technological progress on the production function
- derive the equilibrium condition of a single product firm
- learn the graphical way of deriving the cost curves on the basis of the given production function and factor prices.
- derive the production possibility curve and to explain the case of constrained profit maximization of a multiproduct firm.

1.1 INTRODUCTION

Theory of production basically explains how factors of production are combined to produce the outputs or commodities. The technical relationship between the quantity of output and the quantities of inputs is called the production function. It shows how the level of output varies as the factor inputs change.

The laws of production describe the technically possible ways of increasing the level of output.

In the short run, output can be increased by using more of the variable factor(s) while one factor at least is kept constant. The law of variable proportions refers to the short-run analysis of production.

In the long-run, output may be increased by varying all the factors of production. This long-run analysis of production is provided by the laws of returns to scale.

At the very outset, you will acquire some elementary knowledge about the production function, methods of production, isoquants, graphical representation of production function, ridge lines etc. Next, you will find a graphical analysis of the laws of production. Then the unit will deal with the relationship between technological progress and production function. It will also explain how the cost function is derived from the production function .We will also examine the equilibrium of single product firm as well as the equilibrium of multi- product firm.

1.2 PRODUCTION FUNCTION FOR A SINGLE PRODUCT

Production function is purely a technical relation between outputs and factor inputs. It specifies the maximum output that can be produced with given inputs for a given level of technology. Production function reflects the technology of a firm or, as an aggregate production function, the technology of the economy as a whole.

The production function includes all the technically efficient methods of production. A method or process of production is a combination of the factor inputs that is required for the production of one unit of output. A method of production (say A) is said to be technically efficient than any other method (say B), if A uses less of at least one factor and no more from the other factor(s) as compared with B.For instance, let us consider that commodity Q can be produced by the following two methods A and B:

	Process A	Process B
Labour	4	3
Capital	5	5

Method B is technically efficient than method A as it shows the use of less amount of labour input than method A while level of capital input is same for both. You have to remember that only efficient methods of production are used by rational entrepreneurs. But all methods are not directly comparable on the basis of the criterion of technical efficiency. It is the case when a method /process A contains less of some factor(s) and more of some other(s) in comparison to any other method B as shown below:

	Process A	Process B
Labour	3	4
Capital	5	3

In such cases, both the methods are considered as technically efficient and are included in the production function. But you have to remember that the choice of any particular technique by a firm out of the set of technically efficient methods is an economic matter (determined by the prices of factors of production), not a technical one. It is not necessary that a technically efficient method will also be economically efficient.

An isoquant is the locus of all technically efficient methods (or, all the combinations of the factors of production allowed by the existing technology) for producing a particular level of output. In general, the production isoquant is a smooth curve convex to the origin (Diagram : 1).

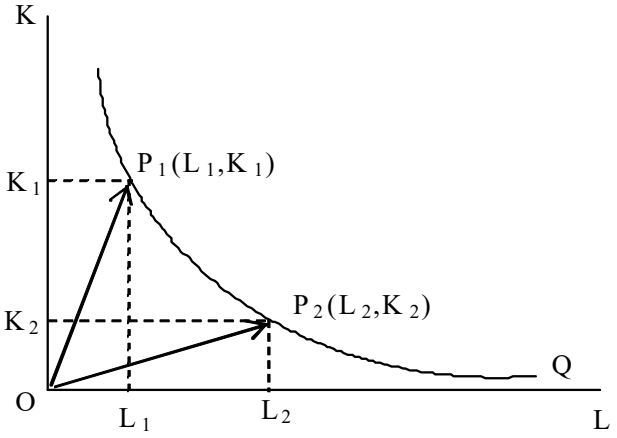


Diagram : 1

This smooth, convex isoquant assumes continuous substitutability of K and L over a certain range of output. The slope of the isoquant ($-\frac{\partial K}{\partial L}$) defines the degree of substitutability of the factors of production (Diagram: 2).

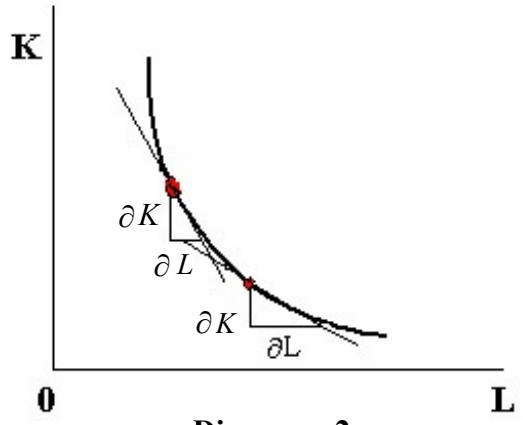


Diagram: 2

As we move downwards along an isoquant we find that the slope of the isoquant decreases in absolute terms. It reflects the fact that substitution of K for L become more and more difficult as we move downwards along an isoquant.The slope of the isoquant is termed as the marginal rate of technical substitution or the marginal rate of substitution (MRS) of the factors:

$$-\frac{\partial K}{\partial L} = MRS_{L,K}$$

Furthermore, the MRS is equal to the ratio of the marginal products of the factors. Mathematically, marginal product of each factor is given by the partial derivative of the production function with respect to this factor. That is,

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

$$MP_K = \frac{\partial Q}{\partial K}$$

Hence we can write,

$$MRS_{L,K} = -\frac{\partial K}{\partial L} = \frac{\partial Q/\partial L}{\partial Q/\partial K} = MP_L/MP_K$$

The MRS depends on the units of measurement of the factors. It makes MRS inferior as a measure of the degree of substitutability of the factors. The elasticity of substitution (σ) which is independent of the units of measurement of the factors, provides a better measure of the ease of factor substitution. It is defined as the percentage change in the capital-labour ratio divided by the percentage change in the marginal rate of technical substitution.

The whole array of production isoquants is described by the production function. Each isoquant of the array refers to a different level of output. It explains the variation of output with the variation in factor inputs.

In the traditional economic theory the production function assumes the form :

$$Q = f(L, K, \upsilon, \gamma)$$

Where

- Q = output
- L = labour input
- K = capital input
- υ = returns to scale.
- γ = efficiency parameter, refers to the entrepreneurial-organizational aspects of production.

Production function for a single product is usually represented by a two dimensional curve as shown in Diagram: 3 and Diagram: 4.

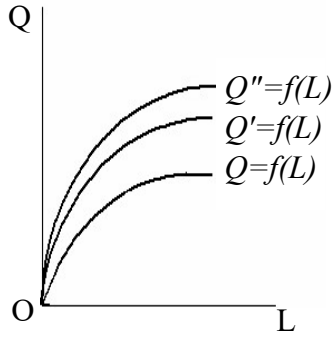


Diagram: 3

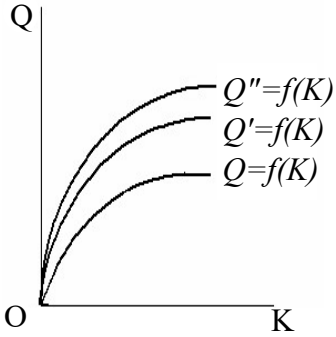


Diagram: 4

In Diagram: 3, each curve depicts the relation between Q and L for given K, υ and γ. The movement along the curve shows that as labour increases, other things remaining the same, output increases. If K (and/or υ, and / or γ) increase the production function Q= f (L) shifts upward as shown by the shifted production functions Q'=f(L) and Q" =f(L) (Diagram : 3).

Similarly, in Diagram : 4, each curve depicts the relation between Q and K for given L, υ and γ. The movement along a curve shows that as capital increases, other things remaining the same, output increases. If L (and/or υ, and/or γ) increase the production function Q= f (K) shifts upward as shown by the shifted production functions Q'=f(K) and Q"=f(K) (Diagram : 4).

The slopes of the curves depicting the production function Q=f (L) (Diagram: 3) and Q=f(K) (Diagram: 4) give the marginal productivity of labour input (MP_L) and the marginal productivity of capital input (MP_K) respectively. Mathematically, marginal product of each factor is given by the partial derivative of the production function with respect to this factor. Thus,

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

$$MP_K = \frac{\partial Q}{\partial K}$$

Marginal product of a factor may be positive, zero, or negative. But you must remember that the basic theory of production usually concentrates only on the efficient range of output. It refers to the ranges of output over which the marginal products of the factors are positive but diminishing. In other words, you can also say that the basic theory of production usually concentrates on the levels of employment of factors of production over which the marginal products of the factors are positive but diminishing. The ranges over which marginal products of the factors become negative are not considered by the rational producers. Mathematically, the ranges of interest of the rational producers are featured by:

$$MP_L > 0 \quad \text{but,} \quad \frac{\partial (MP_L)}{\delta L} < 0 \quad \text{--------- (i)}$$

$$MP_K > 0 \quad \text{but,} \quad \frac{\partial (MP_K)}{\delta K} < 0 \quad \text{-------- (ii)}$$

The conditions (i) and (ii) means that that traditional theory of production concentrates on the range of isoquants over which their slope is negative and convex to the origin.

Let us now present the production function in the form of a set of isoquants possessing usual properties. By joining the points on the successive isoquants where the marginal products of the factors are zero, the two ridge lines are obtained (Diagram : 5).

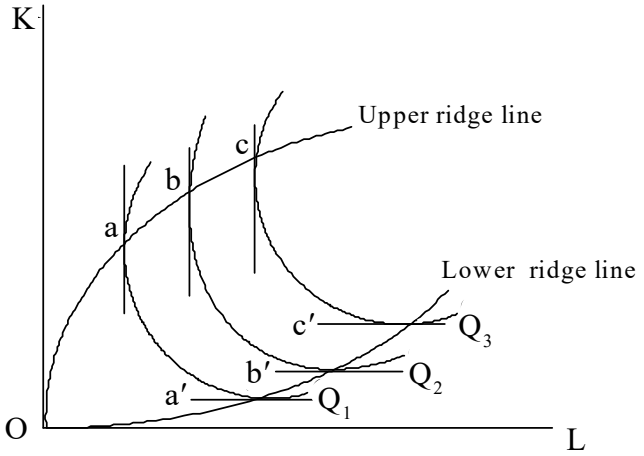


Diagram 5 : Ridge lines

At each point on the upper ridge line the marginal product of capital is zero. On the other hand, each point on the lower ridge line implies that the marginal product of labour is zero. Production methods are technically efficient only inside the ridge lines. Outside the ridge lines the marginal products of the factors are negative. It implies technical inefficiency of the methods of production as they need more quantity of both capital and labour inputs for producing a given level of output. Such inefficient methods are not considered by rational producers. The range of isoquants over which they are convex to the origin indicates the range of efficient production.

Check Your Progress-1.1

Note: **a) Use the space given below for your answer.**

b) Check your answer with the model answer given at the end of the unit.

1) The following statements are based on the text you have already read. Indicate whether these statements are true or false by putting a tick mark (✓) in the relevant box.

		True	False
i)	Production function refers to a technical relationship between output and factor inputs.	[ ]	[ ]
ii)	An isoquant is the locus of all efficient methods (or all the combinations of the factors of production allowed by the existing technology) for producing a particular level of output.	[ ]	[ ]
iii)	The basic theory of production usually concentrates on the range over which marginal products of the factors are negative.	[ ]	[ ]

1.3 LAWS OF PRODUCTION

The laws of production describe the technically possible ways of expanding the level of output. The level of output may be expanded in different ways.

In the long-run output may be increased by changing the levels of all the factors of production. This long-run analysis of production is provided by the laws of returns to scale.

In the short run output can be increased by using more of the variable factor(s) while one factor at least is kept constant. The law of variable proportions refers to the short-run analysis of production. This law is also termed as the law of (eventually) diminishing returns of the variable factor because the marginal product of the variable factor(s) will decline eventually as more and more quantities of the variable factor(s) is (are) combined with the fixed factor(s).

Let us examine the laws of returns to scale and the law of variable proportions one by one.

1.3.1 LAWS OF RETURNS TO SCALE

The laws of returns to scale refer to the long-run analysis of production. You know that in the long run, all factors are variable and therefore, output may be increased by changing all the factors of production. The factors may be varied either by the same proportion or by different proportions. You have to remember that 'returns to scale' refers to the changes in output when all the factor inputs change by the same proportion.

We can explain the concept of 'returns to scale' with the help of the following production function:

$$Q=f(L, K).$$

Let us consider that both the inputs i.e., labour (L) and capital (K) are allowed to increase by the same proportion λ. Then, it is obvious that the initial level of output will also increase. The resulting level of output (say, Q*) will obviously be greater than the original level of output (Q).

$$Q^*=f(\lambda L, \lambda K) \quad \text{.....(1)}$$

If Q* increases by the same proportion λ with which the inputs are increased, it is said that there are constant returns to scale.

If Q* increases less than proportionately with the increase in the inputs, it implies that there are decreasing returns to scale.

And if Q* increases more than proportionately with the increase in the inputs, it refers to the case of increasing returns to scale.

Let us now learn the meaning of homogeneous production function and its relation with the returns to scale. If in (1), λ,the term by which each factor inputs were multiplied, can be completely taken out of the brackets as a common factor, then the given production function Q=f(L,K) is termed as homogeneous production function (otherwise non- homogeneous). Then it will be possible to write the new level of output Q* as a function of λ^{υ} and the original level of output Q as given below :

$$Q^* = \lambda^{\upsilon} f(L, K)$$

or, $$Q^* = \lambda^{\upsilon} Q \text{ ----------(2)}$$

The power υ of the term λ in (2) is termed as the degree of homogeneity of the function. υ provides a measure of the returns to scale as follows:

If $\upsilon = 1$, we get constant returns to scale which means that output increases by the same proportion λ as the inputs. The production function in this case is termed as linear homogeneous.

If $\upsilon < 1$, it implies decreasing returns to scale which means that output increases less than proportionately with the increase in the inputs.

If $\upsilon > 1$, it refers to increasing returns to scale which implies that output increases more than proportionately with the increase in the inputs.

Product Lines

In order to understand the graphical representation of the returns to scale, you need to learn about concepts of product line and isocline beforehand.

The expansion of output can be shown graphically in two ways: you can either use the concept of product lines (in two-dimensional framework), or, introduce a third dimension. The first option is easier than the latter. Now let us try it.

A product line shows the (physical) movement from one isoquant to another in response to a change in both factors or a single factor of production. You have to remember one significant point that the concept of product line is independent of factor prices. Hence, it only describes various technically possible alternative ways of expanding output. But what path out of these will be the actual choice of the corresponding firm will be determined by the prices of the factors of production.

When all factors are variable, the product line passes through the origin Diagram: 7 & Diagram : 8. If only one factor is allowed to vary keeping the others at a constant level, the product line becomes a straight line parallel to the axis of the variable factor (Diagram : 6).

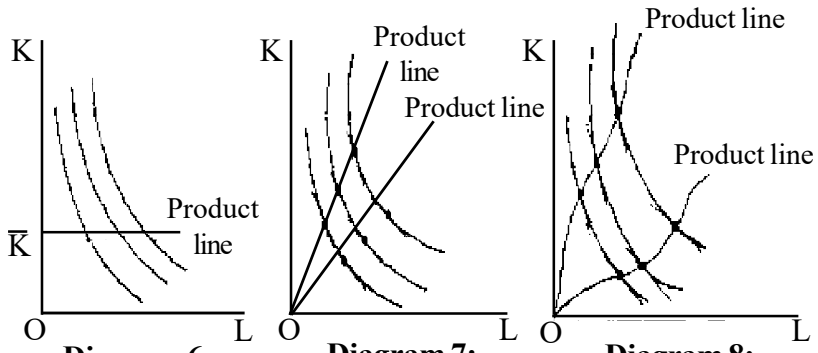


Diagram 6: Product line when level of capital (K) is given

Diagram 7: Product line for homogeneous production function

Diagram 8: Product line for non-homogeneous production function

Some product lines which are useful for the choice of the firm are termed as isoclines. An isocline is the locus of points of different isoquants at which the marginal rate of substitution (MRS) of factors i.e., slope of the isoquants is constant.

For homogenous production the isoclines are straight lines through the origin. Along such an isocline the K/L ratio is constant (along with the MRS of the factors). Obviously for different isoclines, the K/L ratio (as well as the MRS of the factors) is different (Diagram :7)

When the production function is non-homogeneous, the shape of the isocline will be non-linear (Diagram : 8).It is obvious that the K/L ratio varies along each such isocline (as well as on different isoclines).

Graphical Presentation of the Returns to Scale for a Homogeneous Production Function

After being equipped with the concept of isocline, it will now be easier for you to understand the graphical explanation of the returns to scale. To explain the returns to scale graphically for a homogeneous production function you just have to measure the distances between successive multiple-level-of-output isoquants along an isocline. Here, multiple-level-of-output isoquants refer to the isoquants which represent the levels of output that are multiple of some base level of output, e.g. Q, 2Q, 3Q, 4Q, etc. Let us discuss the cases of constant, decreasing and increasing returns to scale one by one.

(i) Constant returns to scale: When along any isocline the distance between successive 'multiple-level-of-output' isoquants remain constant, it refers to constant returns to scale. Output increases by the same proportion as the inputs. By doubling the

initial levels of factor inputs double of the level of initial output is obtained. In Diagram : 9 point 'b' defined by 2K and 2L, lies on the isoquant which represents 2Q level of output i.e. just twice the base level of output Q.

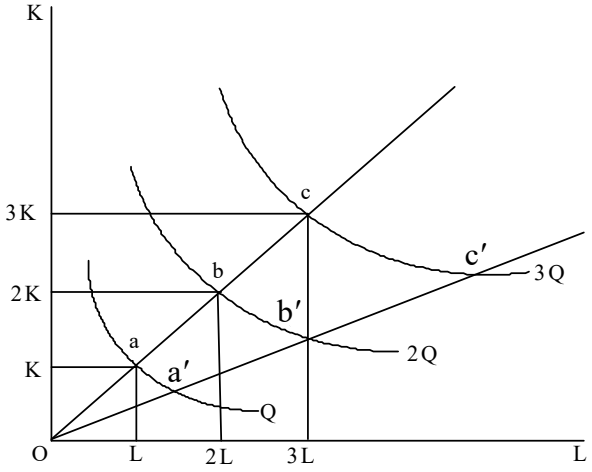


Diagram : 9 - Constant returns to scale (Oa=ab=bc)

(ii) Decreasing returns to scale: When along any isocline the distance between consecutive 'multiple-level-of-output' isoquant increases, it refers to decreasing returns to scale. Doubling the factor inputs results in the level of output which is less than double of the initial level of output. Diagram : 10 explains it clearly. The point b′, defined by 2K and 2L, lies on an isoquant below the one that represents 2Q level of output.

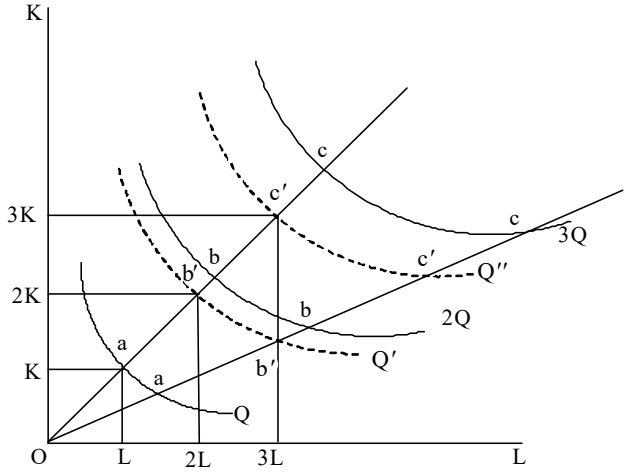


Diagram : 10- Decreasing returns scale (Oa < ab < bc)

(iii) Increasing returns to scale: When along any isocline the distance between consecutive 'multiple-level-of-output' isoquant decreases, it implies increasing returns to scale. In Diagram :11 depicts that by doubling K and L, a point b′ is achieved which lies upon an isoquant above the one representing 2Q. It means that output increases more than proportionately with the increase in the inputs.

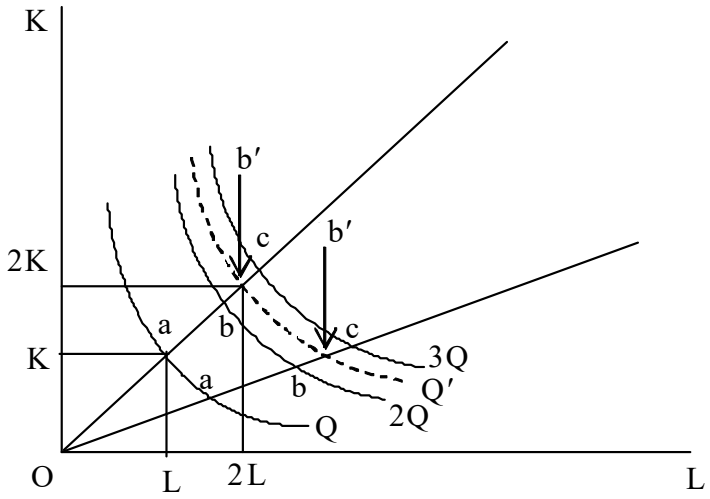


Diagram : 11- Increasing returns to scale (Oa > ab > bc)

(iv) Varying returns to scale: In general it is assumed that returns to scale are same along all the expansion-product lines. But technological conditions of production may lead to variation in the returns to scale over different ranges of output. In Diagram :12, it is shown that up to the output level 4Q, returns to scale are constant and beyond that level of output returns to scale are diminishing.

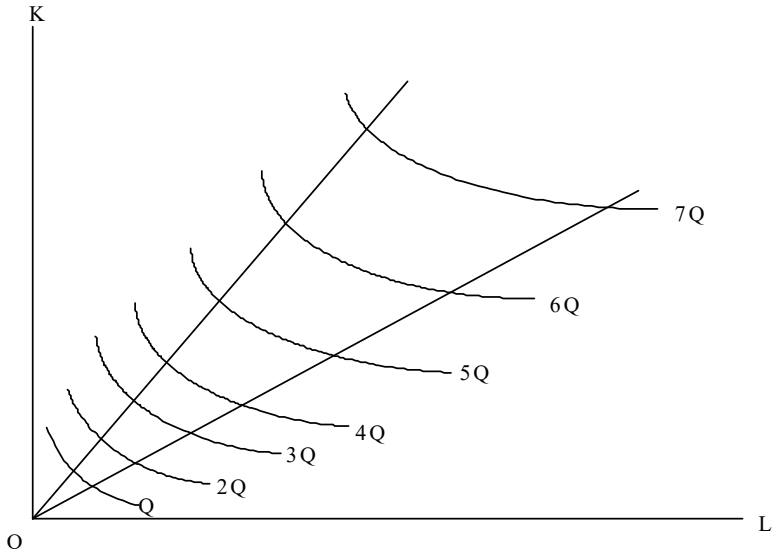


Diagram : 12 - Varying returns to scale (up to 4Q returns to scale are constant & beyond 4Q level of output returns to scale are decreasing)

Causes of Increasing and Diminishing Returns to Scale:

The main causes of increasing returns to scale are technical and/or managerial indivisibilities. In general, duplication is possible

for most of the processes but they may not be halved. To illustrate it let us consider that we have the following three processes:

	L (Men)	K (machines)	Q (output in tons)
A: Small scale process	1	1	1
B: Medium scale process	40	40	80
C: Large scale process	80	80	300

It is clear that for all three processes K/L ratio is the same. For each process duplication is possible but it is not possible to halve them. 'Unit'- level appears to be different for different methods. Each process shows constant returns to scale. It is clear that for output levels less than 40 (i.e., $Q<40$) tons, firms will use the small scale process. Depending upon the requirement, they will increase the input levels to get the proportionate increase in output level. For instance if the market requirement is 35 tons of output, the firm will increase the base level of inputs (corresponding to process A) by 35 times to get 35 tons of output (on the basis of constant returns to scale assumption). For $40<Q<80$, the firm will use the process B. The shift from process A to process B gives a discontinuous increase in output (from 39 tons of output produced with 39 units of L and 39 units of K, to 80 tons produced with 40 units of L and 40 units of K). If the market demand is only 60 tons, the firm would still use the process B inefficiently (producing only 60 units, or producing 80 units and throwing away the additional 20 units). It is because the process B, even though inefficiently used, is still more productive or relatively efficient than process A. Duplication of process A for the range of output in between 40 tons and 80 tons ($40<Q<80$) will not be profitable for the firm when relatively more productive process B is available. The same type of interpretation is valid in case of a shift from process B to process C. Though each process individually exhibits constant returns to scale, their indivisible nature tends to result in increasing returns to scale.

On the other hand, 'Diminishing returns to management' are the most common causes of decreasing returns to scale. With the growth in output, when the top management of the firm becomes eventually overburdened, and begin to perform less efficiently its role as coordinator and ultimate decision maker, decreasing returns to scale takes place. In spite of the development of modern management science, it continues to be a major observation that as firms grow beyond the appropriate optimal level, management diseconomies starts. Exhaustible natural resources may also give rise to decreasing returns to scale.

1.3.2 THE LAW OF VARIABLE PROPORTIONS: SHORT-RUN ANALYSIS OF PRODUCTION

You have already learned that if only one factor is allowed to vary keeping the other(s) at a constant level, the product line will be a straight line parallel to the axis of the variable factor (Diagram : 6)

In general, in the short-run the level of output can be increased by increasing the use of the variable factor(s) while keeping one factor (at least) constant. In this case the marginal product of the variable factor(s) will decline after a certain range of output as more and more quantities of the variable factor(s) are combined with the other constant factor(s). We have already discussed under section (1.2) that the basic theory of production concentrates on the ranges of output over which the marginal products of the factors are positive but diminishing. The range of increasing returns to a factor and the range of negative productivity are not equilibrium ranges of output.

Now let us examine the law of variable proportions or, the law of (eventually) diminishing returns of the variable factor in some detail, which describes the expansion of output with one factor (at least) constant. We will consider all the three cases:

(i) When the production function is homogeneous and exhibits constant returns to scale.

(ii) When the production function is homogeneous and exhibits diminishing returns to scale and

(iii) When the production function is homogeneous with increasing returns to scale.

Case-I : When the production function is homogeneous and exhibits constant returns to scale everywhere on the production surface, the returns to a single variable factor will be diminishing. This is implied by the negativity and convexity of the slope of the isoquants. With the assumption of constant returns to scale everywhere on the production surface, by doubling the initial levels of factor inputs double of the level of initial output is obtained. See how Diagram :13 explains the occurrence of diminishing returns to the single variable factor labour (L).

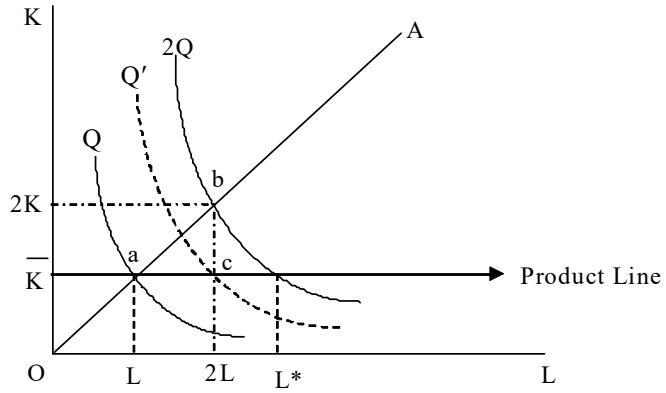


Diagram: 13

In the Diagram: 13, point 'b' on the isocline OA corresponding to 2K and 2L input levels, lies on the isoquant which represents 2Q level of output. But if capital input is kept constant at $\overline{K}$ level and only the labour input is doubled, the point 'c' is achieved which lies on an isoquant that shows the output level Q′ which is less than 2Q (i.e., Q′ <2Q). We can have 2Q level of output with the constant level of capital K only if labour input (L) is increased to L* units. It is obvious that L*>2L. Thus, doubling only L and keeping K constant, less than double of the initial output is obtained. It clears the fact that the variable factor labour (L) exhibits diminishing returns in the case of a production function which is homogeneous with constant returns to scale.

Case-II: When the production function is homogeneous with decreasing returns to scale everywhere on the production surface, the returns to a single variable factor will certainly be diminishing. In this case, obviously, doubling the factor inputs results in the level of output which is less than double of the initial level of output. In Diagram: 14, the point 'd' defined by 2K and 2L, lies on an isoquant below the one that represents 2Q. But if only labour is doubled and capital is kept constant at $\overline{K}$ level output reaches the point 'c' that lies on a still lower isoquant.

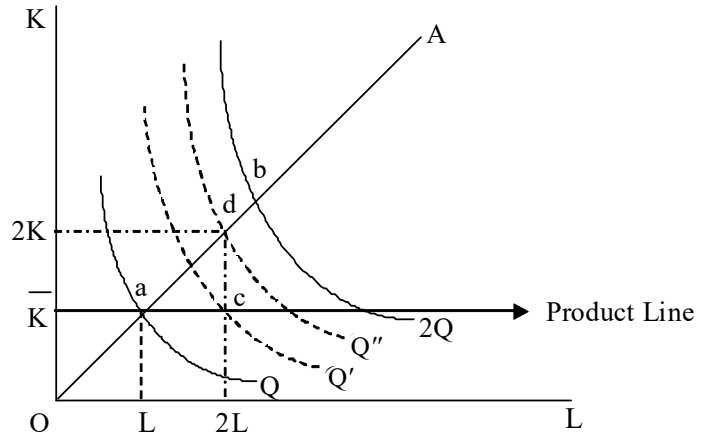


Diagram: 14

Case-iii : If the production function shows increasing returns to scale everywhere on the production surface, the productivity of the variable factor will in general be diminishing . Diagram: 15 explains the usual case of diminishing returns to a single variable factor when the production function is homogeneous and exhibits increasing returns to scale.

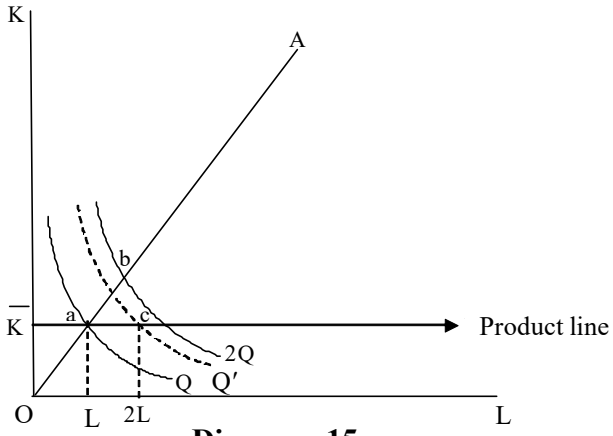


Diagram :15

By doubling the initial levels of factor inputs more than double of the initial level of output is obtained. In other words, double of the initial level of output is obtained by employing less than double of the initial levels of factor inputs. The Diagram : 15 makes it clear that the point defined by 2K and 2L,will obviously be on an isoquant higher than that representing 2Q level of output.Try to draw that isoquant yourself in the above diagram. With the constant level of capital $\bar{K}$ and 2L labour units, we reach the point 'c' in Diagram: 15, which lies on an isoquant lower than that indicating 2Q.

But one significant point you have to keep in mind is that, in this case, the diminishing returns arising from the decreasing marginal product of the variable factor (labour) may be offset if the returns to scale are too strong to offset the former. Diagram: 16 shows that returns to scale are too strong to offset the diminishing productivity of the single variable factor L.

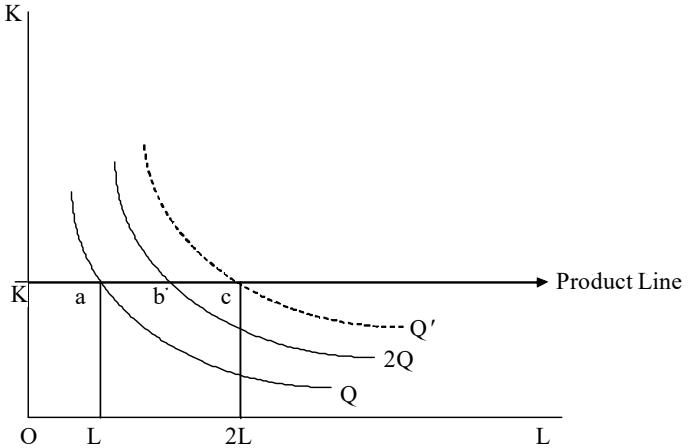
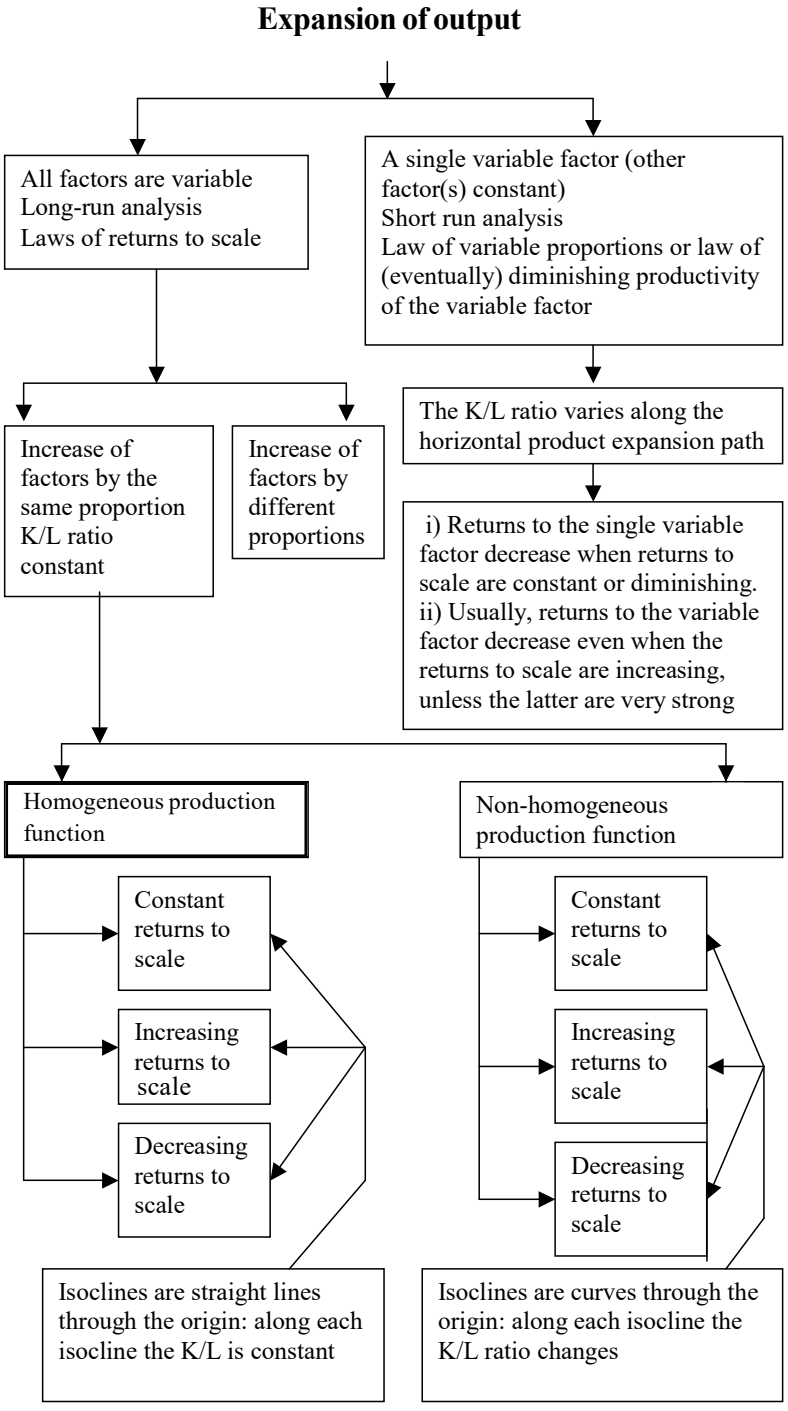


Diagram: 16

Even by keeping capital constant at $\overline{K}$ level and just by doubling labour input we reach a level of output which is higher than 2Q level (point 'c' in the Diagram: 16). However, it is an exceptional case in practice. Hence, you can say that in general, the productivity of a single variable factor (other things remaining the same) is diminishing.

SUMMARY: On the basis of the above analysis we can summarize the laws of production schematically as shown below (Flow Chart-1).



Flow Chart :1 : Schematic representation of the laws of production.

Check Your Progress-1.2

Note:

a) Use the space given below for your answer.

b) Compare your answer with the model answer given at the end of the unit.

1) After reading thoroughly the preceding text complete the following statements:

i) If the degree of homogeneity is equal to one (i.e., $\upsilon=1$), we get ..... returns to scale.

ii) If the degree of homogeneity is less than one (i.e., $\upsilon<1$), it implies..... returns to scale.

iii) If the degree of homogeneity is greater than one (i.e. $\upsilon>1$), it refers to..... returns to scale.

iv) In case of decreasing returns to scale, along any isocline the distance between consecutive 'multiple-level-of-output' isoquants .....

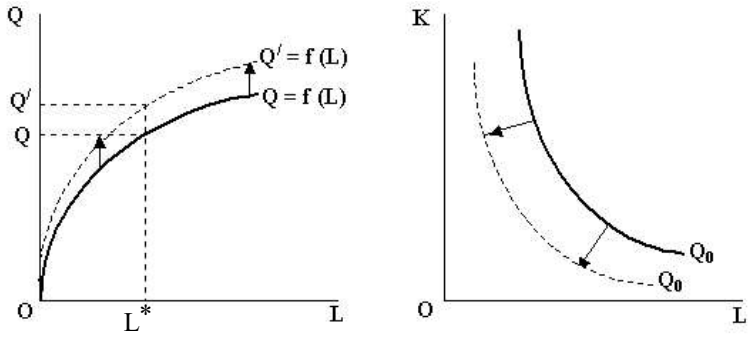
v) Technical and/or managerial indivisibilities are the main causes of ..... returns to scale.

2) Differentiate between the law of variable proportions and the laws of returns to scale.(Hint: see the Text)

1.4 TECHNOLOGICAL PROGRESS AND PRODUCTION FUNCTION:

Technological progress refers to the development of new and more efficient processes of production to produce more and/or improved output from the same set of factor inputs, or, introduction of new products. The introduction of innovation is the most important determining factor of a firm's long-term competitiveness.

Let us now learn how the effect of innovation in processes can be illustrated graphically. Usually it is shown by the upward shift of the production function (Diagram: 17) or, a downward movement of the production isoquants (Diagram: 18). The meaning is that, with technological progress same level of output can be produced with less factor inputs than before .You can also say that with the same level of inputs now it is possible to produce more output.



You have to remember one more thing that technical progress not only produce a shift of the production isoquant but may also change its shape. Here we will discuss the Hicksian classification of technical progress. This classification is based on how the technical progress affects the rate of substitution of the factors of production. Let us illustrate the concepts graphically.

i) Capital deepening technical progress: When along a line passing through the origin on which K/L ratio is constant, the $MRS_{L,K}$ (=MP_L/ MP_K) decreases in absolute terms (but increases when the negative sign of the slope of the isoquant is taken into account), the technical progress is termed as capital deepening technical progress or capital using technical progress. It means that as a result of technical progress the marginal product of capital increases more than the marginal product of labour. Along any given radius the slope of the shifting isoquant becomes more flat for successive shifts as shown in Diagram : 19.

ii) Labour deepening technical progress: When along a line passing through the origin on which K/L ratio is constant, the $MRS_{L,K}\left(=\frac{MP_L}{MP_K}\right)$ increases in absolute terms (but decreases when the negative sign of the slope of the isoquant is taken into account),the technical progress is termed as labour deepening technical progress or labour using technical progress. It means that as a result of technical progress the marginal product of labour increases more than the marginal product of capital. Along any given radius, the slope of the shifting isoquant becomes more steep for successive shifts as shown in Diagram: 20.

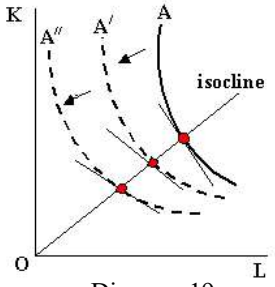


Diagram: 19
Capital - deepening
Technical Progress

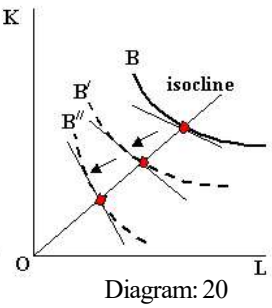


Diagram: 20
Labour - deepening
Technical Progress

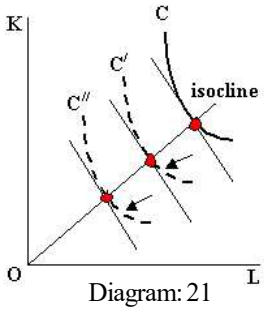


Diagram: 21
Neutral Technical
Progress

iii) Neutral technical progress : When along a line passing through the origin on which K/L ratio is constant, the $MRS_{L,K}\left(=\frac{MP_L}{MP_K}\right)$ remains constant, the technical progress is termed as neutral technical progress. It means that as a result of the technical progress the marginal products of labour and capital increases at the same rate. Along any given radius, the slope of the shifting isoquant remains the same as shown in Diagram: 21.

Check Your Progress-1.3

Note:

a) Use the space given below for your answer.

b) Compare your answer with the model answer given at the end of the unit.

1) The following statements are based on the text you have already read. Indicate whether these statements are true or false by putting a tick mark (✓) in the relevant box.

	True	False
(i) For capital deepening technical progress, along a line passing through the origin on which K/L ratio is constant, the $MRS_{L,K}$ increases in absolute terms.	[ ]	[ ]
(ii) For labour deepening technical progress, along a line passing through the origin on which K/L ratio is constant, the $MRS_{L,K}$ increases in absolute terms.	[ ]	[ ]
(iii) For neutral technical progress, along a line passing through the origin on which K/L ratio is constant, the $MRS_{L,K}$ remains constant.	[ ]	[ ]

1.5 EQUILIBRIUM OF THE FIRM : CHOICE OF OPTIMAL COMBINATION OF FACTORS OF PRODUCTION

We will examine here the choice of optimal combination of factors of production under two heads: (i) constrained profit maximization in a single period (either by maximizing output subject to a cost constraint or by minimizing cost for a given output constraint) and (ii) unconstrained profit maximization via expansion of output over time.

For all the cases the basic objective of the firm is assumed to be the maximisation of profit (π). It means that the difference between total revenue (R) and total costs (C) is to be maximised. Prices of the factors of production i.e., wage rate (w) and price of capital services (r), price of output (P_Q) are assumed to be given and these given levels are denoted by $\overline{w}$, $\overline{r}$, and $\overline{P}_Q$ respectively.

1.5.1 CONSTRAINED PROFIT MAXIMISATION

Here we will examine following two forms of constrained profit maximisation problem faced by the firm:

(a) Maximisation of profit (π) is obtained by maximisation of output given a cost constraint. Total cost and prices of inputs as well as output are assumed to be given at $\overline{C}$, $\overline{w}$, $\overline{r}$ and $\overline{P}_Q$ level respectively.

The statement of the problem can be made as -

$$\text{Maximise} \quad \pi = R - C$$
$$= \overline{P}_Q.Q - \overline{C}$$

As $\overline{P}_Q$, $\overline{C}$ are given constants, π can be maximised only by maximising the level of output (Q).

(b) In this form, the maximisation of profit is obtained via cost minimisation subject to an output constraint. The statement of the problem can be made as:

$$\text{Maximise} \quad \pi = R - C$$
$$= \overline{P}_Q.\overline{Q} - C$$

It is obvious that the value of profit (π) will be maximised if cost (C) is minimised as Q and P_Q are assumed to be constant at $\overline{Q}$ and $\overline{P}_Q$ level. Let us examine both the cases graphically.

Maximisation of output subject to a cost (financial) constraint:

Let us assume that the production function is given by:

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

Total cost outlay (C) and prices of factor inputs i.e., wage rate (w) and price of capital services (r) are assumed to be constant.

The firm will reach its equilibrium position when it will maximise its output level subject to the given cost constraint. Formally, the problem can be stated as:

Maximise $Q = f(L, K)$

Subject to $\bar{C} = \bar{w}L + \bar{r}K$ (cost constraint)

The production function will be represented by an isoquant map and the isocost line AB will represent the cost constraint (Diagram : 22).

You have already learned that the slope of the isoquant is given by -

$$-\frac{\partial K}{\partial L} = MRS_{L,K} = \frac{MP_L}{MP_K} = \frac{\partial Q/\partial L}{\partial Q/\partial K}$$

Now we have to define the isocost line and its slope.

The isocost line is defined by the locus of all combinations of factor inputs (here L and K) that can be purchased by the firm with a given monetary cost expenditure. In equation form, it is given by the cost equation.

$$C = r.K + w.L$$

And the slope of the isocost line is given by the ratio of the factor prices

$$\text{i.e., slope of the isocost line} = \frac{w}{r}$$

Let us assume that the given total cost outlay of the firm is $\bar{C}$. If the firm spends the whole amount of money on K- input, then the maximum amount of the factor K that it can purchase will be

$$OA = \frac{\bar{C}}{r}$$

On the other hand if the firm spends the entire amount $\bar{C}$ on L-input then the maximum quantity of L- input it can purchase will be

$$OB = \frac{\bar{C}}{w}$$

Hence the slope if the isocost line AB (see Diagram : 22) will be

$$\frac{OA}{OB} = \frac{\overline{C}/r}{\overline{C}/w} = \frac{w}{r}$$

With these learning about the slope of isoquant and isocost line, it will now be easier for you to deal with the graphical solution of the problem. The maximum level of output which the firm can produce subject to the given cost constraint (represented by the isocost line AB) is Q_2. In Diagram: 22 the equilibrium position of the firm (i.e., its profit maximising position) is defined by the point 'e'which is the point of tangency of the given isocost line AB and the highest possible isoquant Q_2 that can be reached subject to the given cost constraint.The coordinate of the point 'e' i.e. (L_2, K_2) represent the optimal combination of the factors of production. Given the factor prices w and r, L_2 units of labour and K_2 units of capital will be the profit maximising combination of factor inputs. With the aforesaid assumptions of given production function (represented by the isoquant map), given cost outlay and factor prices (represented by the isocost line AB), Q_2 is the maximum possible output that can be achieved.

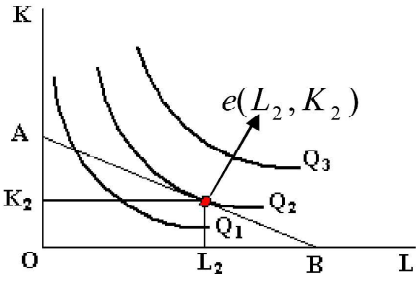


Diagram : 22

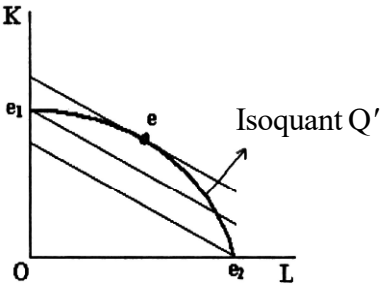


Diagram : 23

It is because of the fact that though higher levels of output (represented by the isoquants right to the point 'e'; like Q_3 in Diagram : 22) are desirable, it is not possible to achieve them because of the assumed cost constraint. Except 'e', all other points on AB or below AB lie on lower isoquant than that representing Q_2 and hence correspond to lower levels of output.

The first condition for the firm's equilibrium is that at the tangential point 'e', the slope of the isocost line AB must be equal to the slope of the isoquant Q_2. The second order condition states that the isoquants must be convex to the origin.

Thus the two conditions for the equilibrium of the firm are as follows :

Condition -I :

Slope of the isocost = slope of the isoquant.

$$\Rightarrow \frac{w}{r} = \frac{MP_L}{MP_K} = \frac{\partial Q/\partial L}{\partial Q/\partial K} = MRS_{L,K}$$

Condition -II:

The isoquants must be convex to the origin as shown in the Diagram : 22. From Diagram : 23, it is clear that if the isoquant is concave, the tangential point between the isocost and the isoquant will not be the equilibrium point. It is obvious from the Diagram:23 that Q′ level of output represented by the concave isoquant can be produced with lower cost at e_1 position, as e_1 corresponds to a lower isocost line than e.The same level of output can also be produced at e_2 which is upon a further lower isocost line. The solution that is obtained in case of a concave isoquant is termed as "corner solution".

Case - 2: Minimisation of cost subject to an output constraint

Formally the problem can be stated as:

Minimise $C = f(Q) = wL + rK$

Subject to $\bar{Q} = f(L.K)$.

In this case it is obvious that we will have a single isoquant representing the given level of output ($\bar{Q}$) and a set of isocost lines (Diagram : 24). The isocost lines are parallel to each other reflecting the assumption of constant factor prices. Since the factor prices w and r are assumed to be constant, their ratio (w/r) which gives the slope of the isocost line will obviously remain constant. Here the conditions for equilibrium of the firm are:

(i) at the point of equilibrium, the given isoquant must be tangent to the lowest possible isocost line and

(ii) the isoquant must be convex to the origin (Diagram: 24).

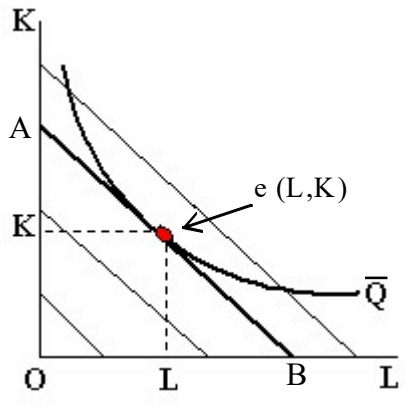


Diagram : 24

Unlike case-1, here the firm has to produce a given output (say $\bar{Q}$) at minimum possible cost. In the Diagram : 24, 'e' denotes the point of equilibrium which is the point of tangency between the isoquant representing the given level of output ($\bar{Q}$) and the lowest possible isocost line AB. It is desirable for the firm to reach isocost lines below AB as they represent lower cost than that shown by AB, but these remain unattainable because of the output constraint ($\bar{Q}$). On the other hand, points above 'e' or, to the right the isocost line AB represent higher cost than that indicated by point 'e' or, isocost line AB. Hence 'e' represents the point of minimum cost possible subject to the output constraint. The co-ordinate of the point 'e' i.e., (L, K) gives the least-cost combination of inputs for producing the given level of output .

From the above analysis it becomes clear that equilibrium conditions for both Case - 1 and Case - 2 are the same. Least-cost combinations of factor inputs for profit maximisation are determined by (i) the point of equality of the slopes of the isoquant and the isocost line and (ii) the isoquant must be convex to the origin.

1.5.2 Choice of optimal expansion path

We will discuss this problem under two heads - (i) Choice of optimal expansion path in the long run and (ii) Choice of optimal expansion path in the short run. You have to remember one significant point that out of all technically possible alternative paths of expanding output (defined by different product lines), what path will actually be chosen by a rational producer is determined by the prices of factors of production and is represented by the expansion path.

Optimal expansion path in the long run

The long run is featured by the variability of all factors of production. In order to serve its basic objective of maximisation of profit, the firm can expand its level of output without facing any constraint either financial or technical. It can employ any level of inputs to produce the desired level of output to ensure profit maximisation. Hence it is a case of unconstrained profit maximisation. Out of all alternative technical possibilities, the firm has to choose the optimal way of expanding output. Given the production function (represented by the isoquant map) and factor prices (w and r), the locus of the tangential points between successive parallel isocost lines and the successive isoquants depicts the optimal way of expanding the level of output in order to maximise the firm's profits. For given factor prices w and r , the successive isocost lines will be of constant slope equal to the factor price

ratio w/r and hence become parallel to each other. The expansion path in the long-run emanates from the origin. For homogeneous production function it takes the shape of a straight line. The slope of the expansion path determines the optimal factor input ratio (K/L) and it is itself determined by the ratio of the given factor prices.

Look at the Diagram : 25. It depicts the optimal expansion path as OA for the given factor prices w and r, which is obtained by joining the points of tangency between the successive isocost lines with constant slope w/r and the successive isoquants. If the factor price ratio changes (say from w/r to w′/r′) the slope of the isocost lines will also change and the optimal expansion path will be a different straight line emanating from the origin and passing through the tangential points of the successive isoquants and the new set of successive parallel isocost lines with the changed slope w′/r′ (OB in Diagram : 25).

For non-homogeneous production function, the optimal expansion path will not be a straight-line in spite of the assumption of given factor price ratio. As shown in the Diagram: 26, the optimal expansion path in this case will take a non-linear /curved shape. The reason is that, in equilibrium, w/r, i.e. the constant factor price ratio representing the slope of the isocost line must be equated to $MRS_{L, K}$, i.e., the slope of the respective isoquant which is the same on a curved isocline.

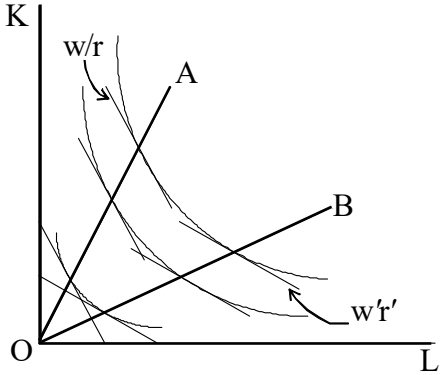


Diagram: 25-Optimal expansion path for homogeneous production function

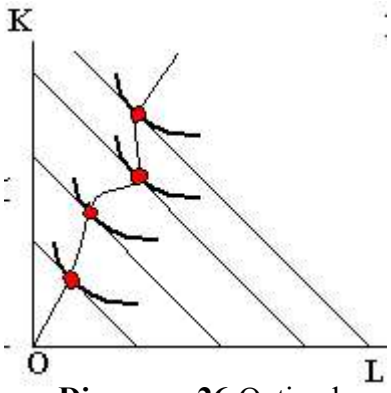


Diagram : 26-Optimal expansion path for non-homogeneous production function

Optimal Expansion Path in the Short -Run

In the short run all factors cannot be varied. Therefore in the short-run expansion of output can be carried on by using more of variable factor (here assumed to be the labour input, L) against the fixed factor (here capital input, K). Graphically, the expansion of

output is shown along a straight line which is parallel to the axis of the variable factor (L) and which is at the level of constant capital ($\bar{K}$). With the given factor prices, the profit of the firm cannot be maximised in the short-run because of the constraint set by the given level of capital . The Diagram : 27 shows that if it were possible to vary the level of capital input ($\bar{K}$) along with the labour input (L), the optimal expansion path would be OA. However, because of the constraint set by the fixed level of capital at $\bar{K}$ level, in the short-run the firm can expand only along $\bar{K}\,\bar{K}$.

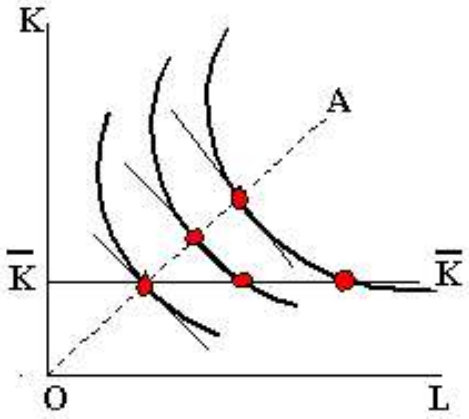


Diagram : 27

Check Your Progress-1.4

Note:

a) Use the space given below for your answer.

b) Compare your answer with the model answer given at the end of the unit.

1) The following statements are based on the text you have already read. Indicate whether these statements are true or false by putting a tick mark (✓) in the relevant box.

	True	False
(i) The slope of the expansion path in the long-run determines the optimal factor input ratio (K/L).	[ ]	[ ]

(ii) For non-homogeneous production function, the optimal expansion path will be a straight-line. [ ] [ ]

(iii) With the given factor prices and the given level of capital, the profit of the firm cannot be maximised in the short-run. [ ] [ ]

2) What are the equilibrium conditions for the problem of constrained profit maximisation? (Hint: see the Text).

1.6 DERIVATION OF COST FUNCTIONS FROM PRODUCTION FUNCTION

In this section you will learn about the derivation of cost functions from the production function. Given the production function, you will be able to derive graphically the cost curves from it. Let us illustrate the derivation of cost curves from the given production function with the help of a suitable example.

Here we will assume :

i) The production function is given (which implies that technology remains constant) with constant returns to scale.

ii) Factor prices are also given as follows :

w = Rs. 10 per labour hour

r = Rs. 10 per machine hour

Now let us consider that the existing technology of the firm includes the following six methods of production. As already explained in section-1.2, the methods of production refer to the combinations of the factor inputs labour (L) and capital(K) necessary to produce unit level of output.

	P_1	P_2	P_3	P_4	P_5	P_6
Labour : hours	2	3	3.5	4.5	5.5	6
Capital : hours	5	4.5	4.2	3.5	3	2.7

Now you can calculate the total cost associated with each of the above methods for the production of one 'unit' of output on the basis of given factor prices as follows:

	P_1	P_2	P_3	P_4	P_5	P_6
Labour cost	20	30	35	45	55	60
Capital cost	50	45	42	35	30	27
Total cost	70	75	77	80	85	87

It is obvious that, given the production function (that means constant technology) with constant returns to scale & given the factor prices, the first method P_1 is the least-cost method of production. Hence it is obvious that, a rational producer will opt P_1 method of production for all levels of output. Given the assumption of constant returns to scale, output levels can be increased according to the need just by increasing the input levels proportionately. In the following Table-1, some levels of output and their respective total cost are shown for the chosen least-cost method of production

P_1 .On the basis of the total cost of production of unit level of output for P_1 method (which is equal to Rs.70 as shown in the Table-1), the total cost for different levels of output can be calculated easily.

So far as the graphical representation is concerned, first we have to derive the product expansion path. It is obtained by joining the tangential points of the successive parallel isocost lines with the successive isoquants (Diagram : 28).These points of tangency provide necessary information on output and costs on the basis of which the total cost (TC) curve may be derived. Let us illustrate the point in some detail as follows :

at point a, Q = 5, TC = 350

at point b, Q = 10, TC = 700

at point c, Q = 15, TC = 1050

at point d, Q = 20, TC = 1400 etc.

TABLE - 1: Output Levels and Respective Total Costs (TC) and average Costs (AC) for P1 method production.

Output Q(in tons)	Total cost C(in rupees)	AC(Rs. Per ton)
0	0	0
5	350	70
10	700	70
15	1050	70
20	1400	70
25	1750	70
30	2100	70
35	2450	70
40	2800	70
45	3150	70
50	3500	70
55	3850	70
60	4200	70
65	4550	70

Now by plotting these points on a two-dimensional diagram with output (Q) on the horizontal axis and total costs (TC) on the vertical axis, you can have the total cost curve as shown in the Diagram : 29.

AC will remain constant for all levels of output (Rs. 70 per unit level of output) because of our assumption of constant returns

to scale and given factor prices. Obviously, the AC will be a straight line parallel to the output axis (Diagram : 30).

One notable point is that before defining the cost curves the problem of choice of the least-cost combination of factors of production must be solved.

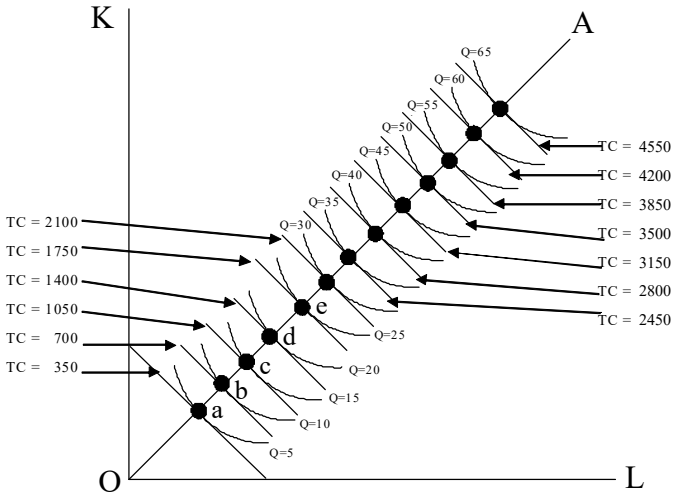


Diagram : 28-Product expansion path which is obtained by joining the tangential points between the successive isocosts and isoquants.

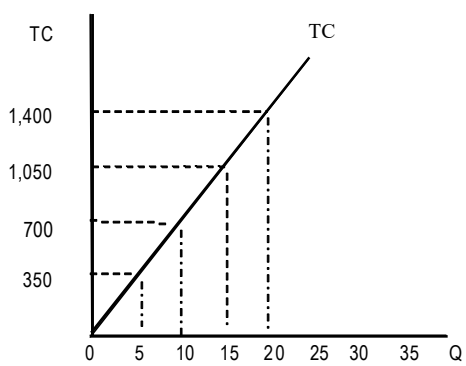


Diagram : 29
Total Cost (TC) Curve

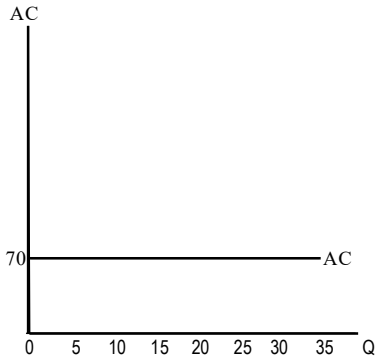


Diagram : 30
Average Cost (AC) curve

Check Your Progress-1.5

Note:

a) Use the space given below for your answer.

b) Compare your answer with the model answer given at the end of the unit.

1) The following statements are based on the text you have already read. Indicate whether these statements are true or false by putting a tick mark (✓) in the relevant box.

	True	False
(i) Cost curves are derived from the production function on the assumption of changing technology.	[ ]	[ ]
(ii) The problem of finding the least- cost input combination must be solved before defining the cost curve.	[ ]	[ ]

1.7 THE PRODUCTION FUNCTION OF A MULTIPRODUCT FIRM

In analyzing the case of a multiproduct firm we will consider a firm that produces only two products for technical convenience. The analysis can be extended to any number of products. The production - possibility curve and the iso-revenue curve are the two important concepts that are used in the determination of the equilibrium position of a multiproduct firm.

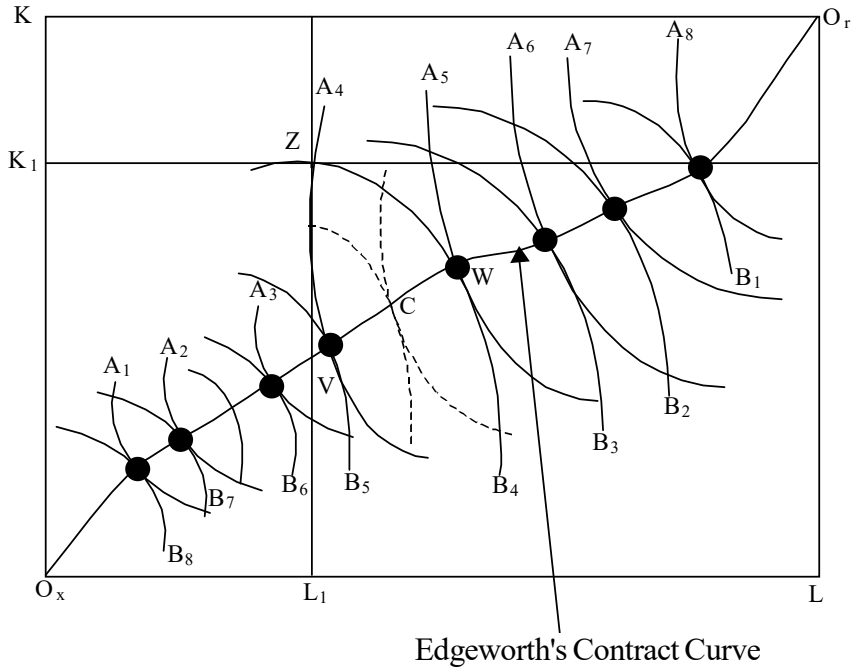
1.7.1 The Production - Possibility Curve of a Firm Producing Two Products X and Y

Let us consider that the firm produces two goods X and Y by using two factors of production labour (L) and capital (K). The production functions for the two products are given by :

$$X = f_1(L, K)$$
$$Y = f_2(L, K)$$

In Diagram : 31, the set of isoquants denoted by A (which are convex to the origin O_X) represents the production function for X while the set of isoquants denoted by B (which are convex to the origin O_Y) represents the production function for Y. Here we will derive the production possibility curve of the firm by using Edgeworth box diagram. Along the sides of the Edgeworth box,

the total quantities of the factors of production are shown. It is assumed that the firm has OL units of labour and OK units of capital available to it (Diagram : 31).



Digram : 31-Edgeworth's box

Any point in the Edgeworth box represents a certain combination of quantities of the two goods X and Y which are produced by employing the entire amount of the available factors L.C. i.e., OL amount of labour and OK amount of capital. By joining the tangential points between the two sets of isoquants, we obtain the contract curve. The importance of the contract curve lies in the fact that only the points upon the contract curve are efficient. It is because any other point off the contract curve represents a combination of the two products X and Y produced by employing all available resources, which contains less quantity of at least one product in comparison to the corresponding points on the contract curve. Let us explain with the help of the Diagram : 31. Let us assume that the initial point of production for the firm is 'Z'. The point 'Z' implies that using all available resources, the firm is producing A_4 level of X and B_4 level of Y. The firm utilises O_xL_1 units of labour and O_xK_1 units of capital for producing A_4 level of X and the remaining amount of the factors i.e., L_1L units of labour and K_1K units of capital are used to produce B_4 level of Y. The firm can produce more quantity of either good X or good Y or both of them just by reallocating its available resources so as to move to any point between V and W on the contract curve from point Z. Let us assume that by reallocating its available resources the firm chooses

to move to point W. At point W it is obvious from the diagram that the firm will be able to produce A_5 level of output of good X, which is higher than the A_4 level that was produced at the point Z and an equal level of output of good Y (B_4 level).Similarly a movement from Z to V makes it possible for the firm to produce a better combination of the two goods in comparison to point Z (B_5 level of output of Y, which is higher than the B_4 level that was produced at the point Z with an equal level of output of good X (A_4 level). If the firm chooses any point in between the range V and W on the contract curve, it will be able to produce higher levels of both the goods in comparison to point Z. For instance at point C, the firm achieves higher levels of isoquants for both the products X and Y as compared to point Z. What point on the contract curve will actually be chosen by the firm will be determined by the ratio of the prices of the two goods X and Y (i.e. by P_x / P_y). In order to determine this actual choice of the combination of commodity levels of X and Y, we have to derive the production possibility curve (in short PPC, which is also known as product transformation curve). PPC is the locus of points defining the commodity levels of X and Y which are produced by using all available resources of the firm. It is derived from the contract curve. Any point on the contract curve (i.e. each point of tangency between the two sets of isoquants) defines a combination of the levels of two goods X and Y and constitutes a point on the production possibility curve. For instance, point V on the contract curve (Diagram : 31) which represents the combination of A_4 level of X and B_5 level of Y is the point V′ in Diagram : 32. Point W on the contract curve (Diagram : 31) is represented by point W' in Diagram: 32. Thus plotting each point of the contract curve on the two dimensional plane where the two axes represent the commodities X and Y and taking the locus of these points, we obtain the production possibility curve (PPC) as shown in Diagram : 32.

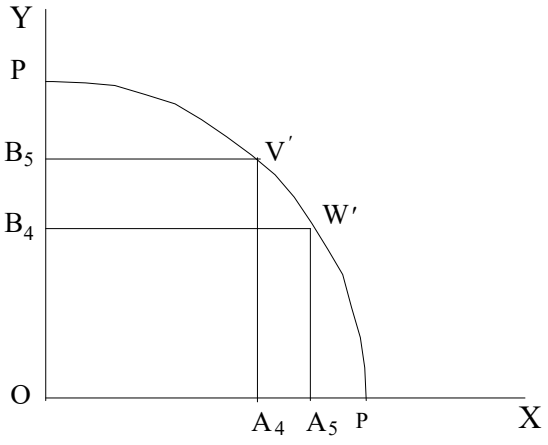


Diagram : 32- Production possibility curve

The slope of the production possibility curve is given by

$$-\frac{\partial Y}{\partial X} = MRPT_{X,Y}$$

Where $MRPT_{X,Y}$ means the marginal rate of product transformation. It can be shown that the slope of the production possibility curve (or, product transformation curve) is

$$-\frac{\partial Y}{\partial X} = \frac{MP_{L,Y}}{MP_{L,X}} = \frac{MP_{K,Y}}{MP_{K,X}}$$

where $MP_{L,Y}$, $MP_{L,X}$, $MP_{K,Y}$, $MP_{K,X}$ are the marginal products of the two factors in producing the two goods X and Y.

In order to find out graphically the equilibrium position of a multi-product firm, besides the PPC, you should also acquire the concept of the isorevenue curve.

THE ISOREVENUE CURVE OF THE MULTI-PRODUCT FIRM

An isorevenue curve is the locus of the points defining different combinations of quantities of X and Y, the sale of which yields the same revenue to the firm. Look at the Diagram : 33.

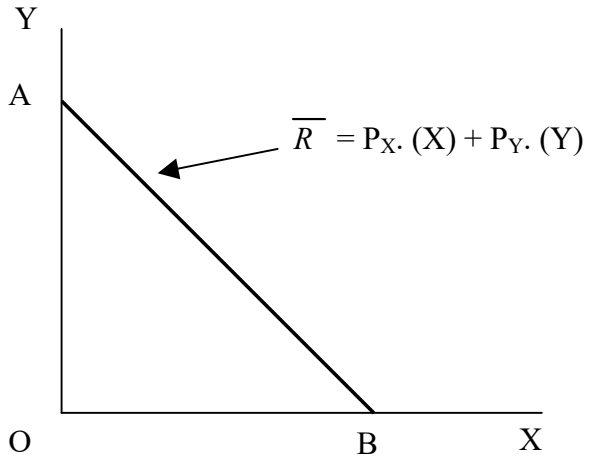


Diagram : 33 - Isorevenue curve

The point A on the line AB implies that the firm can earn total revenue $\overline{R}$ (represented by the line AB) by selling only good Y (i.e. a combination of OA units of good Y and zero unit of good X) .Similarly, The point B on the line AB implies that the firm can earn total revenue $\overline{R}$ (represented by the line AB) by selling only good X (i.e. a combination of OB units of good X and zero unit of good Y).Or, the firm can earn the same by choosing any point on AB, the coordinate of which will define a particular combination of the two goods X and Y.

The slope of the isorevenue curve is given by the ratio of the prices of the two commodities, P_X and P_Y respectively.

$$\text{i.e.} \quad \begin{bmatrix}\text{slope of the} \\ \text{isorevenue line AB}\end{bmatrix} = \frac{OA}{OB} = \frac{P_X}{P_Y}$$

You can prove it easily as you have done in case of the iso-cost line in the section-1.5. Let us first consider the point A. If we sell only commodity Y, the total revenue in terms of the Diagram : 33 will be (OA).$P_y = \overline{R}$,

$$\text{Or, } OA = \overline{R} / P_Y \qquad \text{---------- (i)}$$

Where OA is the quantity of Y which yields $\overline{R}$.

Similarly the quantity of X that yields $\overline{R}$ will be given by

$$OB = \overline{R} / P_x \qquad \text{---------- (ii)}$$

Dividing (i) by (ii) we will obtain the slope of the isorevenue curve as

$$\begin{bmatrix}\text{slope of the} \\ \text{isorevenue line AB}\end{bmatrix} = \frac{OA}{OB} = \frac{\frac{\overline{R}}{P_Y}}{\frac{\overline{R}}{P_X}} = \frac{P_X}{P_Y}$$

As the distance of the isorevenue curves from the origin increase, the total revenue of the firm will also increase.

1.7.2 EQUILIBRIUM OF THE MULTI PRODUCT FIRM

The firm's objective is to maximise its profit subject to the following constraints set by:

(i) the factors of production,
(ii) the transformation curve, and

(iii) the prices of the commodities (P_X and P_Y) and of the factors of production (w and r).

On the basis of the assumptions that quantity of the factors and their prices are given, the maximisation of π is achieved by maximising the revenue R. Graphically the equilibrium of the firm is defined by the tangential point of the given product transformation curve and the highest possible isorevenue curve as shown by point 'e' in Diagram : 34.

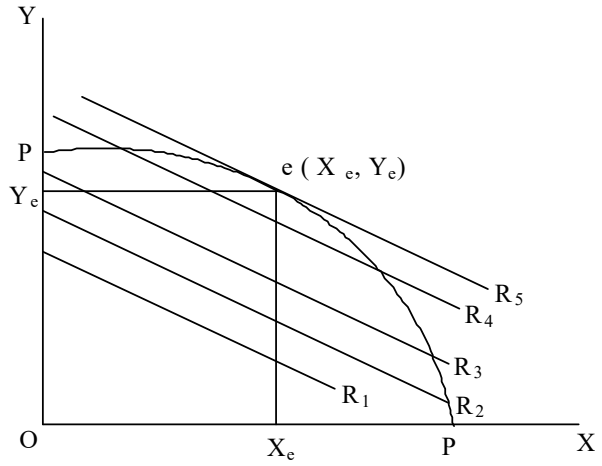


Diagram : 34-Equilibrium of the multiproduct firm

The coordinate of the point 'e' gives the optimal combination of the two goods X and Y as (X_e ,Y_e). At the point of the tangency 'e' the slope of the isorevenue line and the slope of the product-transformation curve must be equal. Thus, the condition for equilibrium of the multiproduct firm is given by:

Slope of the PPC= Slope of the isorevenue curve

$$\text{Or, } -\frac{dY}{dX} = \frac{MP_{L,Y}}{MP_{L,X}} = \frac{MP_{K,Y}}{MP_{K,X}} = \frac{P_X}{P_Y}$$

Check Your Progress-1.6

Note:

a) Use the space given below for your answer.

b) Compare your answer with the model answer given at the end of the unit.

1) The following statements are based on the text you have already read. Indicate whether these statements are true or false by putting a tick mark (✓) in the relevant box.

	True	False
(i)The points upon the Edgeworth's contract curve are inefficient in comparison to any other point off the curve.	[ ]	[ ]
(ii) Each point on the PPC gives the commodity levels of X and Y which use up all available resources of the firm.	[ ]	[ ]

(2) What is the equilibrium condition of a multiproduct firm ? (Hint: see the Text.)

1.8 LET US SUM UP

In the beginning of this unit, we have discussed some of the fundamental concepts related to the theory of production. Before dealing with the detailed discussions of the principal themes such as the laws of production and other complex derivations, we have explained certain basic concepts (mainly production function for a single product and other relevant concepts).

Next, we have examined the laws of production. In the short run output can be increased by using more of the variable factor(s) while one factor at least is kept constant. The law of variable proportions refers to the short-run analysis of production. In the long-run output may be increased by changing the levels of all the factors of production. This long-run analysis of production is provided by the laws of returns to scale.

Later, we have briefly dealt with the relationship between technological progress and production function. We explained graphically three different types of technical progress namely capital deepening technical progress, labour deepening technical progress and neutral technical progress as classified by Hicks.

Our next effort was to show the use of the production function in the choice of optimal combination of the factors of production. First, we have examined two cases of constrained profit maximisation in a single period and then considered the case of unconstrained profit maximisation by the expansion of output over time.

In the subsequent section we derived graphically the cost curves from production function. We have explained to you that with the assumptions of given production function with constant returns to scale and given factor prices, the tangential points of the successive parallel isocost lines with the successive isoquants provide necessary information on output and costs on the basis of which TC-curve could be derived. AC curve was found as a straight line parallel to the horizontal axis.

Finally we have dealt with the production function of a multiproduct firm .In this last section we have acquainted you with the concepts of production possibility curve and isorevenue curve and also found out the equilibrium condition for the multiproduct firm.

1.9 KEY WORDS

Production function: Production function is purely a technical relationship between outputs and factor inputs. It specifies the maximum output that can be produced with given inputs for a given

level of technology. The production function includes all the technically efficient methods of production.

Marginal product: Marginal product of each factor is defined as the change in output resulting from a very small change of the factor, other factors assuming constant for the time being.

Ridge lines: The two ridge lines are obtained by joining the points on the successive isoquants (the set of isoquants represents the given production function) where the marginal products of the factors are zero. Production methods are technically efficient only inside the ridge lines. Outside the ridge lines the marginal products of the factors are negative. The range of isoquants over which they are convex to the origin indicates the range of efficient production.

Marginal rate of substitution (MRS) : The absolute value of the slope of the isoquant is termed as the marginal rate of technical substitution or the marginal rate of substitution (MRS) of the factors. It is equal to the ratio of the marginal products of the factors. As a result of movement along an isoquant in downward direction the absolute value of its slope or, MRS, declines reflecting the fact that substitution of K for L become more and more difficult as we move downwards along an isoquant.

Product line : A product line shows the (physical) movement from one isoquant to another in response to a change in both factors or a single factor of production. The concept of product line is independent of factor prices. Hence, it only describes various technically possible alternative ways of expanding output.

Expansion path : Out of all technically possible alternative paths of expanding output (defined by different product lines), what path will actually be chosen by a rational producer is determined by the prices of factors of production and is represented by the expansion path. Given the production function and factor prices , the locus of the tangential points between successive parallel isocost lines and the successive isoquants depicts the optimal way of expanding the level of output i.e., the optimal expansion path in order to maximise the firm's profit.

1.10 SUGGESTED READINGS

1. Koutsoyannis, A., Modern Microeconomics, ELBS with Macmillan, London.
2. Maddala and Miller, Microeconomics, Tata Graw Hill.
3. Salvatore, D., "Microeconomics Theory and Applications", Oxford University Press, New Delhi.
4. Rubinfield & Pyndick, Microeconomics, 5th Edition, Pearson.
5. Ahuja, H.L., Advanced Economic Theory

MODEL ANSWERS

Check Your Progress-1.1

1) i) True ii) True iii) False

Check Your Progress-1.2

1) i) constant ii) diminishing iii) increasing iv)increases v) increasing.

Check Your Progress-1.3

1) i) False ii)True iii)True

Check Your Progress-1.4

1) i) True ii) False iii) True

Check Your Progress-1.5

1) i) False ii) True

Check Your Progress-1.6

1) i) False ii) True

UNIT 2: THEORY OF COSTS

STRUCTURE

2.0 OBJECTIVES

The main aim of this unit is to acquaint you with some very important aspects of the theory of costs. After studying this unit you should be able to:

- understand the meaning and significance of the cost function.
- define unit and total cost functions and to give their graphical representations.
- explain how the traditional theory of cost differentiates between short-run and long-run cost.
- understand how the average total cost, average variable cost and marginal cost are related to each other.
- know the causes of formation of the 'envelope' curve under the traditional theory of cost, interpret it and derive it graphically.
- discuss the causes of formation of the L-shaped planning curve under the modern theory of cost, interpret its significance and derive it graphically.
- differentiate real and pecuniary economies of scale
- understand the relevance of the shape of costs in decision making.

2.1 INTRODUCTION

Cost functions are derived from the production function. Total cost functions both in the short-run and in the long -run, are multivariable functions. The long-run cost function can be written as-

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

While the short-run cost function takes the form-

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

Where

C =Total cost

Q = Output

T = Technology

P_f = Prices of factors.

$\overline{K}$ = Fixed factor (s)

For graphical representation of the cost functions in two dimensional plane, we just consider that cost is a function of output alone i.e., C = f (Q); other determining factor (s) are assumed to be constant for the time being. If the factors other than output change, the impact of their changes is shown graphically by a shift of the cost curve.

2.2 THE TRADITIONAL THEORY OF COSTS

Let us first discuss the traditional theory of costs: how it differentiates the short run costs and the long-run costs. The short-run is characterised by the existence of some fixed factor (s) (such as capital input and entrepreneurship) while in the long-run all factors are assumed to be variable.

2.2.1 SHORT-RUN COSTS OF THE TRADITIONAL THEORY

Before going into detail analysis, let us first recollect some basic concepts of the theory of cost along with their graphical representation which are already known to you.

As said above, in the short-run some factor(s) is (are) fixed, that is can not be changed with the change in the level of output. Accordingly, the traditional theory of costs considers total costs (TC) to be composed of total fixed costs (TFC) and total variable costs (TVC) :

TC=TFC+TVC.

The fixed costs refer to the costs that a firm would incur even if its output for the period in question were zero. Examples of total

fixed costs are: salaries of administrative staff, interest payments, expenses for building depreciation and repairs etc. On the other hand, the variable costs refer to the costs that vary with the level of output. Examples include the cost of raw materials, labour costs, fuel costs etc. Diagram-1, Diagram-2 and Diagram-3 depict the shapes of TFC, TVC and TC-curves respectively.

Diagram -1 : Total Fixed Cost (TFC) Curve

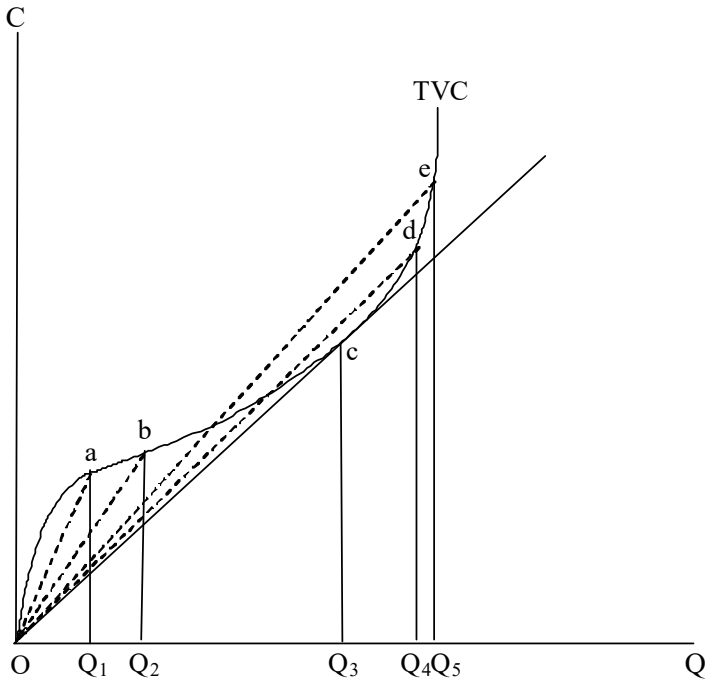


Diagram -2 : Total Variable Cost (TVC) curve

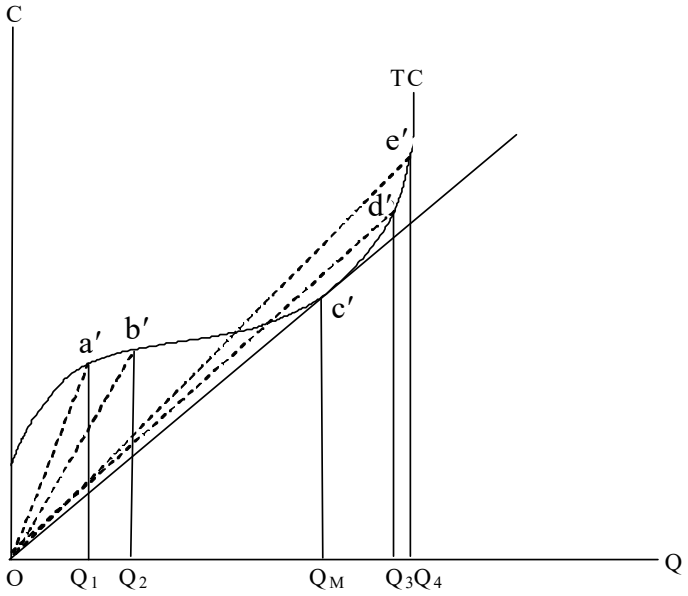


Diagram-3: Total Cost (TC) curve

The TFC-curve is represented by a straight line parallel to the output axis which means that TFC do not vary with the variation of output. TVC-curve is an inverse S-shaped curve. It shows that TVC is zero when output is zero and rise as output (Q) rises. Remember that the shape of the TVC-curve reflects the law of variable proportions. This law states that, at the initial stages of production, as quantity of the variable factor (s) is (are) increased to combine with the fixed factor (s), the productivity of the variable factor(s) increases and the average variable cost (AVC) declines until the optimal combination of the fixed and variable factors is reached. Beyond this point, any increase in the amount of the variable factor (s) to combine with the given level of fixed factor(s), causes a fall in the productivity of the variable factor(s) and AVC rises, TC- curve is obtained just by adding TVC and TFC curves. As TC = TFC + TVC , the TC- curve has the same shape as the TVC curve but lies above the TVC curve by a vertical distance equal to the amount of the TFC. The average fixed cost (AFC) is obtained by dividing TFC by the level of output

$$i.e., \qquad AFC = \frac{TFC}{Q}$$

The shape of AFC-curve is given by a rectangular hyperbola (Diagram-4). The shape means that every point on the curve gives the same value of TFC (given by AFC × Q). On the other hand average variable cost (AVC) is obtained by dividing TVC by the

level of output (Q), i.e. AVC=TVC/Q. Average total cost (AC) equals TC/Q. Also AC = AFC +AVC. Marginal cost (MC) implies the rate of change of TC or, TVC per unit change in output. Graphically, the AVC-curve is obtained at each level of output in terms of the slope of a line drawn from the origin to the point on the TVC curve corresponding to the given level of output (see Diagram-2 and Diagram-5). Similarly, the AC-curve is obtained at each level of output in terms of the slope of a line drawn from the origin to the point on the TC curve corresponding to the given level of output (see Diagram-3 and Diagram-6). At each level of output, MC is given by the slope of the TC or the TVC curve corresponding to the given level of output (see Diagram-7 and Diagram-8).

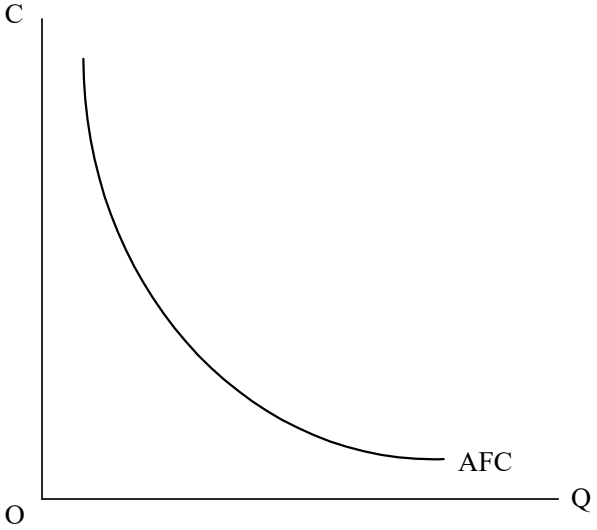


Diagram-4: Average Fixed Cost (AFC) curve of rectangular hyperbolic shape

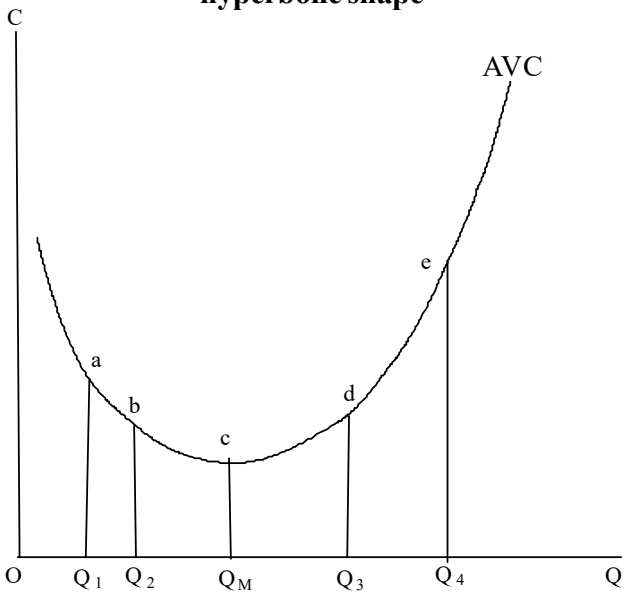


Diagram-5: U-shaped Short-run Average Variable Cost (AVC) curve.

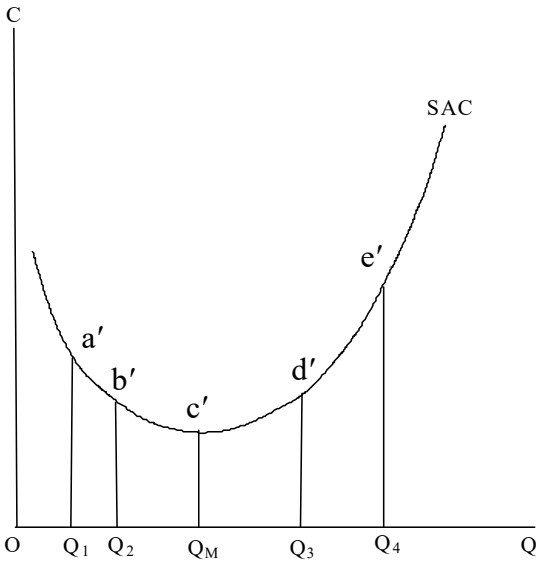


Diagram-6: U-shaped Short-run Average Cost (SAC) curve.

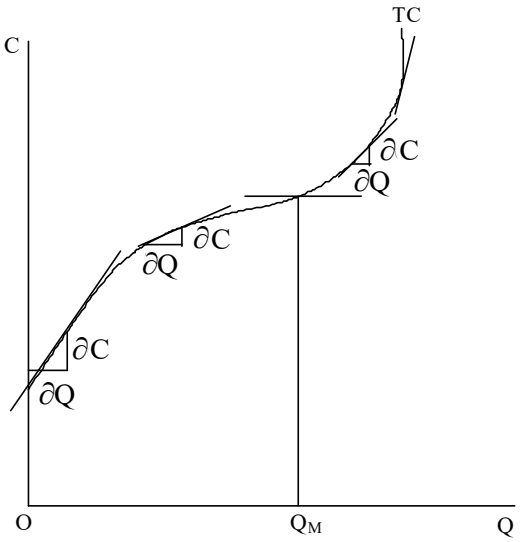


Diagram-7: Total Cost (TC) curve with different slopes at different points

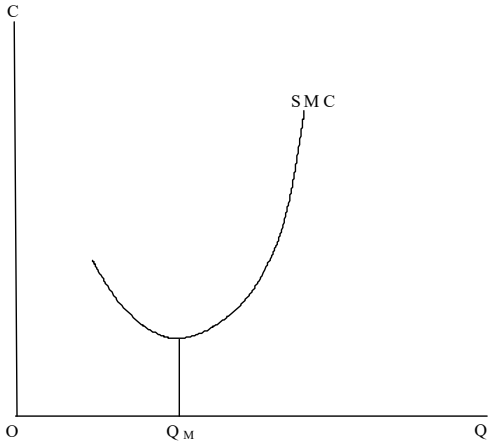


Diagram-8: Short-run Marginal Cost (SMC) curve.

Diagram -2 shows that the AVC corresponding to the level of output Q_1 is the slope of the ray Oa, the AVC corresponding to Q_2 level of output is the slope of the ray Ob and so on. The Diagram-2 also makes it clear that the slope of the successive rays falls continuously until the ray becomes tangent to the TVC-curve at point c and beyond c the slope of the rays starts increasing. Accordingly the short-run average variable cost (AVC) curve falls at the initial stage of production as the productivity of the variable factor (s) rises, then reaches a minimum point (point c in Diagram-5) where the plant is operated optimally making use of the optimal combination of fixed and variable factors, and beyond that point starts rising. You have to follow the same kind of interpretation in case of the graphical derivation of the AC curve from TC-curve (see Diagram-3 and Diagram-6). Mathematically the MC is the first order derivative of the TC-function, i.e.

$$MC = \frac{\partial C}{\partial Q}$$

You should know that the slope of a curve at any point on it is the slope of the tangent at that point. The Diagram-7 shows that the slope of the tangent to the TC-curve falls gradually for successive units of output until the slope becomes zero corresponding to the level of output Q_M and then begins to rise. Accordingly we get the U-shaped MC curve (Diagram-8). You can derive the MC from TVC curve graphically following the same method.

Thus, according to the traditional theory of costs, the short-run cost curves namely AVC, AC and MC are U- shaped. This shape follows directly from the law of variable proportions. It implies that, in the short-run, with (some) fixed factor(s), there is a phase of increasing productivity (which means falling unit costs) and a phase of diminishing productivity (which means rising unit costs) of the variable factor (s). Between these two phases of plant operation characterised by increasing and diminishing productivity of the variable factor (s), there exists a single point at which unit costs are minimum. At this minimum point (i.e., at the minimum point of the short run average total cost curve), the plant is said to be utilized optimally. That is, at the minimum point of the SAC-curve, the plant operates with the optimal proportions of the fixed and the variable factors.

The relationship between AC and AVC and MC:

Diagram-9 explains graphically the relationship between AC, AVC and MC. It is clear from the Diagram-9 that the minimum point of the U-Shaped AC curve occurs to the right of the minimum point of the U-shaped AVC curve.

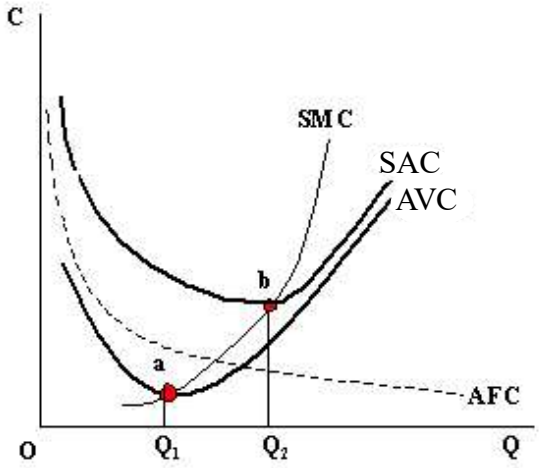


Diagram-9: Relationship among SAC, AVC and SMC curves.

The reason is also obvious from the diagram. The AFC-curve falls continuously with the increases in output. The AVC reaches its lowest point corresponding to the level of output Q_1 and then begins to rise. The rise in AVC over a certain range of output (from Q_1 level of output to Q_2 level of output) is offset by the fall in the AFC. Because of this fact the AC (which is the summation of AFC and AC) also falls continuously over that range of output (from Q_1 to Q_2) in spite of the increase in AVC. Beyond Q_2 level of output, however, the rise in the AVC more than offsets the fall in the AFC. Hence AC begins to rise beyond Q_2 level of output. As Q increases the AVC approaches AC asymptotically.

The relationship between the MC and AC:

The relationship between the MC and AC curves can be easily understood by using the tool of differential calculus.

Let us assume that total cost (C) is given by the following equation

$$C=bQ \quad \ldots\ldots\ldots\ldots\ldots \quad (i)$$

Where b implies AC and Q is level of output. Obviously b=f (Q), i.e., the AC is a function of output Q.

If we differentiate (i) with respect to output (Q) partially, we get

$$\frac{\partial C}{\partial Q}=\frac{\partial (bQ)}{\partial Q} \quad \text{----------}(ii)$$

Obviously other things remaining the same, $\frac{\partial C}{\partial Q}$ will give us the rate of change of total cost (C) with respect to the unitary change in the level of output (Q), i.e. MC.

Applying the product rule of differentiation (according to which if y = uv, then $\frac{\partial y}{\partial x} = u\frac{\partial}{\partial x}(v) + v\frac{\partial}{\partial x}(u)$ where u and v both are functions of x) to (ii) we get :

$$MC = \frac{\partial C}{\partial Q} = \frac{\partial (bQ)}{\partial Q} = b\frac{\partial Q}{\partial Q} + Q\frac{\partial b}{\partial Q}$$

$$or, \quad MC = b + Q\frac{\partial b}{\partial Q} \qquad \ldots\ldots\ldots\ldots\ldots \text{(iii)}$$

where, $\frac{\partial b}{\partial Q}$ is the slope of AC. The derivative $\frac{\partial b}{\partial Q}$ implies how b (i.e., AC) changes with unitary change in Q (i.e., output), Thus we can write (iii) as follows :

MC = AC + (Q) (Slope of AC) ..... (iv)

As Q>0 and AC>0, we can draw the following conclusions from the relation (iv):

a) If (slope of AC) <0, i.e., the slope of AC is negative (implying that AC falls when output rises), then we get, MC<AC.

b) If (Slope of AC)>0, i.e., the slope of AC is positive (implying that AC rises when output (Q) rises, then we get, MC>AC.

c) If (Slope of AC) = 0, (i.e., the change in AC with respect to changes in output is equal to zero), then MC=AC. The slope of AC becomes zero at the minimum point of the given U-shaped AC-curve (as shown by point 'b' corresponding to the level of output Q_2 in Diagram-9). Hence, at the minimum point of the AC-curve MC and AC become equal.

The Diagram-9 makes it clear that to the left of the point 'b', the MC-curve lies below the AC-curve and hence AC falls downwards for successive units of output (which implies that the slope of the AC-curve is negative to the left of point 'b'). To the right of the point 'b', MC-curve lies above the AC-curve and hence AC rises as output increases (which means that the slope of the AC-curve is positive to the right of 'b'). Hence, the point 'b', where the MC-curve intersects the AC-curve, is the minimum point of the AC-curve. Similar type of interpretation is applicable so far as the relationship between MC and AVC is concerned.

2.2.2 LONG-RUN COSTS OF THE TRADITIONAL THEORY: THE ENVELOPE CURVE

In this section we will discuss how under traditional theory

of costs, the long-run average cost curve (LAC) is derived from the short-run average cost curves (SAC).

Let us assume that the existing technology at a particular point of time allows the firm to work with three methods of production. Each method corresponds to a different plant size. In Diagram-10, the SAC_1, SAC_2 and SAC_3 represent the average cost curves of the small, medium and large size plants respectively. The firm will consider the small plant if it has to produce Q_1 level of output. Similarly in order to produce Q_2 and Q_3 levels of output the firm will choose the medium plant size and large plant size respectively. Let us assume that the firm starts with Q_1 level of output. If the demand for the product increases further, the firm will be able to produce successive units of output with lower cost using the small plant size up to the level of output Q_M corresponding to the minimum point of SAC_1 curve with which the small plant size operates.

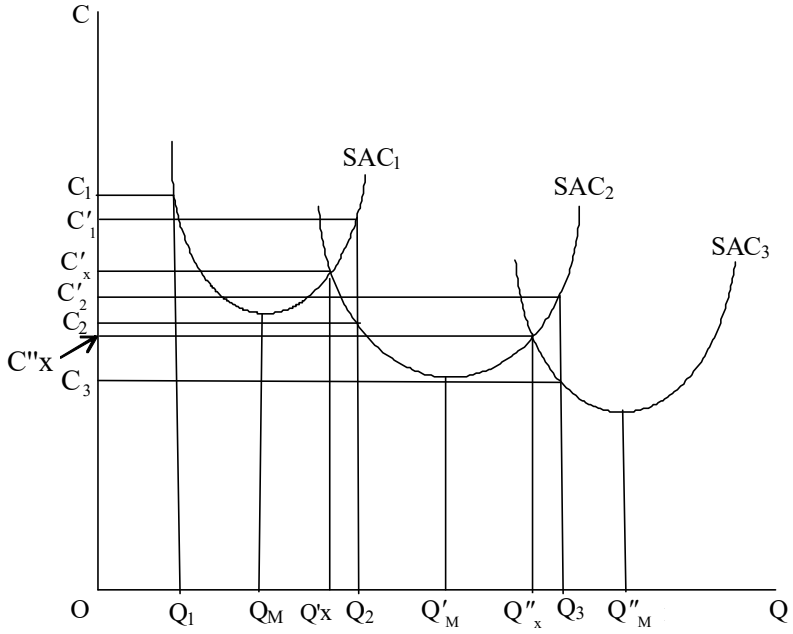


Diagram-10

Points beyond Q_M imply higher costs. If demand for the product increases up to the level Q'_x, the firm can work with either the same small plant size or can install the medium plant size. If the firm's expectation regarding the future demand is such that it will cross the limit of Q'_X, only then the firm will install the medium size plant. It is because operating with SAC_2 (i.e., with medium plant size), output levels greater than Q'_x can be produced with lower costs in comparison to the small plant size up to a certain level. On the other hand if the firm expects that future market demand for its product will not increase beyond Q'_x, then it will continue to work with the small plant. Because installation of a new plant implies a larger investment and it will be profitable only

if demand cross the limit of Q'_x. Similar considerations are applicable for the decision of the firm whether to continue with the medium size plant or to install the large one when it arrives at the level of output Q''_x.

The traditional U-shaped cost curves assume that each plant size is designed to produce a single level of output optimally. This point of optimal production refers to the minimum point of the U-shaped cost curves. Any departure from this point in any direction (i.e., either an increase, or, a decrease from the optimal level of output) results in increased costs. In this sense, the plant is quite rigid or inflexible. No reserve capacity is assumed and hence no adjustment of production can be made even to face seasonal fluctuations in demand for the product. Because of this assumption, the LAC-curve 'envelopes' the 'SAC' curves. Each point of the LAC-curve is a tangential point with the corresponding. SAC curve. Different plant sizes operate with different SAC curves. The point of optimal production corresponding to a particular plant size refers to the minimum point of the SAC -curve with which it operates. Each SAC curve represents the plant size to be used to produce a particular level of output at minimum costs. The LAC curve is then the tangent to these SAC-curves and each point of this curve shows the minimum (optimal) cost for producing the corresponding level of output. You can define the LAC curve to be the locus of points denoting the least cost of producing the corresponding level of output. While deriving graphically the LAC-curve, you have to be careful about certain points. To the left of the minimum point M of the LAC- curve, since the slope of the LAC curve is negative, the slope of the SAC curves must also be negative at the tangential points. It is because of the fact that at the tangential points the slope of both the curves must be equal. Hence to the left of M, as Diagram -11 explains, the tangential points occur to the falling part of the SAC-curves, as the LAC falls.

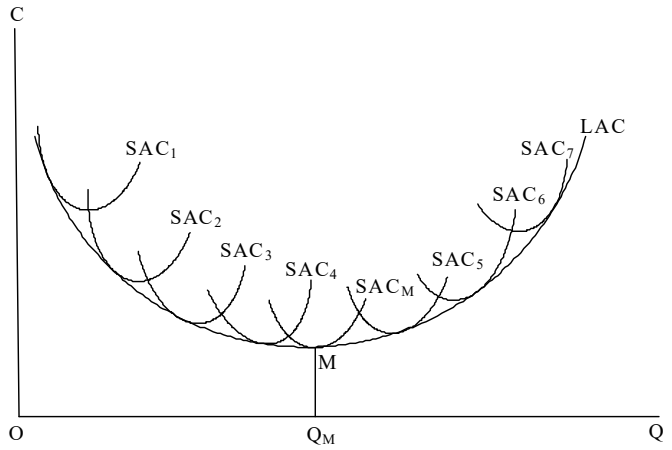


Diagram-11: The Long-run Average Cost (LAC) curve under traditional theory which envelopes the SAC-curves. (The Envelope curve.)

On the other hand, for output levels greater than Q_M, as the LAC-curve rises, the points of tangency must occur to the rising part of the SAC-curves. Only at the minimum point (M) of the LAC-curve, the corresponding SAC-curve also reaches its minimum point indicating the optimal use of the plant. The analysis makes it clear that to the left of the minimum point M of the LAC-curve, i.e., on the falling part of the LAC, the plants are not worked with full capacity or, are not used optimally; on the other hand, on the rising portion of the LAC, the plants are over worked. Only at the minimum point M the plant is used optimally.

Now let us explain how the long-run marginal cost (LMC) curve is derived from the short-run marginal cost (SMC) curves. The LMC-curve is the locus of the points of intersection between the SMC-curves and the vertical lines upon the output axis drawn from the points of tangency of the corresponding SAC-curves and the LAC-curve (Diagram-12). It means that the LMC must be equal to the SMC for the output levels at which the corresponding SAC curve is tangent to the LAC-curve.

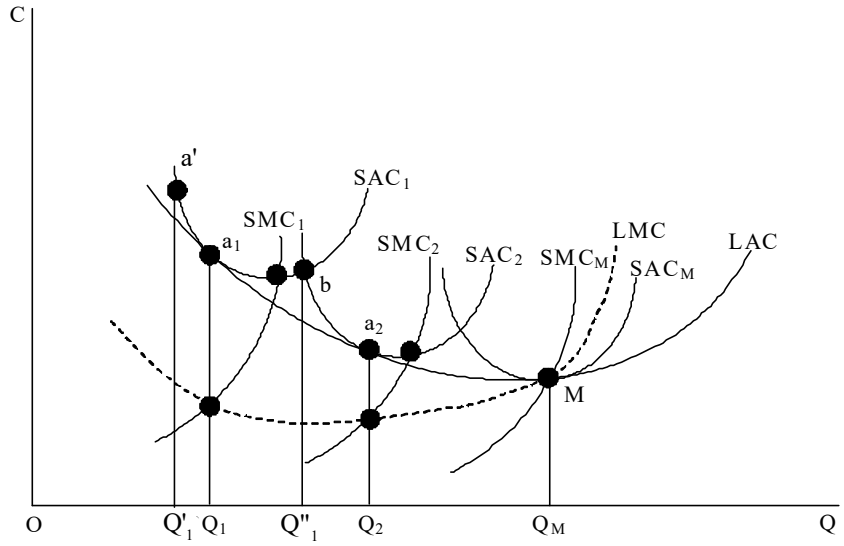


Diagram-12: Derivation of Long-run Marginal Cost (LMC) curve

For the levels of output to the left of the minimum point M of the LAC-curve the LMC curve lies below the LAC-curve and for the levels of output to the right of M , the LMC curve lies above the LAC-curve. The LMC-curve intersects the LAC curve at its minimum point M and hence at M, the LMC becomes equal to LAC. Thus, at the minimum point M of the LAC-curve, we have

SAC = SMC= LAC = LMC.

If as the limiting case we assume that the available technology allows very large number of plants, a continuous curve is obtained which will be the LAC-curve of the firm. It is a U- shaped curve.

As the LAC-curve under traditional theory envelopes the SAC-curves, it is often termed as the 'envelope' curve. LAC-curve is also said to be a planning curve, because considering it as a guide, the firm can decide what particular short-run plant is to be installed in order to produce optimally (i.e., at least possible cost) the anticipated level of output (in the long run). The U-shape of the LAC-curve reflects the laws of returns to scale according to which as plant size increases, the unit costs of production decrease because of the economies of scale generated by larger plant sizes. The economies of scale is assumed to exist only up to a certain size of plant. This particular plant size is termed as optimal plant size because of the fact that with this plant size all possible economies of scale are fully absorbed. Beyond this optimal plant size, diseconomies of scale arise due to managerial inefficiencies.

Check Your Progress- 2.1

Note:

a) Use the space given below for your answer.

b) Compare your answer with the model answer given at the end of the unit.

1) Match the following and give the correct answer from the code given below:

A)	If the (slope of AC) <0	a)	then MC > AC
B)	If the (slope of AC) >0	b)	then MC < AC
C)	If the (slope of AC) =0	c)	then MC £ AC
		d)	then MC = AC

Code:

	(A)(B)	(C)	
(i)	(a)	(b)	(c)
(ii)	(b)	(c)	(a)
(iii)	(b)	(a)	(d)
(iv)	(b)	(a)	(c)

2) The following statements are based on the text you have already read. Indicate whether these statements are true or false by putting a tick mark (✓) in the relevant box.

		True	False
i)	Graphically the AFC is a rectangular hyperbola	[ ]	[ ]
ii)	The U-shape of the short-run cost curves under traditional theory reflect the laws of returns to scale.	[ ]	[ ]
iii)	The U-shape of the long-run average cost curve under traditional theory reflects the law of variable proportions.	[ ]	[ ]
iv)	A serious implicit assumption of the traditional U-shaped cost curves is that each plant size is designed to produce optimally a single level of output.	[ ]	[ ]
v)	The 'envelope curve' is a planning curve.	[ ]	[ ]

2.3 MODERN THEORY OF COSTS

Like the traditional theory, modern theory of costs also differentiates between short-run and long-run costs. But so far as the shape of cost curves are concerned, there exist certain differences between the two theories. The modern theory of costs suggests that the LAC-curve is L-shaped rather than U-shaped. In modern theory of costs, the short-run average variable cost (AVC) has a saucer-type shape which implies that firms set up plants with some flexibility in their productive capacity. The empirical evidence is also in support of the view that there are no diseconomies of scale at large scale of output. However, no final conclusion has been drawn whether costs remain constant beyond a certain minimum optimal scale, or fall continuously with the expansion of output.

2.3.1 THE SHORT-RUN COSTS UNDER MODERN THEORY

The short-run total cost (TC) is composed of total variable costs (TVC), total fixed costs (TFC) and normal profit. The TVC include labour costs, costs of raw materials, running expenses of machinery etc.TFC refer to salaries of administrative and production staff, fixed expenses of plant, depreciation cost of fixed capital etc.

Let us first explain the shape of AVC-curve according to the recent developments in microeconomic theory. According to the modern theory the AVC has a saucer type shape. It has a flat stretch over a certain range of output. This flat stretch is due to the planned reserve capacity which is built in the plant while designing it. This reserve capacity provides the firm maximum flexibility in operation over a certain range of output (Q_1Q_2- range of output in Diagram-14).

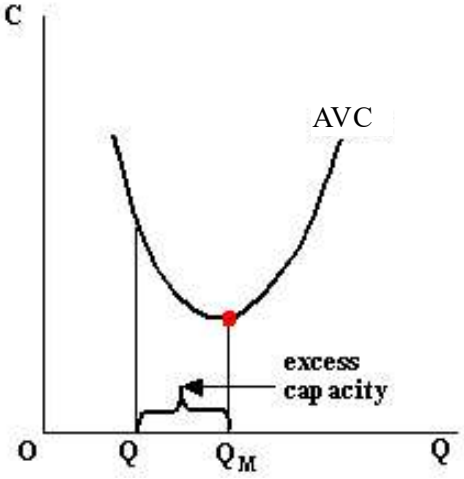


Diagram-13: Excess capacity which arises with the U-shaped AVC curve of the traditional theory

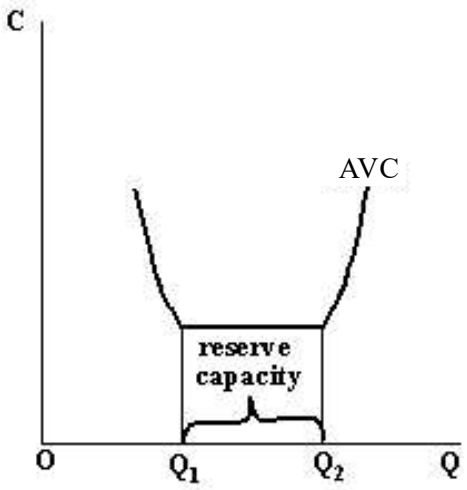


Diagram-14: Reserve capacity that arises with the saucer shaped AVC curve of the modern theory of cost

It means that within that range, the firm can increase its level of output without any rise in cost to meet any increase in demand. You have already learned that according to the traditional theory of costs each plant size is designed to produce optimally only a single level of output which implies zero flexibility in production (Diagram- 13). If the firm produces Q units of output, there is excess capacity equal to the difference (Q_M -Q). Unlike reserve capacity, the excess capacity is not desirable for the firm as it leads to higher unit costs. But the planned reserve capacity (represented by the range of output Q_1Q_2 in Diagram-14) does not lead to increase in costs. In general the firms' expectation is to operate in between two-thirds and three quarters of their capacity. In terms of the range Q_1Q_2, obviously the firm expects to utilise its plant at a level which is closer to Q_2 than Q_1 level of output. This level of utilisation of the plant, which is considered to be normal by the firms, is termed as 'load factor' of the firm.

Let us now explain the shape of the SAC-curve under modern theory. AC is obtained by adding the AFC (including the normal profit) and the AVC at each level of output. As shown in the Diagram-15, the SAC curve falls continuously with the expansion of output up to Q_A level of output at which the reserve capacity is fully exhausted. If output rises further beyond Q_A, SAC will also increase with that. The MC will intersect the SAC-curve at its minimum point. You must remember that the minimum point of SAC occurs to the right of Q_A, the level of output which corresponds to the end point of the flat stretch of the saucer-shaped AVC-curve.

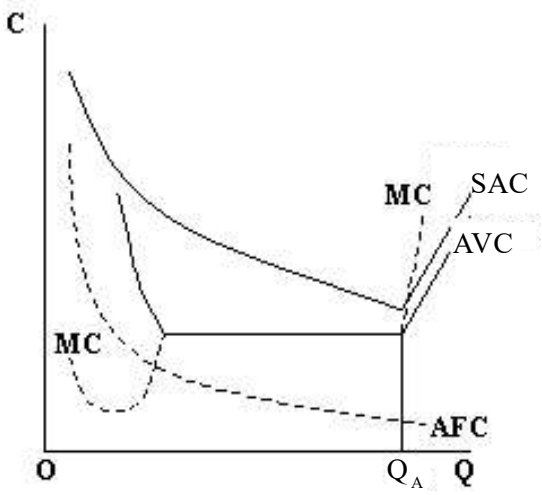


Diagram-15: Relationship among AVC, SAC and MC curves under modern theory of cost

Over the flat stretch representing the planned reserve capacity the AVC is equal to the MC; both remain constant with the expansion of output. To the left of the range of reserve capacity, MC lies below the AVC and to the right of the range of reserve capacity, MC lies above AVC. Up to the Q_1 level of output the AVC falls continuously with the expansion of output (Diagram-14) because of (i) the better utilisation of the fixed factor combined with the increasing level of variable factor, (ii) improvement in skills and productivity of the variable factor labour, which will ensure reduction in waste of raw materials and better utilisation of plant. On the other hand, the increasing trend of the AVC to the right of the flat stretch is seen due to the fall in labour productivity which is caused by longer hours of working, rise in cost because of overtime payment, wastage of raw materials and frequent breakdown of machinery caused by over utilisation of the plant.

2.3.2 LONG-RUN COSTS IN MODERN THEORY: THE L-SHAPED SCALE CURVE

One significant outcome of the modern theory of costs is that the LAC-curve is almost L-shaped, not U-shaped as suggested by the traditional theory of costs. Two cases may occur:

(i) The LAC may decline rapidly up to a certain level of output and beyond that level become constant throughout and

(ii) It may fall continuously (though the rate of falling becomes smooth at very large scales of output)

Production and managerial costs falls continuously with the

expansion of output. Though at very large scale of output managerial costs may show a slightly increasing trend, the long-run average total cost (assumed to be composed of average production and average managerial cost) does not rise. It is because of the fact that the rate of fall of production costs is more than sufficient to nullify the rising trend in managerial cost even at very large scales of output.

Let us now explain in some detail how the variation in production and managerial costs contribute to the L-shape of the LAC-curve.

(i) PRODUCTION COST : Initially the production cost falls at a much higher rate with the expansion of output (because of which we get the comparatively steep portion of the L-shaped LAC-curve). Even for very large scale of output production costs continues to fall but the rate of fall become slower (which leads to the comparatively flat portion of the L-shaped LAC-curve).Unit cost of production falls continuously because of some technical economies of large scale production. At the initial level of expansion of output these economies are enjoyed substantially by the firm. But if the firm reaches a minimum optimal scale which defines a certain level of output at which all or most economies of scale are exhausted (assuming constant technology), the LAC remains constant with expansion of output beyond that scale. If technical progress occurs and makes it possible for the firm to produce large scale of output with new better techniques, it will obviously ensure lower unit cost of production for the firm. However, even in the absence of any technical progress, using the techniques allowed by the existing technology certain economies can always be enjoyed by the firm for large scale of production. Examples are:

(i) Economies generated by further decentralization and improvement in skills.

(ii) Economies from lower repair cost that may be attained after achieving a certain scale.

(iii) Economies from the ability of the firm to produce certain raw materials or equipment at lower costs which are needed for its own production instead of purchasing from other firms.

Managerial costs: Recent developments in management science makes it possible to overcome most of the managerial diseconomies which were said to be responsible for the rising part of the U-shaped LAC under traditional theory of costs. Smooth, efficient operation of each plant size now can be ensured by suitable organizational and administrative set-up.

For different sizes of plant, different appropriate management techniques are made available by the modern developments in

management science. To begin with, the cost of various management techniques falls up to a certain plant size. Only at very large scale a gradual rising tendency in management costs may be seen.

But production economies more than offsets the managerial diseconomies that may appear at very large scale of production. Hence, the LAC-curve falls smoothly or remains constant (after reaching the minimum optimal scale) even at very large scales of output.

Let us now explain the procedure of drawing the LAC-curve under modern theory of cost. First, let us assume that the existing technology offers four different plant sizes. Higher the plant size lower the cost associated with it. For example, the smallest plant size operates with SAC_1 while SAC_4-curve corresponds to the largest plant size out of the four available plant sizes. Remember one particular point which you have already learned: usually firms consider that the 'normal' level of utilization of their plant (which is termed as the load factor of the plant) lies somewhere in between two-thirds and three-quarters of their capacity. Let us consider that the load factor of each plant in this case is two-thirds of its total capacity. Then you just have to join the points on the SAC-curves corresponding to the two-thirds of the full capacity of each plant size to get the LAC-curve. If instead of four plant sizes, we consider that the existing technology includes a very large number of plant sizes, the LAC-curve will be a continuous curve as shown in Diagram-16.

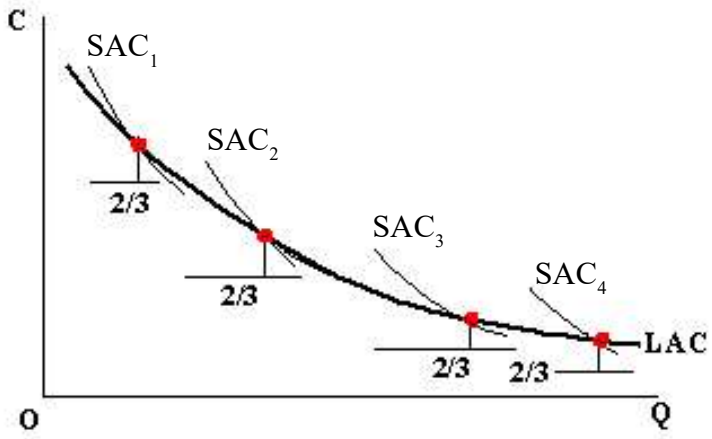


Diagram-16: Long-run Average Cost (LAC) curve under modern theory of costs which is drawn by joining the points on the short-run average cost curves corresponding to the two thirds of the full capacity (representing the load factor) of each plant size.

The main features of the LAC-curve under modern theory which can differentiate it from U-shaped LAC-curve of the traditional theory are that:

(i) it has no upward rising portion like the U-shaped LAC-curve even at very large scale of output; and

(ii) it does not envelope the SAC-curves (which the U-shaped LAC-curve does), rather intersect them at the points which correspond to the levels of output defined by the given load factor.

You have to remember one significant point: if the LAC-curve falls continuously throughout (rapidly to begin with and then smoothly at very large scales of output),the long-run marginal cost (LMC) curve will lie below the LAC at all scales (Diagram-17). On the other hand if the firm reaches a minimum optimal plant size ($\bar{Q}$ in Diagram-18) which implies that most or all possible economies are exhausted, beyond that scale the LAC become a straight line indicating a constant value for all levels of output. The corresponding LMC-curve then will lie below the LAC until the minimum optimal scale is reached (at $\bar{Q}$) and beyond that level merge with the LAC as depicted by Diagram-18.

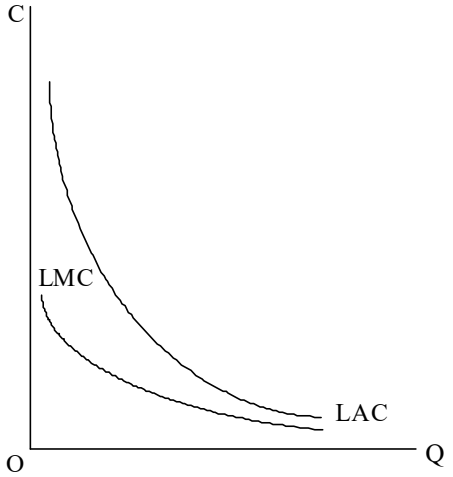


Diagram-17: LAC falls continuously and LMC lies below LAC for all levels of output.

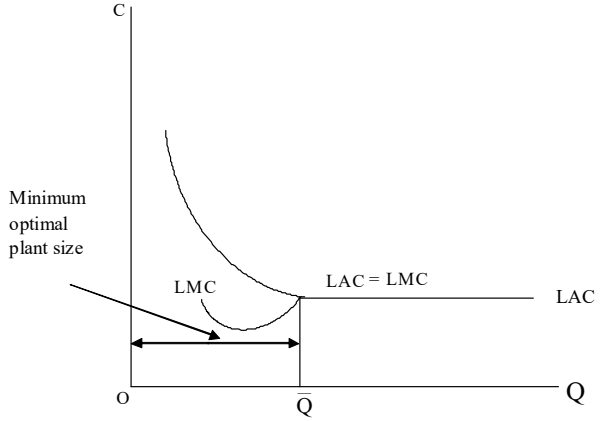


Diagram-18: LMC lies below LAC until (the minimum optimal scale is reached and beyond that level coincides with LAC.

Check Your Progress- 2.2

Note:

a) Use the space given below for your answer.

b) Compare your answer with the model answer given at the end of the unit.

1) After reading thoroughly the preceding text complete the following statements:

i) Under modern theory of cost the AVC-curve is..... .....shaped.

ii) Excess capacity that arises with the U-shaped costs under traditional theory is obviously undesirable because it leads to.....

iii) The flat stretch over a range of output of the AVC-curve under modern theory corresponds to the built-in-the-plant.....

2) Choose the correct option and put a tick mark (✓) in the relevant box.

	Option (a)	Option (b)
i) Under modern theory of cost theLAC-curve		
a) envelopes the SAC-curves.	[ ]	[ ]
b) intersects the SAC-curves.		
ii) Most of the empirical studies on cost have provided evidence which substantiates the hypotheses of		
a) a U-shaped AVC and a U-shaped LAC-curve.	[ ]	[ ]
b) a flat-bottomed AVC and an L-shaped LAC-curve.		

3) What is meant by the 'load factor' of a plant ? (Hint: see the Text.)

---

4)What are the main features of the LAC-curve under modern theory of cost ? (Hint: see the Text.)

2.4 THE ANALYSIS OF THE ECONOMIES OF SCALE

The economies of scale refer to both internal and external economies scale. The internal economies of scale are gained by the firm through its own action when it expands its output level in the long-run and are built in to the shape of the long-run cost curves. External economies on the other hand are independent of the actions of the firm. The changes of the environmental condition in which the firm operates, give rise to external economies. The firm in question can realize these economies from actions of other firms in the same or in different industries. If the external economies are realized in the form of a change in the factor prices and /or a change to the production function (improvement in the techniques of production), then both short-run and long-run cost curves will shift. We will discuss mainly about the internal economies of scale. Emphasis will be given on the intra-plant economies, i.e., economies that may be achieved within a particular plant.

Economies of scale are broadly classified into real and pecuniary economies of scale. Real economies are generated by a reduction in the physical quantity of inputs-raw materials, different types of labour and capital etc. Real economies of scale may be of different forms: production economies, selling or marketing economies, managerial economies, transport and storage economies etc. Different factors give rise to different types of economies. For instance, specialisation of the capital equipment for large scales of production and indivisibilities of modern industrial techniques of production results in technical economies of scale. Labour economies arise from division of labour and specialization of labour force for larger scale of production. These measures improve the skills of labour and thus accelerate productivity of labour in case of large scale production.

It is asserted that, larger the output the smaller the advertising cost per unit. It gives rise to selling or marketing economies of scale. Also, costs on other types of selling activities for large scale promotion, such as the distribution of samples etc. increase but less than proportionately with the level of output at least up to a certain scale. Managerial economies mainly arise from specialization of management and mechanization of managerial function. Again it is obvious that storage costs fall with size. Unit transportation cost would fall if the firm uses its own vehicles for transportation of materials and output up to the point of their full capacity. Pecuniary economies of scale are realized by the firm because of the discounts that it can obtain due to its large scale

operations. These economies are enjoyed by larger firms in the forms of (i) lower prices for bulk-buying of raw materials, (ii) lower cost of finance (for instance, banks usually offer loans to large corporations at a lower rate of interest), (iii) lower advertising and other selling costs per unit , (iv) lower transportation costs that are realized when larger quantity of materials and output are transported (v) lower wages and salaries due to monopolistic power attained by a large firm, or, due to the prestige which is assumed to rise with the association with large firms. The total average cost is the summation of all these costs of production, marketing, managerial, transport etc. It is generally accepted that the total LAC-curve will fall with the increase of scale of the plant (and of the firm), at least up to a certain plant (or firm) size. Disagreement still continues among economists regarding whether

(i) there are diseconomies at very large scales of output (which gives rise to the traditional envelope-shaped LAC);

(ii) there is a minimum optimal scale of output at which all possible economies of scale have been absorbed so that beyond that level the LAC become constant (the case of L-shaped LAC);

(iii) there are economies of scale at all levels of output though beyond a certain scale they are realised in smaller amount (the case of inverse J-shaped cost curve).

Check Your Progress- 2.3

Note: Use the space given below for your answer.

1) Write a short note on real economies of scale. (Hint : see the Text)

---

2) 2) Mention the main factors that give rise to pecuniary economies of scale for larger firms. (Hint: see the Text)

2.5 THE RELEVANCE OF THE SHAPE OF COSTS IN DECISION-MAKING

In the process of decision making by the firm as well as the government, the knowledge of cost functions plays a very crucial rule. For example, short-run costs are important for pricing and output decisions. On the other hand, long-run costs play the role of a guide of the entrepreneur by providing necessary information for proper planning of the growth and investment policies of the firm.

In all market structure costs are one of the main determinants of price and output. For instance, the pure competition model is valid only for U-shaped cost curves. Otherwise, the optimal output become indeterminate [as AR (=MR) is given by a straight line parallel to the output axis].In monopolistic competition the shape of costs does not possess such special significance; the size of the firm is determinable provided that the slope of the MC-curve is smaller than the slope of the MR-curve. But costs affect the price-output decision of the firm explicitly both in the short-run and in the long-run as the profit-maximising position is determined at the level where MC=MR. The monopolist sets his equilibrium price at the level where the MC-curve intersects the MR-curve from below. In the traditional theory of price leadership, the firm with lowest costs plays the role of a price leader. Costs and the concept of barriers to entry are also closely related. The lower the costs of production for all levels of output and the larger the minimum optimal scale of output i.e. greater the economies of scale, the higher the entry barriers. Therefore, the established firms in an oligopolistic industry can set the limit price at a higher level above the competitive price level without attracting the potential entrants.

If the market size is known, the direction of growth of a firm is primarily influenced by costs. For instance, LAC-curve under traditional theory is said to be a planning curve; because considering it as a guide, the firm can decide what particular short-run plant is to be installed in order to produce optimally a particular level of output. After reaching the minimum point of the U-shaped LAC-curve, i.e., after absorbing all possible economies of scale, the expectation of further expansion of output in the same market to meet the increased market demand makes it necessary to set up a new plant of adequately higher size. If the market becomes stagnant, the firm facing a U-shaped scale curve will try to make investment in other markets. Adequate knowledge of costs is required for the regulation of industry by the government. On the basis of the information on the costs of various firms, the regulatory authorities can decide whether to split up large firms, to encourage or prohibit mergers etc.

Check Your Progress: 2.4

Note:

a) Use the space given below for your answer.

b) Compare your answer with the model answer given at the end of the unit.

1) After reading thoroughly the preceding text complete the following statements:

i) The pure competition model breaks down unless cost curves are .....

ii) The lower the costs of production for all levels of output and the larger the minimum optimal scale the ..... the entry barriers.

iii) Given the market size, the direction of growth of a firm is primarily determined by .....

2) Write short note on the relevance of the shape of costs in 'Decision Making Process'.(Hint: see the Text)

2.6 LET US SUM UP

In the beginning of this unit, we have discussed some of the fundamental concepts related to the theory of costs. First we have examined the traditional theory of cost. In the section 2.2 we have mainly dealt with the graphical derivation of the unit cost curves and with their mutual relationship. You have learned about the formation and significance of the U-shaped LAC which is also termed as the envelope curve. We have also explained to you why it is termed as the planning curve. Finally in section 2.2 we derived graphically the LMC-curve and learned about its relation with LAC-curve. Next, we have examined the modern theory of costs in section 2.3, explored its differences with the traditional theory of cost mainly with respect to the shape of AVC and LAC curves. We have learned

that most of the empirical studies on cost provided evidence in support of a flat bottomed AVC and L-shaped LAC as suggested by the modern theory of cost. Next, we have dealt with the analysis of economies of scale and concentrated upon real and pecuniary economies of scale. At last, we have discussed about the relevance of the shape of costs in decision making.

2.7 KEY WORDS

Excess capacity : The difference between the output level indicated by the lowest point of the U-shaped cost curve and the output actually produced by a monopolistically competitive firm.

Economies of scale : Increases in productivity, decreases in average cost of production, which arise from increasing all the factors of production in the same proportion.

Minimum optimal scale: The smallest level of output at which the long-run average cost is at a minimum; the smallest output required to achieve all economies of scale in production.

Envelope curve: The U-shaped Long-run Average Cost curve which envelopes the short-run average cost curves under traditional theory of costs ; it is a planning curve.

L-shaped scale curve: The L-shaped Long-run Average Cost curve which is formed by joining the points on the Short-run Average Cost curves corresponding to the typical load factor of each plant size.

2.8 SUGGESTED READINGS

1. Koutsoyannis, A., Modern Microeconomics, ELBS with Macmillan, London.
2. Maddala and Miller, Microeconomics, Tata Graw Hill.
3. Salvatore, D., "Microeconomics Theory and Applications", Oxford University Press, New Delhi.
4. Rubinfield & Pyndick, Microeconomics, 5th Edition, Pearson.
5. Ahuja, H.L., Advanced Economic Theory, S.Chand.

2.9 MODEL ANSWERS

Check Your Progress: 2.1

1) (iii) 2) (i) True (ii) False (iii) False (iv) True (v) True

Check Your Progress: 2.2

1) (i) saucer (ii) higher unit costs (iii) reserve capacity

2) (i) b (ii) b

Check Your Progress: 2.4

1) (i) U-shaped (ii) greater (iii) cost considerations.

SELF LEARNING MATERIAL

ECONOMICS

COURSE : ECO - 101

MICROECONOMIC THEORY

BLOCK - III

Directorate of Distance Education
DIBRUGARH UNIVERSITY
DIBRUGARH - 786 004

ECONOMICS

COURSE : ECO - 101

MICROECONOMIC THEORY

Contributors :

Prof. J. K. Gogoi
P. P. Buragohain
J. Borgohain
Department of Economics
Dibrugarh University

Editor :

Prof. K. C. Borah
Department of Economics
Dibrugarh University

Reprint : August, 2015
Copies printed : 1000

Published on behalf of the Directorate of Distance Education, Dibrugarh University by the Director, DDE, D.U. and printed at Maliyata offset Press, Mirza, Kamrup (Assam)

MICROECONOMIC THEORY

BLOCK - III

THEORY OF PRODUCT PRICING

UNIT:I THEORY OF FIRM

Structure

1.0 OBJECTIVES

There are two broad objectives of this particular unit:

(i) To know the basic differences between the traditional and modern theories of firm.

(ii) To discuss the concept of break even point and its importance.

1.1 INTRODUCTION

It is of utmost importance to know the basic differences between the traditional theories of firm and the modern theories of firm. It is also necessary to have knowledge about the break-even point. Therefore realizing the importance we have discussed here about break-even point and its importance.

1.2 TRADITIONAL AND MODERN THEORIES: THE BASIC DIFFERENCES

The basic differences between the traditional theories of firm and the modern theories of firm are that the former's basic objective is to maximize profit; while the basic objective of the modern theories of firms are not the profit maximization but sales revenue maximization and others, for example Baumol's theory of sales revenue maximization.

1.3 BREAK EVEN POINT:

Generally a business's concern with regard to profit can be of three types, i.e., it may incur lose, it may just meet the expenses or it may make profit. Break-even point is the mid-course of the above three cases. When the output reaches such a point that it's total

revenue meets all the expenses of production, inclusive of depreciation and research, it is called the break-even point. It can also be defined as the level of output at which total revenue is equal to total cost, which implies that the net income of the firm is zero.

The cost of production of a particular product can be split in to two main parts: fixed cost and variable cost. Fixed costs remain constant over the output range under consideration while variable costs vary proportionately with the volume of output. Thus, from the cost accountant point of view, the break even sales volume is that which ensures all fixed and variable costs are covered, given a particular selling price.

Mathematically,

$$\text{Break-even Volume} = \frac{Fixed\,Cost}{Selling\,\text{Price} - Variable\,Cost\,Per\,Unit}$$

For example, if a business firm produces a product, the sale price of which is Rs. 10 per unit and the variable costs are Rs. 6 per unit. Suppose the firm's annual fixed costs are Rs. 200,000.

$$\therefore Break-Even\,Volume = \frac{Rs.200{,}000}{Rs.(10-6)} = 50{,}000 \text{ Units}$$

In this case, the net income of the firm is equal to zero at 50,000 units of output. More specifically, it can be shown as bellow:

Revenue (50000 units @ Rs. 10)	= Rs. 50,0000
Variable Cost (50,000 units @ Rs. 6)	= Rs. 30,0000
Variable Profit	= Rs. 20,0000
Fixed Cost	= Rs. 20,0000
Net Income	= Rs. 0

We can also explain the concept of break-even point with the help of Diagram 1.1

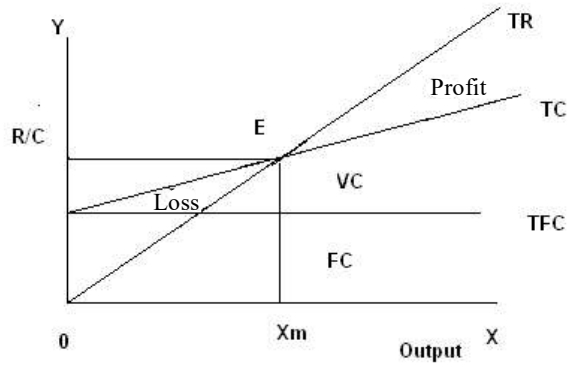


Diagram 1

In Diagram 1, on the horizontal axis we measure the units of output and on the vertical axis we measure the revenue/cost. The point E represents the break even point, where the total cost curve intersects the total revenue curve. Thus it is clear from the figure that the firm will incur loss up to X_m level of output and beyond that it will make profit.

The concept of break-even point can also be explained with the help of consumption function analysis. The point at which consumption expenditure are just equal to income as illustrated by the point at which the consumption function intersects the 45^0 line in the income-expenditure model. Below this point of interaction between the consumption function and the 45^0 line, consumption is greater than the income and therefore dissaving occurs. Above this point, consumption is less than income, i.e., saving occurs.

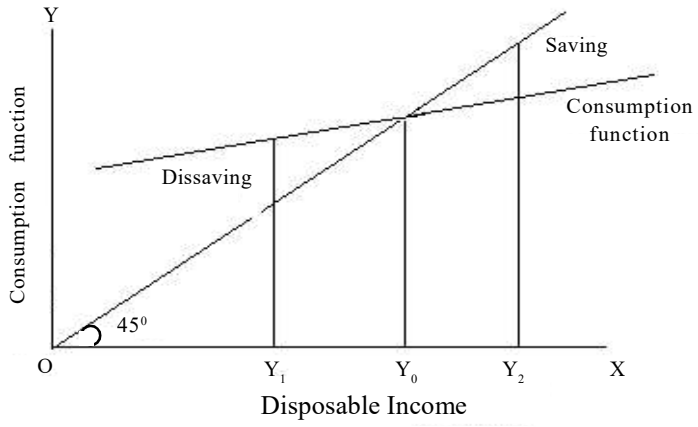


Diagram 2

Importance:

The break-even point analysis bears a significant role in case of production activities. For example, if we have to judge the efficiency between two industries than we must have to know the break-even point of production, i.e., the break-even volume of output of the each industry. If for industry I, 40% of total production is needed to achieve break-even point and 60% is needed to achieve the same for the industry II. Obviously, the industry I will be relatively non-efficient than the industry II from the point of view of the industrial production, because just to ensure the fact that the total revenue becomes equal to total cost (i.e., the break-even point is attained). The First industry has to produce 40% of his total production, which implies that normal profit is ensured when 40%

of his total production is produced. On the other hand, the industry has to produce comparatively a greater percentage (60%) of its total production in order to achieve the break-even point, which ensures normal profit. The super normal profit occurs only after the attainment of break-even point. Besides the usefulness of break-even point in analyzing the production process to make profit, it is also playing important role in making decisions in some other respects such as, the choice between buying or leasing equipment, whether to buy an item or make it within a company, or submitting tenders for winning a contract etc. Break-even analysis evolved an attempt to enable management to use the economic model of the firm by processing cost and revenue data compiled by the accountants.

CHECK YOUR PROGRESS

1. What is break-even point?

Answer:

.....
.....
.....
.....
.....
.....
.....

2. Briefly write down the importance of break-even point in economics.

Answer:

.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....

SUGGESTED READINGS

1. Modern Microeconomics, A. Koutsoyiannis, Macmillan
2. Advanced Economic Theory, H.L Ahuja, S. Chand
3. Microeconomics: Theory and Application, Salvatore, Oxford

LET US SUM UP

This unit basically deals with the concept of break-even point and the differences between traditional and modern theory of firm. When the output reaches such a point that it's total revenue meets all the expenses of production, inclusive of depreciation and research, it is called the break-even point.

KEY WORDS

Break-Even Point : When the output reaches such a point that it's total revenue meets all the expenses of production, inclusive of depreciation and research, it is called the break-even point.

UNIT: II PERFECT COMPETITION

Structure

2.0 OBJECTIVES:

The basic objectives of the unit are mentioned below:

(i) to discuss the meaning of perfect competition

(ii) to discuss the short run and long run equilibrium of firm and industry.

2.1 INTRODUCTION:

In this unit we will discuss about the meaning, assumptions behind the perfect competition. Then we will discuss about the short run and long run equilibrium of firm and industry.

2.2 MEANING OF PERFECT COMPETITION:

Perfect competition is one of the important market structures. It refers to a market structure where: there are large number of buyers and sellers of a commodity, each too small to affect the price of the commodity; the commodity is homogeneous; there is perfect mobility of the resources and the economic agents have the perfect knowledge about the market conditions.

A perfectly competitive market is based on the following assumptions:

(1) There are a large number of buyers and sellers in the industry/ market, so that no individual buyer or seller, however large, can influence the price by changing the purchase or output.

This means an individual buyer or a seller plays a very insignificant role in the market.

(2) All firms produce a homogeneous product. No way a buyer can differentiate the products available in the market.

(3) There is free entry or exit of the firm in to the industry. That is, there is no restriction on the entry or exit from the industry.

(4) There is perfect knowledge about the market condition on the part of the both buyers and sellers.

The assumption of large number of buyers and sellers and homogeneity of the products imply that the individual firm in perfect competition is a price taker. Its demand curve is infinitely elastic implying that it can sell any amount of product at the prevailing market price. In the following figure, it is seen that the demand curve of a firm in a perfectly competitive market is horizontal at price P. The level of market price is determined by the intersection of the demand and supply forces. For the firm, P= AR = MR, since the price is constant.

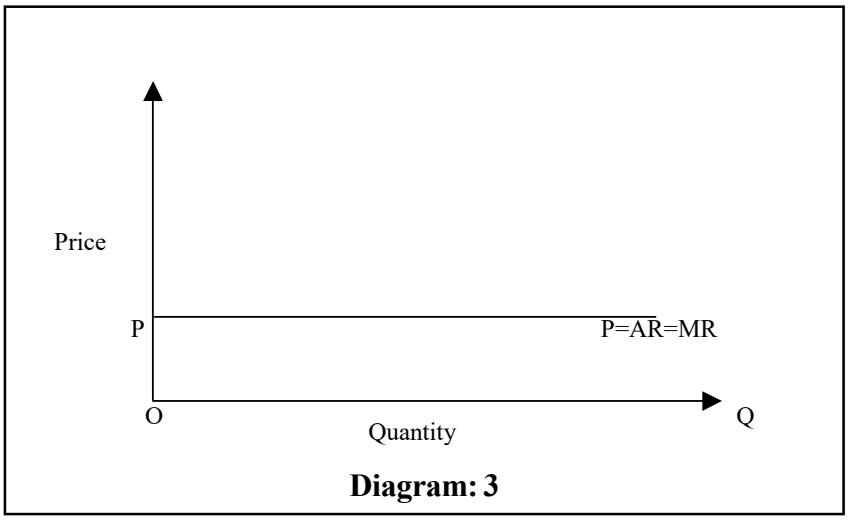


Diagram: 3

Perfect competition among the seller prevails if an individual seller does not have any influence on the market price and action of the others. Among the buyers also similar conditions must hold. A market is perfectly competitive if perfect competition prevails in both sellers and buyers sides of the market.

The perfectly competitive market model is very much important to analyze the market situation how it approximate perfect competition. It provides the point of reference or standard against which the model departs from the perfect competition. The models departing from the model of perfect competition are monopoly, monopolistic competition or oligopoly. In case of monopoly, there

is a single seller of the commodity for which there are no close substitutes. On the other hand in case of monopolistic competition, there are large numbers of sellers with differentiated products. And, oligopoly is a case of a few sellers with either a homogenous or a differentiated product.

2.3 SHORT RUN AND LONG RUN EQUILIBRIUM OF FIRM AND INDUSTRY:

A firm is in equilibrium when it has no tendency to change its output. Thus it is earning maximum profit by equating marginal cost with marginal revenue. In other words the objective of the firm is to make profit. Thus for a firm, the profit function is given by:

$$\prod = R - C$$

Where,

$\prod$ = Profit

R = Total Revenue

C = Total Cost

The total revenue and cost both are the function of the quantity. Now the objective of the firm is to maximize profit. To maximize profit it has to satisfy the first order as well as the second order conditions. The first order condition requires that :

$$\frac{\partial \prod}{\partial Q} = 0 \; ; \; \frac{\partial \prod}{\partial Q} = \frac{\partial R}{\partial Q} - \frac{\partial C}{\partial Q} = 0$$

or $= \frac{\partial \prod}{\partial Q} = \frac{\partial C}{\partial Q}$

i.e., MR=MC

But in perfect competition, the demand curve faced by a firm is horizontal at the price determined by industry supply and demand forces. Therefore, for a firm under perfect competition, P=MR=AR. Thus the first order condition for profit maximization is:

MC=MR=AR=P

The second order condition for profit maximization requires that:

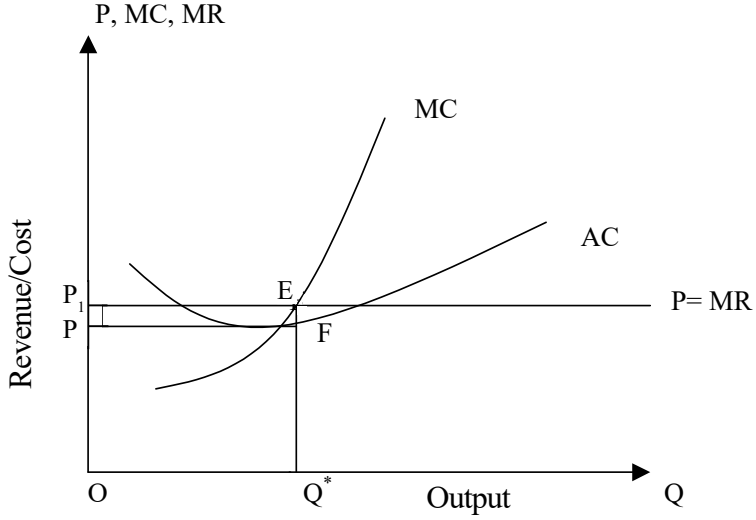
$$\frac{\partial^2 \prod}{\partial Q^2} \pi \; 0$$

$$\frac{\partial^2 R}{\partial Q^2} - \frac{\partial^2 C}{\partial Q^2} \pi\ 0$$

$$\frac{\partial^2 R}{\partial Q^2} \pi \frac{\partial^2 C}{\partial Q^2}$$

which means that the slope of MR curve should be less than the slope of MC curve. In perfect competition MR=P, which is a constant. Hence the slope of the MR curve is zero and that of the MC is positive or rising upward.

Thus for profit maximization two conditions must be fulfilled: (a) MR curve must be equal to the MC curve and (b) MC curve must be rising. These two conditions can be explained in Diagram 4.



Short run equilibrium of firm

Diagram: 4

In the above diagram, AC represents the short run average cost curve of the firm and MC the marginal cost. MC intersects AC at the latter's minimum point. Given the price, P, MR and AR curves are given by the horizontal price line. The equilibrium point E is determined by the intersection of the MC curve and price line. Drawing a perpendicular line from the point of intersection to the horizontal line, we will get the equilibrium output OQ^*. The area $PP^/EF$ represents the profit. The point of intersection of MR and MC gives the equilibrium point E. In the above diagram, it is seen that the area $PP^/EF$ represents the profit for the firm in the short run, which is called as the super normal profit.

2.3.1 SHORT RUN EQUILIBRIUM OF THE FIRM:

The firm in the short run can earn super normal profit as discussed above as well as normal profit and at the same time incur loss also. All these can be explained with the help of the following diagram:

In the below diagram, we have taken output in the horizontal axis and on the vertical axis, we have taken AR, MR, AC and MC. Since price is constant AR and MR curves are identical and are represented by a horizontal line, which is the demand curve D. The first condition of profit maximization is satisfied at the point E and E_1 where the MC and MR are equal. But at the point E_1, MC curve is decreasing while at the point E it is increasing. Thus both the conditions are fulfilled at the point E and at this point the firm will earn total revenue OPEQ, while that of the total cost is ON_1NQ there by a profit N_1PEN at OQ output and OP price.

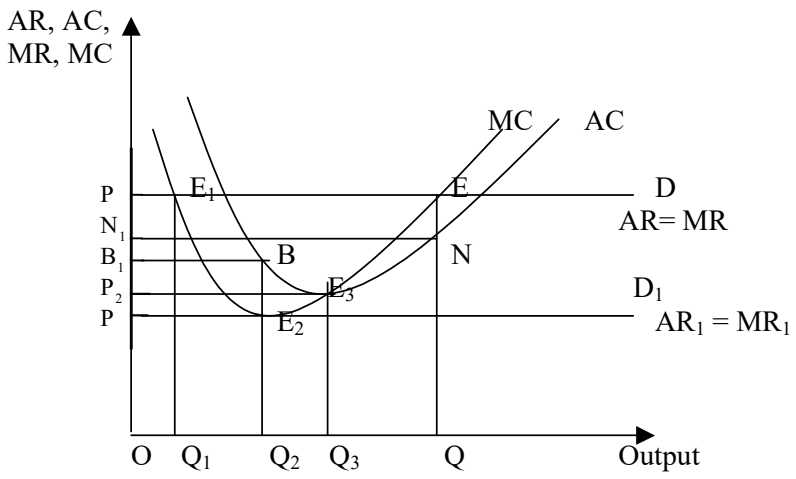


Diagram: 5

However, while discussing about the equilibrium under perfect competition, we must remember two things:

Firstly, there is no definite equilibrium point of the firm. The equilibrium level of output depends on the price, which is considered to be determined by industry demand and supply. In the above diagram it is seen that when price is OP, OQ is the equilibrium level of output and E is the equilibrium point for the firm. Now if the price falls to OP_1, OQ_2 will be equilibrium level of output at

point E_2. Thus the equilibrium of the firm depends on the price of the firm.

Secondly, the firm can earn super normal profit as well loss at equilibrium. For example, if the price is OP_1 and E_2 is the equilibrium point at E_2 both the first order and second order conditions are fulfilled. But at point E_2 the firm is earning negative profits or loss of $P_1B_1BE_2$. The firm can also earn normal profit or zero profit at equilibrium. When price is OP_2, the firm is in equilibrium at point E_3, which is the lowest point of the AC curve. In this case total revenue is equal to the total cost. In this case firm is earning normal profit. It is the profit that is included in the cost of production. If the firm earns positive profit, it is said to be earning excess profit or abnormal profit, or, supernormal profit. A firm operating under short run may continue its production even if it incurs loss because in the long run it can earn profit.

2.3.2 LONG RUN EQUILIBRIUM OF THE FIRM:

In the short run, a firm may earn super-normal profits or normal profits or incur losses. Even if the firm may incur losses, it may continue its production so long as it is able to cover its variable cost. But in the long run the firm may be capable compensating the losses. In the short run there are some fixed costs that are to be met by the firm. Thus even if a firm stops its production it will have to meet the total fixed costs. Now suppose a firm is incurring losses but it is less than or equal to the total fixed cost, the firm will continue its production. But if the firm fails to recover its total variable costs, it will have to stop production. The maximum amount of loss that the firm can bear in the short run is equal to the total fixed cost of the firm. As you know that the total cost has two components: total fixed cost total and variable cost. Now suppose the total revenue can not cover the total variable cost. This means that uncovered portion of the TVC and the entire portion of the TFC will be the loss of the firm. If the losses exceed total fixed cost, the firm will stop production. But if the total revenue covers TVC and a portion of the TFC, then the uncovered portion of the TFC is the loss of the firm in which case the firm will continue its production. Let us discuss it with the help of the following diagram:

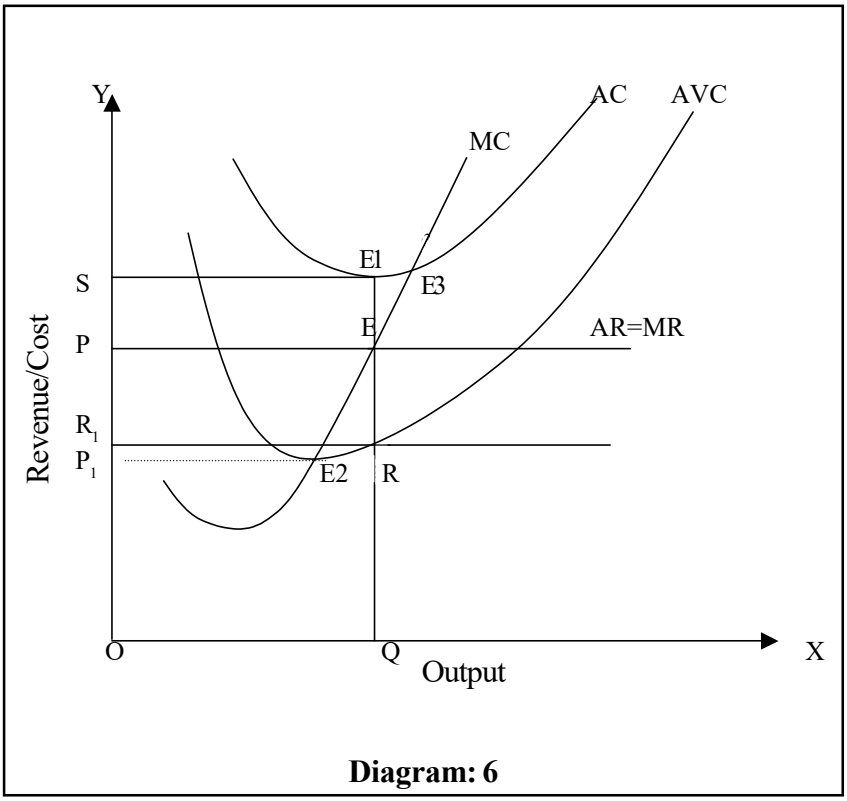


Diagram: 6

From the above diagram you have seen that at the point E, the first order as well as the second order condition of the profit maximization are fulfilled. But at price OP and point E the firm is incurring loses. Here the total revenue of the firm is equal to OPEQ; while the total variable cost and the total fixed cost is equal to RE_1S. Thus the total loss is PEE_1S which is less than the total fixed cost, and therefore the firm will continue its production process in the short run. Now if price falls below P_1, then the firm does not even meet its variable costs. That is, the loss will be equal to the fixed cost plus the uncovered portion of the variable cost. P_1 in the diagram is the lowest point on the AVC curve and where MC = AVC. Thus at the lowest point on AVC, E_2 indicates if price falls below this point it will be better for the firm to close its production. This point is called as the shut down point. Similarly, the lowest point on AC, E_1 indicates the point below which if the price falls will result in losses and the price rise above this point it means increase in profit. On the other hand E_3 represents a break-even point with no profit no loss.

On the other hand in the long run, the firms can earn only normal profit that means that they can operate at the lowest point of the average cost curves. Let us explain it with the help of the Diagram 7.

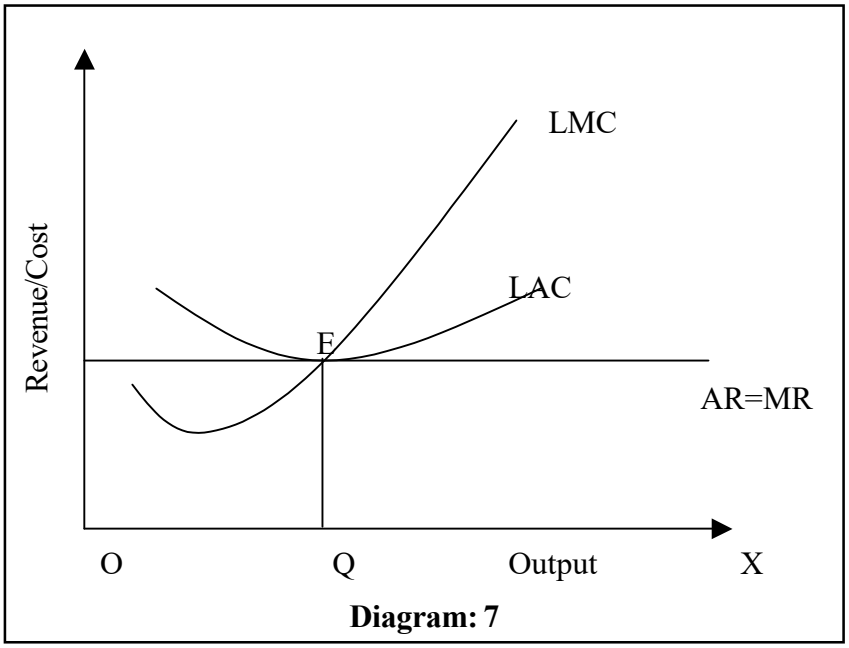


Diagram: 7

In the above diagram, we have seen that the LAC and the LMC curve intersect at the point E where AR = MR = AC = MC. In the long run, this is the only point of equilibrium and the output level OQ is the optimal level of output because in the short run there can several points of equilibrium but in the long run there will be only one point of equilibrium. In the short run the firm will be in equilibrium at any point of the MC curve lying above the AVC curve. But in the long run the firm may be in equilibrium at only one point of the MC curve, where the LAC intersects the LMC curve.

2.4 SUPPLY CURVE OF A FIRM IN PERFECT COMPETITION :

The short run supply curve of a firm in perfectly competitive market is precisely its marginal cost curve for all rates of output equal to or greater than the rate of output associated with minimum average variable cost. For market prices lower than minimum average variable cost, equilibrium quantity is zero. On the other hand, supply curve of an industry is a horizontal summation of supply curves of individual firms. It will have a positive slope to indicate that more quantity will be supplied at a higher price, and vice versa.

CHECK YOUR PROGRESS

1. What is meant by perfect competition?

.....
.....
.....
.....
.....
.....
.....

2. What is the basic difference between the perfect and pure competition?

.....
.....
.....
.....
.....
.....
.....

3. Write down the assumptions of the perfect competition on which it is based.

.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....

SUGGESTED READINGS:

1. Modern Microeconomics, A. Koutsoyiannis, Macmillan
2. Advanced Economic Theory, H.L Ahuja, S. Chand
3. Microeconomics: Theory and Application, Salvatore, Oxford.

2.5 LET US SUM UP

In this unit we have discussed about the perfect competition. Perfect competition refers to a market structure where: there are large number of buyers and sellers of a commodity, each too small to affect the price of the commodity; the commodity is homogeneous; there is perfect mobility of the resources and the economic agents have the perfect knowledge about the market conditions. The individual firm in perfect competition is a price taker. The perfectly competitive market model is very much important to analyze the market situation how it approximate perfect competition. It provides the point of reference or standard against which the model departs from the perfect competition.

KEY WORDS:

Price: The amount of money required, expressed or given in payment of something.

Output: A generic term for tangible good or an intangible service that is the end result of the production/resource transformation process.

Perfect Competition: Perfect competition refers to a market structure where: there are large number of buyers and sellers of a commodity, each too small to affect the price of the commodity; the commodity is homogeneous; there is perfect mobility of the resources and the economic agents have the perfect knowledge about the market conditions.

Price Taker: A buyer or seller that possess to little market power that it has no control over the price of the good, it must take or accept the going market price.

Price Maker: A price maker is one who possesses sufficient market control to affect the price of the good.

UNIT: III MONOPOLY

Structure

3.0 OBJECTIVES:

Once you finished the unit, you will be able to:

(a) determine the profit maximizing price and output combination of a monopolist;

(b) compare the output and price under monopoly and perfect competition;

(c) answer how and why the monopolist charges different prices to different customers; and

(d) price and output decisions in multi-plant and bilateral monopolies.

3.1 INTRODUCTION:

Monopoly is a form of market structure characterized by a single seller who acts as a price setter. The monopolists are called the price setter as they determine their own price and supply the entire quantity demanded by the market. In monopoly there is only one seller producing and selling a product, which does not have any close substitutes. The monopolist firm is the industry and the demand for the monopolist coincides with the industry demand.

The monopolist can change both the price charged and quantity to be produced as they wish. The price elasticity of demand for the monopolist is finite. The monopolist has the firm control over the market and entry is blocked.

3.2 DEFINITION OF MONOPOLY

'Mono' means one and 'poly' means seller. Thus monopoly refers to a market situation in which there is only one seller of a particular product. This means that the firm itself is the industry and whatever is the product produced by it has no substitutes thereby it has the monopoly over the product. A monopoly firm possesses the following important characteristics:

(i) In monopoly market there is only one single or dominant producer.

(ii) In monopoly, a producer produces a particular commodity, which is not produced by the other firm, that is, the product does not have close substitutes.

(iii) The monopoly firm possesses firm control over the market supply.

(iv) The monopolist is a price setter not a price taker.

(v) In monopoly there is no entry of new firm in to the industry because of legal or natural barrier.

(vi) There is absence of competitive advertisements

(vii) Since there is only one firm in the industry, under monopoly the equilibrium of the firm is same as the equilibrium of the industry.

3.3 FACTORS BEHIND THE EXISTENCE OF THE MONOPOLY

There are many factors behind the existence of monopoly. The main causes that lead to monopoly are mentioned below:

(i) Ownership of strategic raw materials or exclusive knowledge of production technique.

(ii) Absolute cost advantage.

(iii) Locational advantage.

(iv) Patent right on product or on the process of production.

(v) Government licensing or imposition of foreign trade barrier to restrict foreign competitors.

(vi) Practice of limit pricing.

3.4 EQUILIBRIUM OF THE MONOPOLIST

A monopoly is a form of market where the monopolist is always guided by the profit motive. There fore the objective of the monopolist is to maximize profit.

3.4.1 SHORT-RUN EQUILIBRIUM:

A monopolist like other producers is always guided by the chief consideration of maximization of net gain or the minimization of the net losses. In the short run, the monopolist cannot adjust its plant size but it can maximize its short run profit by equating marginal revenue curve and marginal cost curve. The second order condition is that MC must cut MR from below. This is explained with the help of the Diagram 8.

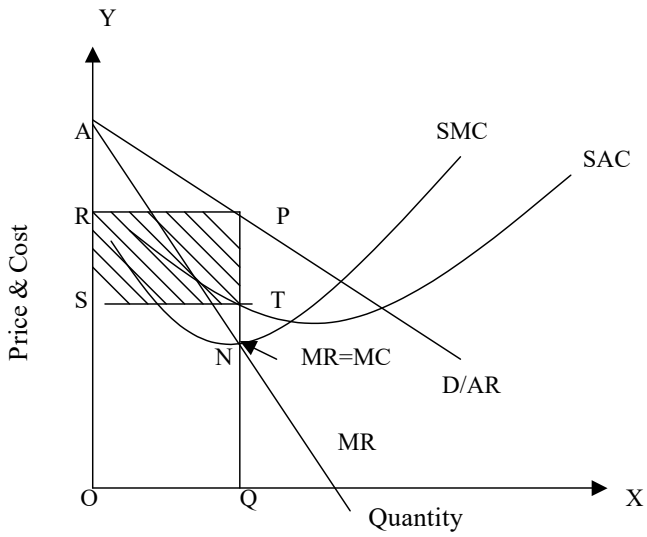


Diagram: 8

In the Diagram 8 you have seen that the price & costs are measured on the vertical axis and quantity is measured on the horizontal axis. You have also noticed that demand curve of the monopolist is downward slopping. The primary reason for this is that since the firm is a price maker not a price taker, if it reduces the supply in the market, the price will rise; and vice versa. So its demand curve or average revenue curve is downward slopping from left to right.

The equilibrium position in the above diagram is established at the output OQ where MC is equal to the MR. For OQ equilibrium output he charges QP price and thereby earns a super normal profit

of PRST. This excess profit is not disappeared in the long run due to entry restriction.

The profit will attract additional firms in to the perfectly competitive industry until all firms just break even in the long run. On the other hand the monopolist can continue to earn profit even in the long run because of blocked entry.

3.4.2 LONG RUN EQUILIBRIUM

In the short run, the monopolist aims at maximization of the profit subject to the given plant size already built and operating. On the other hand, in the long run the first problem faced by the monopolist is whether he should be in the business or not. If the demand and cost conditions are such that the long run average cost (LAC) cannot be covered, the monopolist will abandon his business. Thus the second decision he has to take is that what is the most profitable size of the plant and operating it at the most profitable rate.

A monopolist in the long run will choose such a plant that yields him the abnormal profit. Thus a monopolist firm in the long run will earn abnormal profit in the long run, which is explained in the diagram 9.

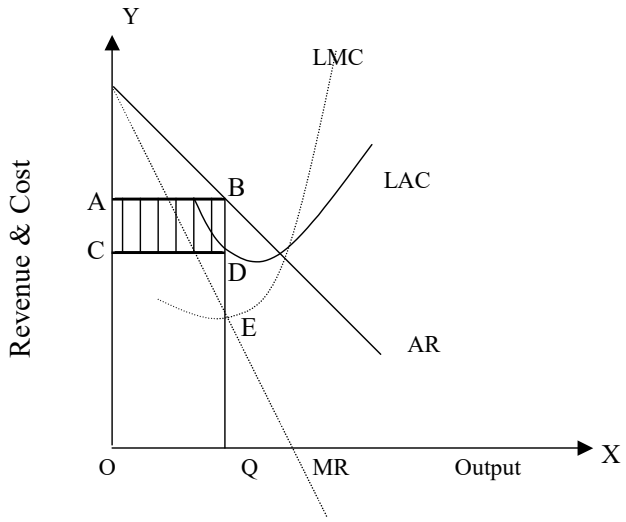


Diagram: 9

In the diagram 9, the monopolist's profit is determined at the point E, where MC=MR, corresponding the output, OQ that is produced at QD cost per unit and is sold at QB per unit. Thus the firm makes profit BD per unit and makes a total profit as shown by the shaded area ABCD.

3.5 PRICE DISCRIMINATION:

When the monopolist charges different prices from different buyers for the same good, he is known as price discriminating monopolist. Price discrimination is not possible under perfect competition because everyone knows the price at which the good is being bought and sold. A monopolist, however, can charge different prices for the same good. There are two conditions that must be fulfilled to be price discrimination. Firstly, the market must be divided in to sub markets with different price elasticities. Secondly, there must be effective separation of the markets so that reselling is not possible.

In the simplest case price discrimination occurs when an identical product is going to two buyers in such a way that Price paid by Buyer A is different from that of the buyer B.

3.5.1 EQUILIBRIUM OF A DISCRIMINATING MONOPOLIST:

A discriminating monopolist, like an ordinary monopolist, tries to maximize his profit. He supplies different amount of goods to different markets to attain his goal of profit maximization. In fact his goal of profit maximization is profitable if the elasticity of demand are different in different markets. If the demand for product of the monopolist in the market A is greater than the market B, he would like to supply more in the market A by reducing supply in market B. If a discriminating monopolist is to be in equilibrium, two conditions are to be fulfilled:

(i) Marginal revenue in both or all markets must be the same.

(ii) The marginal revenue derived from each of these markets must also equal the marginal cost of the monopolist's total output.

The equilibrium under price discrimination can be explained with the help of the following diagram:

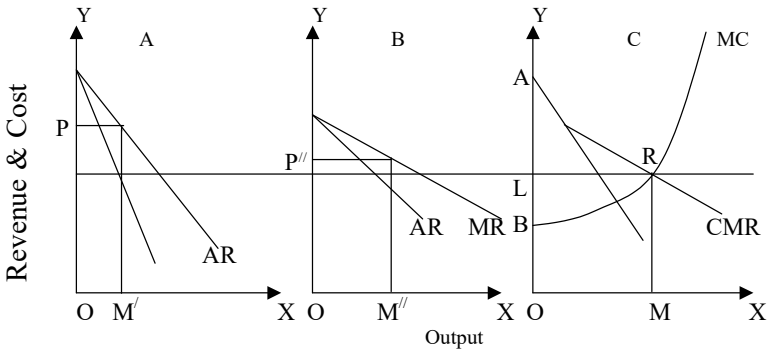


Diagram: 10

In the Diagram 10, panel A and panel B show the average and marginal revenue curves of the firm for two separate markets (sub-market A and sub-market B). These markets have different elasticities of demand at each price. In panel C of the diagram, the profit maximizing output (OM) is shown at the intersection of the marginal cost curve (MC) for the monopolist's whole output, with the curve showing combined marginal revenue (CMR) obtained from the two markets. The curve CMR is obtained by adding the curves MR_1 and MR_2 together sideways.

In this equilibrium situation, the output is OM, and marginal revenue is OL or MR. The output OM has, therefore, to be distributed between the two separate markets in such a way that marginal revenue in each is OL. It means that $OM^{/}$ is to be sold in sub-market A at price OP (marginal revenue is here OL).

Similarly, $OM^{//}$ must be sold in sub-market B at a price of $OP^{/}$ (marginal revenue here is also OL). The monopolist's profit is shown by the area ARB in the panel C of the diagram and here it is at a maximum.

Output under Price Discrimination:

The total output of a monopolist with two or more prices can be either larger or smaller than his total output if he would sell at one price.

In practice, demand and cost relations are such that without discrimination a particular commodity or service will not be produced at all. Take the case of Indian Sugar industry. If free sale of sugar is prohibited, production of sugar will be unprofitable. Some commodities and services might not be produced at all if sellers were not be able or were not allowed to practice price discrimination. The standard and simple example is the physician in a small village.

Preconditions of Price Discrimination:

Although price discrimination is possible under monopoly, yet it is not always possible. Pigou has mentioned two conditions to be required for successful operation of price discrimination by a monopolist.

(i) Price discrimination is possible when there is no possibility of resale of product particularly the service.

(ii) Price discrimination is also possible when there is effective separation of markets from one to another.

Besides these two conditions, price discrimination is also possible under the following conditions:

(a) A monopolist becomes successful in price discrimination on account of consumer's peculiarities; such as consumers' ignorance about the prices, consumers' irrational feeling about the quality of the product, consumers' indifference towards small price differences, etc.

(b) Again, a monopolist becomes successful in price discrimination when the demand for his product has different elasticities in two sub-markets or different markets.

(c) When, there is no state intervention or legal bar.

(d) Finally, price discrimination is possible when buyers and sellers are separated from one another by a great extent.

3.5.2 EFFECTS OF PRICE DISCRIMINATION:

The effects of discriminating monopolist can be explained with the help of the diagram drawn below as follows: so long as the price-discriminating monopolist manages to reap part of the consumer's surplus his total revenue and his total profits will be higher the monopolist can reap consumers surplus OX, which is defined by the intersection of his MC with his (original) aggregate MR curve.

However, the quantity that the discriminating monopolist supplies will be higher than OX if he can charge more than two prices in various sectors of the market, and his total revenue will be higher still. The increase in output is due to the gradual shift of the MR curve upwards and the consequent change of the point of intersection with the given MC curve. The shift of the MR curve, given that $DD^{/}$ does not change, is due to the fact that the MR (at all levels of output) is higher when price discrimination is being adopted, because the lower the price at which the new marginal unit is sold is not the same as for all previously sold units, which have been sold at higher prices via individual negotiations with the buyers. In the diagram if price P were uniform, the monopolist would sell OX amount of output. If price discrimination is applied, only the additional units are sold at gradually lower prices, and hence MR will shift to $MR^{/}$. The new equilibrium will be at $E^{/}$ where output $OX^{/} > OX$. In the limiting case, MR will coincide with the demand curve $DD^{/}$ and the buyer will wish to pay the highest price at $E^{//}$ at output level $OX^{//}$ and the following condition will hold:

$$\mathbf{MC = MR = AR = P}$$

and the seller will have achieved the maximum increase in his revenue, reaping all consumer's surplus.

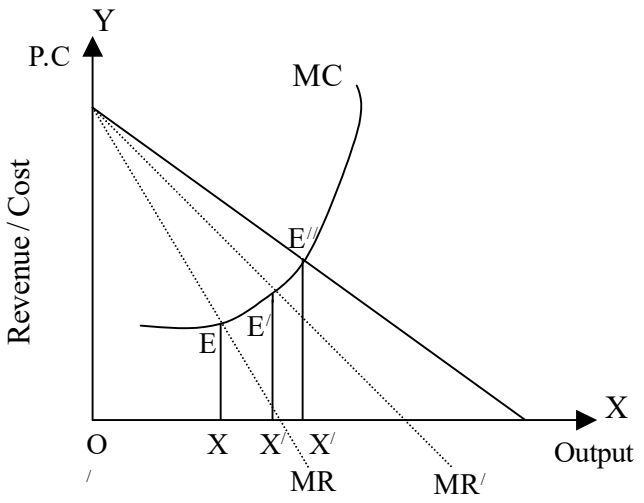


Diagram: 11

3.6 MULTI-PLANT MONOPOLIST :

A monopolist may operate more than one plant to produce a homogeneous product. Such a monopolists that operates more than one plant simultaneously in order to produce a homogeneous product is called as a multi plant monopolist. In the short run a monopolist can operate any number of plants of the same size or of different sizes. But in the long run, it operates only those plants, which bring him the largest profit. Thus to produce profit maximizing output the multi plant monopolist operates in such a way that the MC in each plant is equal to the MR from selling the combined outputs.

3.7 BILATERAL MONOPOLY:

A bilateral monopoly is said to exist when there is only one seller and one buyer of a product. In this form of monopoly, since there is only one buyer and one seller, the price and quantity can be determined by negotiation. We can explain the concepts of bilateral monopoly by considering two situations: (i) the single-seller as all-powerful and (ii) the single buyer as all-powerful. Let us take the help of the following diagram to explain it.

In the below diagram, price is measured on the vertical axis and quantity on the horizontal axis. D is the demand curve, MR is the marginal cost and MC is the marginal cost curve of the single seller. If the monopolist is the most powerful, then she will make the buyer to behave as if there were many buyers and she will equate her MC to MR, produce output Q_S, and charge price P_S.

However, if the single buyer is all-powerful, he can make the monopolist behave like a perfect competitor. Thus, MC_S will be monopolist's supply curve.

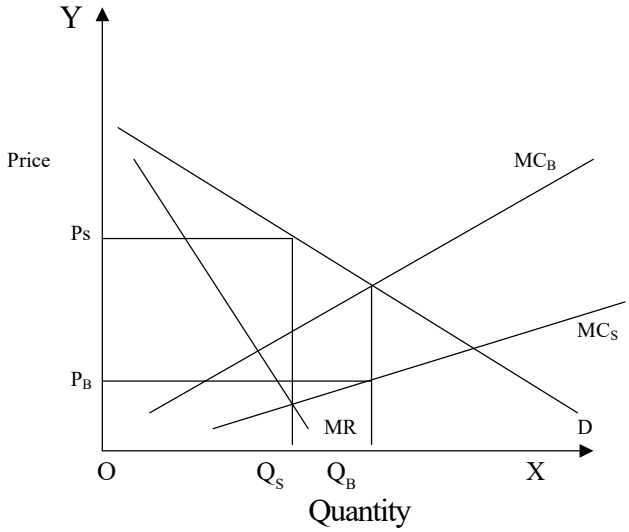


Diagram: 12

CHECK YOUR PROGRESS

1. Define monopoly.

.....
.....
.....
.....
.....

2. Write down any four factors behind the existence of monopoly.

.....
.....
.....
.....
.....
.....

3. What are the conditions required for the existence of short run and the long run equilibrium of a monopolist?

.....
.....
.....
.....
.....

.....
.....
.....

4. What do you mean by discriminating monopolist?

.....
.....
.....
.....
.....
.....

5. Mention the preconditions of price discrimination.

.....
.....
.....
.....
.....
.....
.....

SUGGESTED READINGS

1. Modern Microeconomics, A. Koutsoyiannis, Macmillan
2. Advanced Economic Theory, H.L Ahuja, S. Chand
3. Microeconomics: Theory and Application, Salvatore, Oxford.

LET US SUM UP

In this particular unit we have discussed about monopoly and the various factors responsible for the existence of the monopoly. Monopoly is a form of market structure characterized by single seller of a unique product with no close substitutes.

Thereafter we have discussed about the equilibrium of the monopolist under both short run and the long run conditions. Consequently we have discussed about price discrimination, multi plant monopolist and the bilateral monopoly.

KEY WORDS

Monopoly: Monopoly is a form of market structure characterized by single seller of a unique product with no close substitutes.

Price Discrimination: It is situation when the monopolist charges different prices to different buyers in different markets for the same good.

Multi-Plant Monopoly: A monopolist may operate more than one plant to produce a homogeneous product. Such a monopolists that operates more than one plant simultaneously in order to produce a homogeneous product is called as a multi plant monopolist.

Bilateral Monopoly: It is a market containing a single buyer and a single seller. The buyer and seller and forced to negotiation a price. Where the price ends ups depends on the relative negotiating power of each side.

Price Discrimination: It is situation when the monopolist charges different prices to different buyers in different markets for the same good.

Multi-Plant Monopoly: A monopolist may operate more than one plant to produce a homogeneous product. Such a monopolists that operates more than one plant simultaneously in order to produce a homogeneous product is called as a multi plant monopolist.

Bilateral Monopoly: It is a market containing a single buyer and a single seller. The buyer and seller and forced to negotiation a price. Where the price ends ups depends on the relative negotiating power of each side.

There are several sources of differentiation such as brand name, chemical composition of the product, advertising, location, and packaging etc.

2. Nonprice Competition: Since the products under monopolistic competition are only slightly differentiated and hence the producers are trying to play up the differences in their products in order to increase the demand. They are trying to do this by variety of ways such as advertising.

3. Large Number of Firms and Freedom of Entry and Exit: There are large number of firms, each satisfying a small, but not microscopic, share of the market demand for a similar, but not identical, products. There is also relative freedom of entry and exit of firms in monopolistically competitive markets.

4. Independent Behaviour: The economic impact of one firm's decisions is spread sufficiently evenly across the entire group so that the effect on any single competitor goes unnoticed. This implies that conscious rivalry is missing or that competition is impersonal. Each firm behaves independently.

4.2 DEMAND CURVE OF A MONOPOLISTICALLY COMPETITIVE FIRM:

The demand curve for a monopolistically competitive firm is downward slopping due to it product differentiation. But it is quite elastic since there are close substitutes for the product of a firm. The marginal revenue of the firm will be less than the corresponding price as the demand curve is downward slopping. In case of the monopolistic firms, the marginal revenue curve will be half that of the demand curve which is shown with the help of the following diagram:

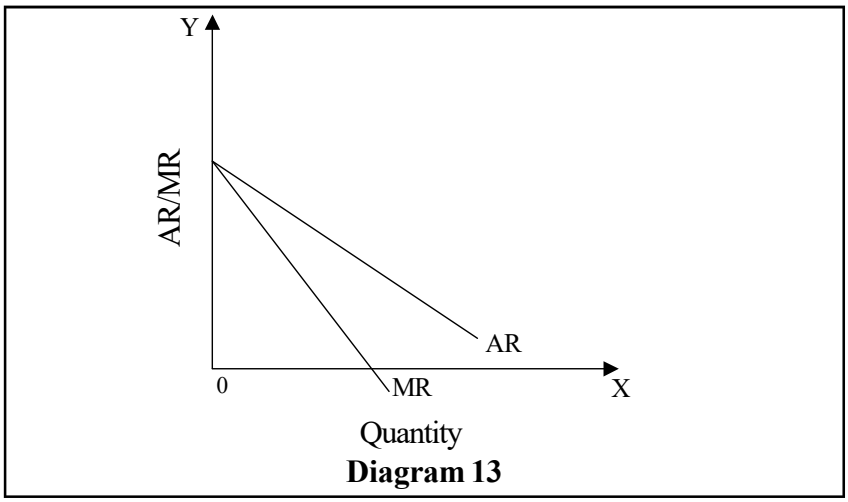


Diagram 13

4.3 EQUILIBRIUM UNDER MONOPOLISTIC COMPETITION:

Under monopolistic competition different firms produce different varieties of products and different prices for them will set at their respective demand and cost conditions. Each firm will set its own price and output of its own product.

Now the question is at what price and output level the monopolistically competitive firm will be in equilibrium. The basic objective of the monopolistic firm is also to maximize profit and therefore the producer will go on producing till the extra receipts to be made from additional production exceed the extra costs to be incurred in the process of production and the producer will stop at the point where the extra receipt and the extra cost is equal. In other words, profits will be maximized when the marginal cost is equal to the marginal revenue. Thus, for a producer, if marginal revenue is greater than the marginal cost then he will go on expanding his production level as he will earn more and more profit. On the other hand if marginal revenue is less than the marginal cost, it means he is incurring losses. Therefore, the producer can go on producing his product up to the point where his marginal revenue is equal to the marginal cost. After this point he will have to reduce production. Thus, in the short run, the firm will be in equilibrium when it is maximizing its profits, i.e., when

Marginal Revenue = Marginal Cost.

Short-run Equilibrium: In the short run a monopolistically competitive firm can earn profit or it may incur loss. Let us discuss the short run equilibrium of monopolistically competitive firm with the help of diagram.

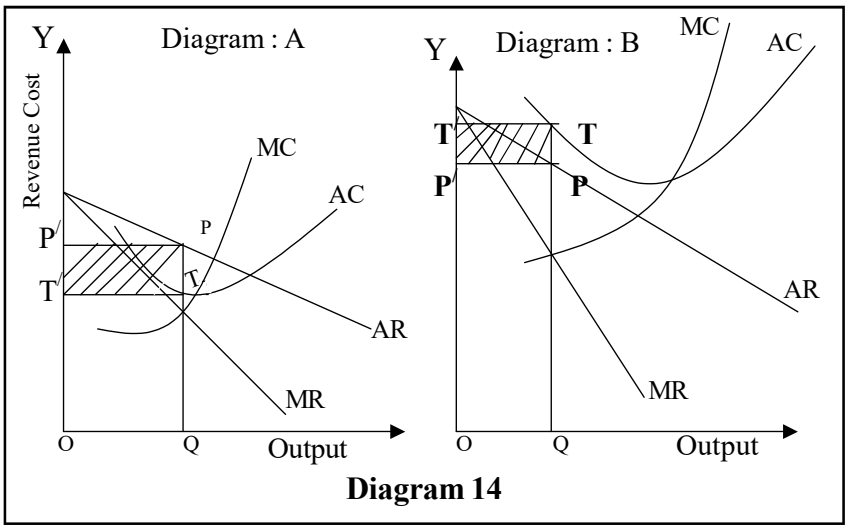


Diagram 14

In the above diagram (A) and (B) on the horizontal axis we measure output and on the vertical axis, we measure the cost/ revenue. AR is the average revenue curve and MR is the marginal revenue curve, AC is the average cost curve and MC is the marginal cost curve.

In the diagram (A) the firm is earning super normal profits. The super normal profit per unit of output is the difference between average revenue and average cost at the equilibrium point. In this case the equilibrium average revenue is QP and the average cost is QT and hence PT is the super normal profit per unit of output and $PTT^{/}P^{/}$ is the total super normal profit. On the other hand at the same time a monopolistically competitive firm may incur loss also in the short run. The firm will incur loss when the demand and the cost conditions are not suitable for the firm. From diagram (B) it is clear that the firm is incurring losses to the amount of $PTT^{/}P^{/}$. Here cost is higher than the revenue. Thus a monopolistically competitive firm may either earn super normal profit or incur loss in the short run.

Long-run Equilibrium: We have noticed from the above explanation that a monopolistically competitive firm can either earn super normal profit or incur loss. But in the long run a monopolistically competitive firm can not earn super normal profit. In the long run such profit disappears. This is due to the fact that we have assumed that entry is free and hence if entry is unrestricted new firm will enter in to the market if the existing firms are earning super normal profits. As new firms enter and start production, the demand curve or the average revenue curve faced by the firm will fall and therefore, the super normal profits will be competed away. And the firms will be earning only super normal profits. On the other hand if the firms are incurring losses in the short run, then some firm will leave the industry so that the remaining firms will earn only normal profits.

Another point to be noted here is that average revenue curve in the long run is more elastic as large number of substitutes are available in the long run. There fore in the long run equilibrium is established when firms are earning only normal profits. Now profits are normal only when average revenue is equal to the average cost. Therefore, equilibrium in the long run under monopolistic competition holds when ***Average Revenue is equal to the Average Cost.*** This is explained with the help of Diagram 15:

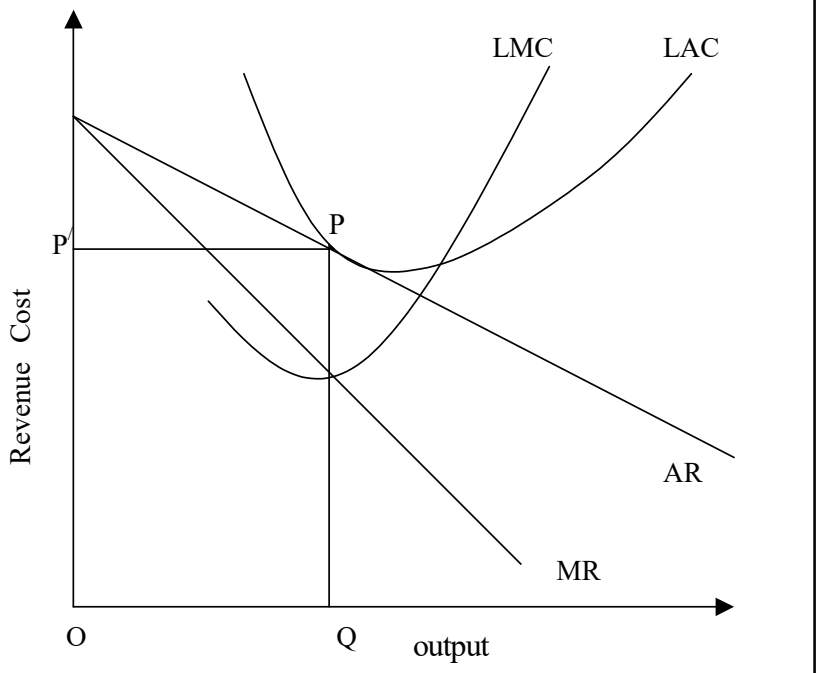


Diagram 15

From the above diagram it is clear that the equilibrium is reached at point P where the average revenue curve is tangent to the average cost curve. The equilibrium output in the long run is OQ and the corresponding price is QP (=OP/). At this point the average cost is also QP and so is average revenue. Thus there are no super normal profits; there are only normal profits. So, in the long run a monopolistically competitive firm is in equilibrium only when:

(i) Marginal Revenue = Marginal Cost.

(ii) Average Revenue = Average Cost.

4.4 MONOPOLISTIC COMPETITION AND ADVERTISING:

An essential characteristic of monopolistic competition is product differentiation. In fact, it is the hallmark of monopolistically competitive markets. In monopolistically competitive form of market, advertising is very much needed to inform the consumer about the differences in its products. Advertisings are both informative as well as the persuasive. Monopolistically competitive markets are characterized by brand names and by continual

development and improvement of the product. With advertisement, the firms want to establish their brand names and thereby maximization of the sales.

Advertising increases cost of production of a firm under monopolistic competition. But at the same time it also shifts demand curve. The net effect could be a rise or a fall in price or a rise or fall in the quantity also. In the following diagram, we show a case of decrease in price and an increase in quantity.

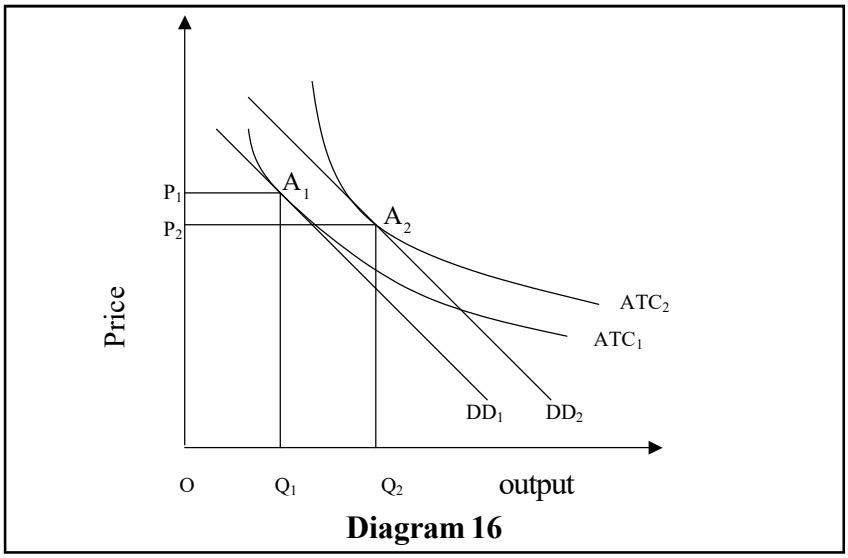


Diagram 16

In the above diagram, DD_1 is the demand curve; P_1 is the price; Q_1 the quantity and the ATC_1 represents average total cost of the firm without advertising. On the other hand DD_2 is the demand curve; Q_2 the quantity; P_2 the price and the ATC_2 the average total cost curve of the firm with advertisement. On the other hand, in the diagram below we show a case of an increase in price and an increase in quantity. Here, without advertising price is P_1 and quantity produced is Q_1; while after advertisement both price and quantity produced increases to P_2 and Q_2 respectively.

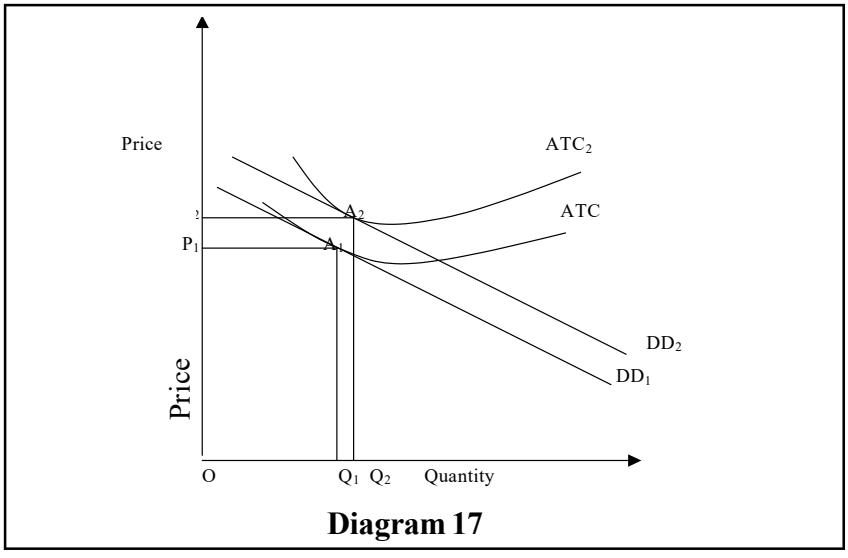


Diagram 17

CHECK YOUR PROGRESS

1. Define Monopolistic competition.

.....

.....

.....

.....

.....

2. Mention the main features of monopolistic competition.

.....

.....

.....

.....

SUGGESTED READINGS :

1. Modern Microeconomics, A. Koutsoyiannis, Macmillan
2. Advanced Economic Theory, H.L Ahuja, S. Chand
3. Microeconomics: Theory and Application, Salvatore, Oxford.

4.5 LET US SUM UP

This unit is devoted to discuss about the monopolistic competition. Professor Chamberlin introduces the concept of monopolistic competition which refers to the form of competition where there are many sellers of a homogeneous or differentiated product, and entry into or exit from the industry is rather easy in the long run. One distinguishing feature of the monopolistic competition is product differentiation. In this particular unit we discuss further about impact of advertising on monopolistic competition. Advertising increases cost of production of a firm under monopolistic competition. But at the same time it also shifts demand curve. The net effect could be a rise or a fall in price or a rise or fall in the quantity also.

KEY WORDS

Monopolistic Competition: Monopolistic competition, which refers to the form of competition where there are many sellers of a homogeneous or differentiated product, and entry into or exit from the industry, is rather easy in the long run.

UNIT: V OLIGOPOLY

Structure

5.0 OBJECTIVES:

After going through the unit you will be able to:

(i) Define oligopoly;

(ii) Explain different models of oligopoly;

(iii) Understand game theory;

(iv) Explain limit pricing and the Baumol's theory of sales revenue maximization.

5.1 INTRODUCTION :

Oligopoly is a form of market structure wherein there are few sellers with either a homogeneous product or a differentiated product. It is one of the most important forms of market structure. The simplest form of oligopoly is duopoly, where there are only

two sellers. If the products are homogeneous, it is called a pure oligopoly and if the products are heterogeneous then we have a differentiated oligopoly. Pure type of oligopoly is found in steel, copper, cement, petrol and a few other industries. On the other hand differentiated form of oligopoly is found in automobiles, tyres, cigarettes, electrical appliances, baby food and others. One very important distinguishing feature of oligopoly is the interdependence or rivalry among the firms in the industry. There are many oligopoly models found in literature. Some important models among these are mentioned below:

(i) The Cournot Model,

(ii) The Edgeworth Model,

(iii) The Chamberlin Model,

(iv) The Stackleberg Model,

(v) The Kinky Demand Model,

(vi) The Centralized Cartel Model,

(vii) The Market Sharing Cartel Model and

(viii) The Price Leadership Model.

5.2 THE COURNOT MODEL:

The French mathematical economist Augustin Cournot gave a solution to the problem of duopoly pricing in 1938 on the assumption that each seller expected his rival never change his output. He had given his solution by taking the example of two mineral springs produced by two persons A and B. Cournot's model is based on the following assumptions:

(i) There are two profit maximizing duopolists, say A and B

(ii) Each duopolist produces an identical product

(iii) Each duopolist sell at identical prices

(iv) Each duopolist fully knows the linear market demand curve

(v) Both duopolists act independently without collusion

(vi) Each duopolist acts under the assumption that its rival's output will remain exactly where it is now.

Under the above assumptions, Cournot explained his model assuming that of two adjoining mineral springs produced by two producers A and B at zero marginal cost. The model can be explained with the help of the following diagram:

In the below diagram, DXn is the demand curve. Since MC is equal to zero. The competitive output is OXn. Let MR_1 be the marginal revenue curve corresponding to the demand curve DXn. Equating MR_1 to MC, we get the monopoly output as OX_1, monopoly price OP_1 and the monopoly profit OP_1AX_1, which would be the solution if two duopolists colluded. If A is the seller, he will behave like a monopolist and sell

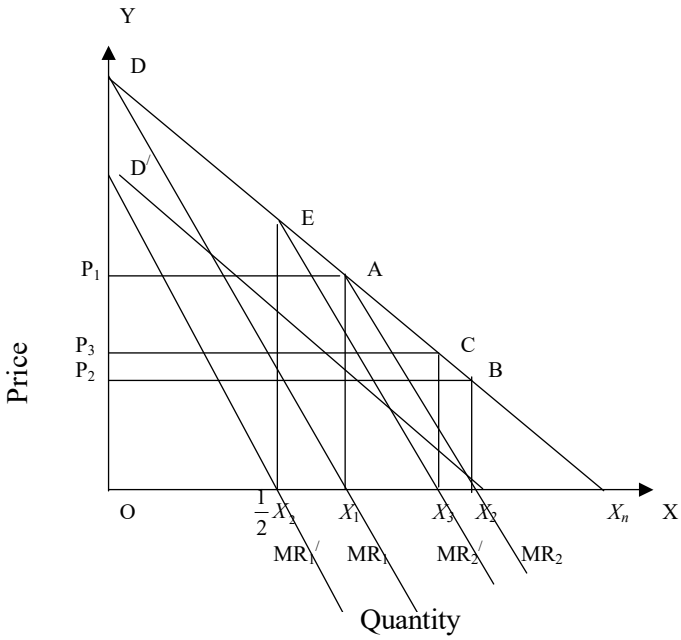


Diagram 1: Cournot's Solution to Duopoly Problem

Output OX_1 = ½ OXn at price OP_1 and makes profit OP_1AX_1. Now if B enters in to the market he will produce half of the A's demand curve. The second duopolist's MR will be equal to the MC at output X_1X_2 and that is the initial profit maximizing output. Total output has gone up from OX_1 to OX_2 and hence price will fall.

Now A realizes that B has entered in to the market and therefore he revaluate the situation and by assuming that B will sell at $X_1X_2 = X_2X_n$, A formed his new demand curve as D/X_2, which is obtained by horizontally subtracting B's output from the market demand curve. The new marginal revenue curve for A is $MR_1/$ which intersects the horizontal axis at ½ (OX_2). The combined output now declines, and product price increases.

Now B revaluates the situation. B's new expectation is that A will always continue to sell ½ (OX_2). B's new demand curve is constructed as before and becomes EXn and he will maximizes output by increasing output to ½ (OX_n-1/2 OX_2). This process will go on until A and B together produce output OX_3= 2/3 OX_n and each seller will produce output 1/3 OX_n.

5.3 EDGEWORTH MODEL:

Edgeworth had developed another model taking the same set of initial conditions as the Cournot model (zero marginal cost and the same product - mineral spring). However, Edgeworth had made two different assumptions, viz., (i) each firm has a maximum but equal rate of output; and second, each firm assumes that the other will keep its price unchanged. Let us explain the model with the help of the following diagram:

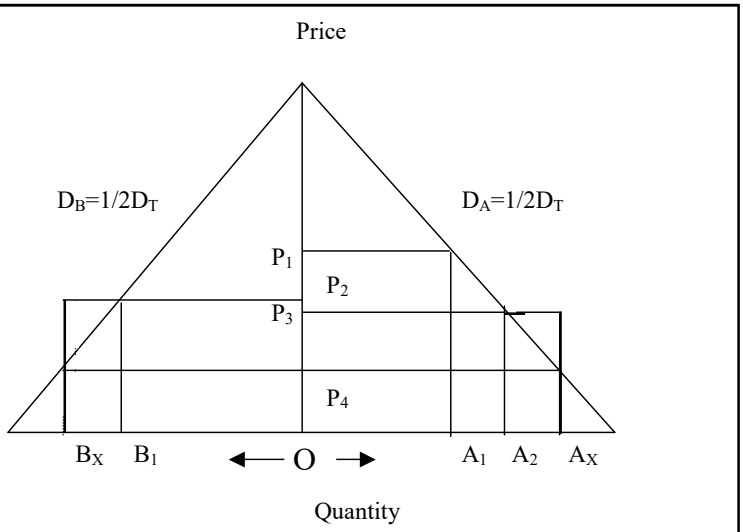


Diagram: 2

In the above diagram, the maximum levels of output for two firms are A_X and B_X. The total market is shared equally between firms that D_A and D_B each represent half of the total market demand D_T. Let us suppose the firm A first starts production and produces profit maximizing output OA_1 and sets price at OP_1. Now Producer B observes that by setting price at OP_2 it can sell B_1 units in its market and as much as it likes in Firm A's market. In diagram this would be B_X minus B_1 units that would be sold in Firm A's market and consequently the sale of the Firm A will reduce. Now suppose the firm B keeps price at P_2 firm A sets its price slightly lower at P_3 and sell the product at A_2 units in its own market covering the amount of market share lost to the firm B.

This price-cutting continues until the price P_4 is reached, at which stage each firm sells its maximum output and each makes an equal economic profit of P_4A_X or P_4B_X. However, sooner or later one firm will realize that raising price to P_1 could increase the profit. Now the other firm will again raise the price slightly lower to P_1 and the price war starts again and the process continues. Thus the

Edgeworth solution does not lead to a stable equilibrium as in the Cournot model.

5.4 CHAMBERLIN MODEL

The Cournot model as well as the Edgeworth model ignores one important factor that the firms in the market are interdependent and to recognize this drawback Chamberlin developed a model that recognized the interdependence between the duopolists in the market place. The model is akin to the Cournot model but close in reality in assuming that both firms are aware that they can do better by sharing the monopoly profit than by any other action. The Chamberlin's model can be explained with the help of the following diagram:

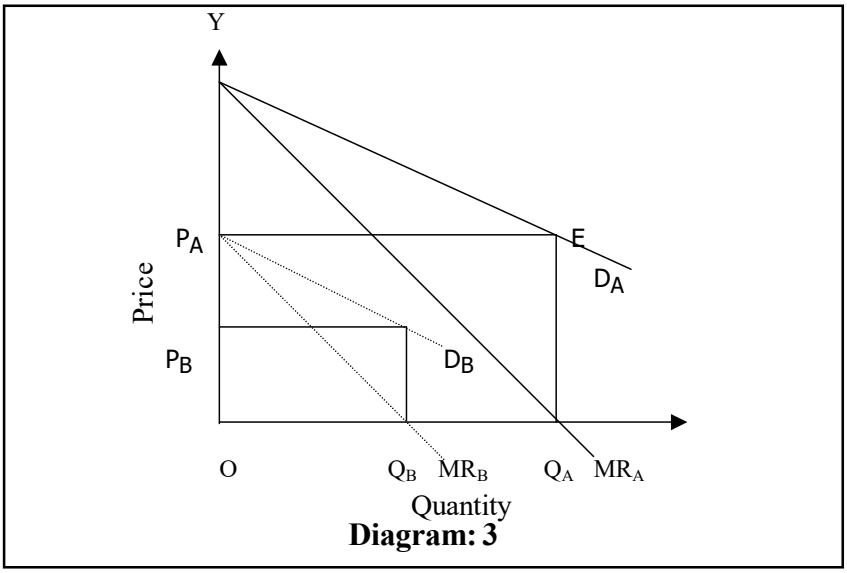


Diagram: 3

Let us suppose firm A enters the market first. With the demand curve D_A, marginal revenue curve MR_A and MC = 0, the profit maximizing level of output is Q_A units and the price is PA. Now firm B enters in to the market with the assumption that its relevant demand curve is D_A minus Q_A units. This is the demand curve D_B. Setting the marginal revenue MR_B= MC= 0, we see that the firm B produces Q_B units at price P_B.

But unlike the Cournot model the firm A recognizes that firm B's decisions are based on what firm A does and, recognizing this market interdependence, cuts output to ½ Q_A (=Q_B) units. Firm B also recognizes the interdependences and keeps output at that level but raises the price to P_A. So firm A reduces output but keeps the price unchanged, while in case firm B the strategy is reversed: it maintains output but increases price. The result is that the two firms share the monopoly profit OP_AEQ_A.

5.5 COLLUSIVE OLIGOPOLY:

The main feature of an oligopolistic market is uncertainty that result from oligopolistic interdependence and to remove this uncertainty the oligopolists may enter in to collusive agreements. There are two main types of collusions:

(i) The cartel

(ii) The price leadership

Cartel: Cartel is a group of firms acting together to control output and price. The main objective of cartel is maximization of joint profits. One of the most famous cartels is Organization of Petroleum Exporting Countries (OPEC). There are two main types of cartels: (i) Joint Profit Maximizing Cartels and (ii) Market Sharing Cartels. However, it is not easy to form a cartel.

The success of a cartel depends on several factors. Some important factors are mentioned below:

(i) The price elasticity of demand

(ii) The stability of demand.

(iii) The ability to control a substantial share of actual and potential output.

(iv) The political climate in the case of international trade.

Joint Profit Maximizing Cartels: The basic objective of the joint profit maximizing cartels is to maximize the joint profit. In this form of cartel the firms produce homogeneous product and there is a central agency that decide the output quotas for the member firms, the prices to be charged and the distribution of industry profit. Since the central board manipulates prices, outputs, sales and distribution of profits, it acts like a single monopoly whose main objective is to maximize the joint profit of the oligopolistic industry. The joint profit maximization cartel is based on the following assumptions:

(i) Only two firms A and B are assumed in the oligopolistic industry that forms the cartel.

(ii) Each firm produces and sells a homogeneous product that is a perfect substitute for each other.

(iii) There are large numbers of buyers.

(iv) The market demand curve for the product is given and is known to the cartel

(v) The cost curves of the firms are different but are known to the cartel.

(vi) The cartel aims at joint profit maximization.

Under these assumptions, the joint profit maximization model can be explained. As you know that the central agency knows the costs of the firms and it can calculate the market demand curve and the corresponding MR curve. From the horizontal summation of the MC curves of the individual firms the market MC curve can be derived. The central agency will decide the price at the point of intersection between the industry MR and the MC curves. The cost structures of the two firms are shown in the figures (a) and (b). From the horizontal summation of the MC curves we obtain the market MC curve that is implied by the profit maximization goal of the cartel: each level of industry output should be produced at the least possible cost. Thus if we add the outputs of A and B that can be produced at the same MC, clearly the resulting total output is the output that can be produced at this common lowest cost. Given the market demands D, in the figure (c) the monopoly solution, which maximizes the joint profits, is determined by the intersection of MC and MR at point E.

With the formation of a cartel, the situation is similar to the multi-plant monopolist. Industry price-output is determined by equating MC and MR in figure (c). Allocation of this output between the two firms A and B is determined by equating the common MR determined in figure (c) with MC_1 and MC_2. Thus, the equilibrium condition is $MR = MC_1 = MC_2$. Thus the firm A will produce output Q_1 and firm B will produce output Q_2. It is to be noted here that the firm with lowest costs would produce a larger amount of output. However, this does not mean that the firm will also take the largest share of profit. The total industry profit is the sum of the profits from the output of the two firms, denoted by the shaded areas of figures (a) and (b). The distribution of the profit of the cartel will be decided by the central agency of the cartel maximizing the joint profit.

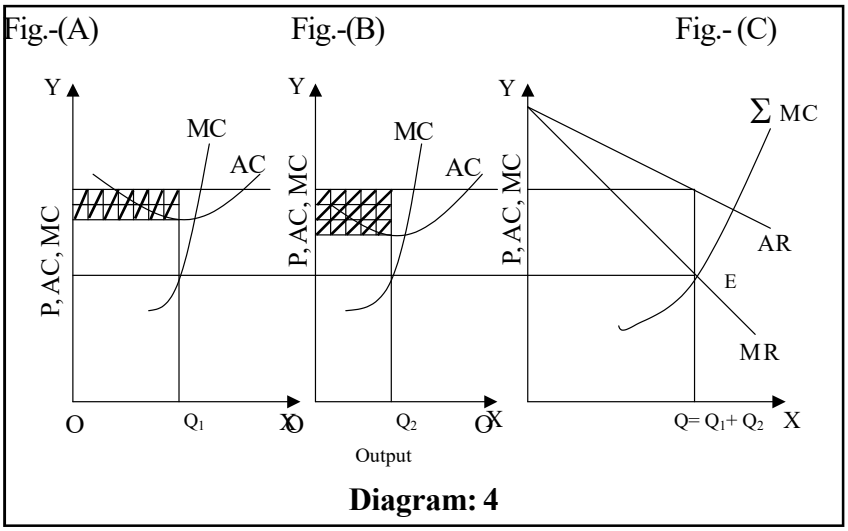


Diagram: 4

Though theoretically the cartel can maximize joint profits, in practice it is not possible for several reasons. Firstly, mistakes may arise in the estimation of the market demand. Secondly, mistakes may also arise in the estimation of the MC curves. Thirdly, the existence of high cost firms sometimes set an obstacle towards joint profit maximization. Fourthly, the cartel may not also maximize total profit for fear of government intervention or for fear of entry. Fifthly, sometimes the cartel likes to maintain a good public image and for this it may not maximize total profit.

Market Sharing Cartels:

This form of collusion is more common in practice because it is more popular. The firms agree to share the market, but keep a considerable degree of freedom concerning the style of their output, their selling activities and other decisions. Market sharing cartels are the form of cartels wherein the firms agrees to share markets on mutual agreement. There are two basic methods for sharing markets. These are: (a) non-price competition agreement and (b) determination of quotas.

(**a) Non-price Competition Agreement :** In this form of cartel, the member firms agree on a common price, at which each of them can sell any quantity demanded. The price is set by bargaining, with the low cost firms pressing for a low price and the high cost firms for a high price. And ultimately all the firm came in to a decision on the basis of which they set a price that allow some profits to all the member firms. The firms agree not to sell products below the cartel price, but they are free to vary their style of their products and the selling techniques.

This form of cartel is indeed loose because it is unstable than the complete cartel aiming joint profit maximization. If all firms have the same cost structure then the price will be set at monopoly level. But the cost differences make the cartels inherently unstable as the low cost firms have a tendency to break away from the cartel.

(b) Sharing of Market by Agreement on Quotas: The second form of market sharing cartel is the sharing of markets by agreement on quotas. In this form of cartel the member firms agree on sharing the market on the basis of the quotas. If the firms have the same cost then the monopoly situation occurs with market being shared with others equally. However, if the costs are different, the quotas and shares of the markets will differ. Allocation of quota-share on the basis of cost will again be unstable. Shares in the case of cost differentials are decided by bargaining. The final quota of each firm will depend on its costs as well as on the bargaining skill.

Price Leadership:

Price leadership is another form of collusion in oligopoly market. In this form of cartel, one firm set the price and the other firms follow it because it is advantageous to them or because they prefer to avoid uncertainty. There are various types of price leadership. However the most common are:

(a) Price Leadership by a low-cost firm

(b) Price leadership by a large (dominant) firm, and

(c) Barometric price leadership.

The price leader is assumed to set his price by applying marginalistic rules, which is by equating MR and the MC. As a result the price leader can maximize profit. The other firms are price takers who take the price set by the leader. At this price the followers may not maximize their profits. If they do, it will be by accident rather by their own independent decisions.

In price leadership by a low cost firm, the oligopolistic firm having lower costs than the other firms sets the price while the other member firms follow it. Thus the low cost firm becomes the price leader. On the other hand, in case of the dominant firm price leadership model, the dominant firm fixes the price and the other member firms have to follow the price set by the dominant firm. Thus, in this form of cartel the dominant firm becomes the price leader.

Barometric price leadership model is another form of oligopolistic model. Here there is no leader firm as such but the firm with the wisest management announces the price change first and the other firms in the industry have to follow it. In this type of model, it is not necessary that the leader firm is either a low-cost firm or a dominant firm.

5.6 THE KINKED DEMAND CURVE:

Paul M. Sweezy developed the concept of kinky demand curve model of oligopoly to explain the rigidity of price in an oligopoly market. The kinky demand model is based on the following assumptions:

(i) There are many firms in the oligopolistic industry.

(ii) Each produces a product, which is a close substitute for that of the other firm.

(iii) Product qualities are constant, advertising expenditures are zero.

(iv) Each oligopolist believes that if he lowers the price of his product, his rivals will also lower the prices of their products and that if he rises, they will maintain the prices at the existing levels.

Under the above assumptions, we can explain the kinked demand model. In kinky demand model the demand curve faced by an individual seller has a kink at the initial price quantity situation. The kinked shaped of the demand curve is based on the assumption that the rivals react differently to a rise or a fall in the price of the product. Further it is assumed that if the producer increases the price of his product, the rival reacts with keeping the price of intact while if his rival reduces the price of the product, he will react by reducing the price of his product also. Thus when the individual producer increases price of the products the rivals will not increase the price of their products so that the sales of the producer increasing the price will be reduced considerably. This means that the demand curve is relatively elastic when price is rising. On the other hand, when the individual producer reduces the price of his product, the rivals will reduce prices of their products so that the producer who reduces price will not gain much from his action. Thus, the kinky demand curve is based on the assumption that a rise in price by one producer will not be followed by a rise in the price of the other producers and vice versa. The concept of kinky demand curve can be explained with the help of the following diagram:

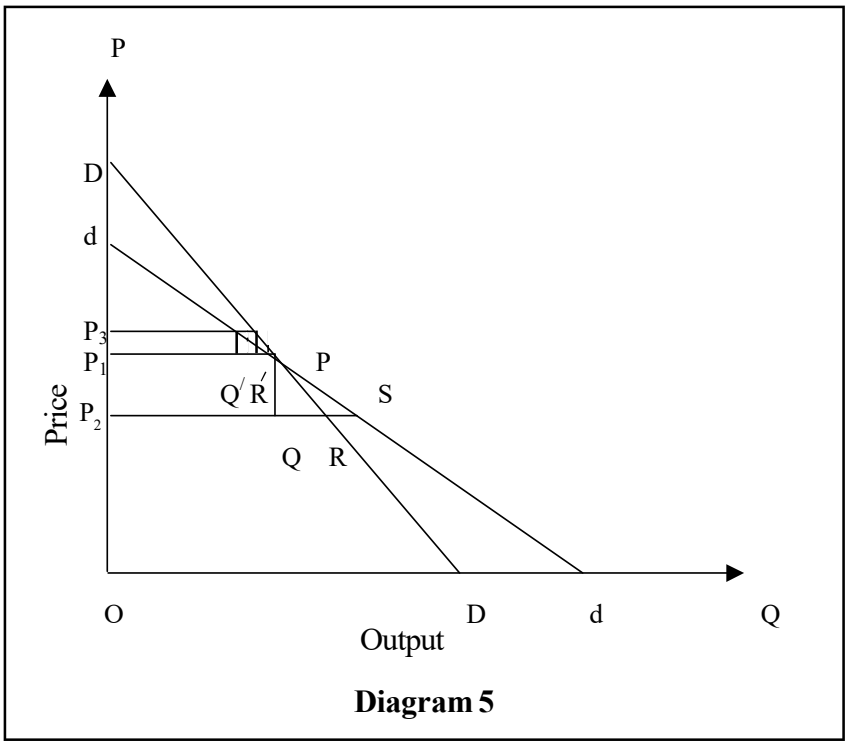


Diagram 5

In the above diagram 5, we have drawn two demand curves dd and DD. The demand curve dd is drawn on the assumption that when one seller changes his price, the other sellers do not change their prices and keep their prices unaffected. The demand curve DD is drawn on the assumption that when one seller changes price, the other sellers also change their price in the same direction. The demand curve dd and DD intersect at point P. In kinky demand analysis it has been assumed that the rise in price will be unrivaled while a fall in the price will be rivaled. Hence the demand curve is dPD at the point of kink P. Let us take a situation where the price is changed from OP_1 to OP_2. If the other sellers also reduce their price the sell of the quantity will increase by QR amount. But if the other sellers do not increase the price of their products the amount to be sold will increase by QS. Similarly, when the price will be increased from OP_1 to OP_3 the quantity will be reduced by $PQ^/$ if other sellers do not increase their prices and the quantity demanded would be $PR^/$ if other sellers also increase prices. Since it is assumed that price decrease by a firm will be matched by a price reduction by rivals but the rivals do not match an increase in the price the demand curve is dPD, which has a kink at point P. The upper section of the kinky demand curve has higher price elasticity than the lower part.

The position of the demand curve is determined by the location of OP_1, the price at which the oligopolist now happens to be selling his product. The price OP_1 is the datum and it is not determined in the model.

Let us now consider the implication of a kink in the demand curve by the seller in the market. If the demand curve is kinked, the corresponding MR curve will be discontinuous. In Diagram 6 dA

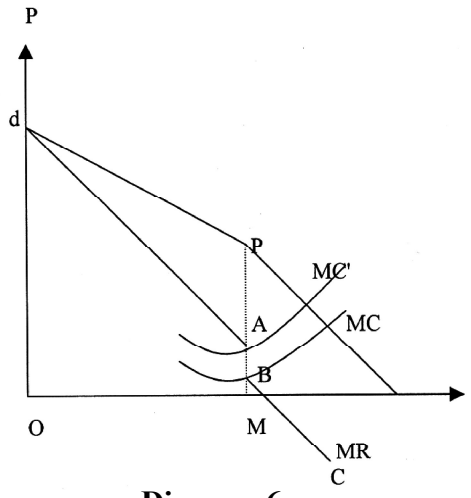


Diagram 6

the demand curve, while the BC portion of the MR curve corresponds to the PD portion of the demand curve. The length of the discontinuity is equal to AB. The point P on the demand curve has two elasticities of demand. If we think that P is a point on dd, we get one elasticity of demand and if we think that P is a point on DD we get another elasticity of demand. The greater the difference between the two elasticities of demand, the greater will be the length of the discontinuity. This is so because, we know from the relation between MR, price and the absolute value of demand elasticity (e) that MR=price [1-1/ e]. Now at the point P, both the demand curves DD and dd have the same output level. The MR will therefore be different because of the differences in the elasticities of demand. Only when the elasticities are equal at point P, the discontinuous range also disappears.

Suppose now that the MC curve of the firm passes through the discontinuous range of the MR curve, in this case, we cannot say that MR equals MC at the equilibrium point. Equality of MR and MC is not possible. All that we can say is that MR cannot be less than MC. In this situation the price and quantity remain the same at the kink point. Even if the Me curve shifts but passes through the discontinuous range AB, the pricequantity combination will remain constant. The price-quantity combination given by the point of the kink remains more or less stable in the oligopoly market. The price rise or the price fall is not profitable for a single seller because of the asymmetrical behavior of sellers for a price rise or a price decrease.

The equilibrium of the firm is defined by the point of the kink because for any output level less than OM, MC is below MR, while for any output level greater than OM, MC is greater than MR. Thus total profit is maximized at the kink though the profit maximizing condition (MR=MC) is not fulfilled at the kink point.

The discontinuity between A and B of the MR curve implies that there is a range within which costs may change without affecting the equilibrium price and output of the firm. This level of price and output is compatible with a wide range of costs. Thus the kink can explain why price and output will not change despite changes in cost within the rangeAB.

If the demand curve is kinked, a shift in the market demand upwards or downwards, will affect the volume of output but not the level of price, so long as the MC curve passes through the discontinuous range of the new MR curve. In this case the demand curve will shift but the kink points lie on the horizontal straight line. As the market expands, the firm will not raise its price, although output will increase.

In conclusion it can, therefore, be said that the kinked demand analysis as a method of price-output determination is not analytically sound. But it can be accepted as a reasonble explanation for the rigidity of price and output in the oligopolistic markets.

CHECK YOUR PROGRESS

1. Define the term oligopoly.

.....

.....

.....

.....

.....

.....

.....

.....

.....

2. What are the different types of collusive oligopoly model?

.....

.....

.....

.....

.....

.....

.....

3. Mention the factors on which the success of a cartel depends?

.....

.....

.....

.....

.....

.....

4. Write down a brief note on the kinky demand curve model.

.....

.....

.....

.....

.....

.....

.....

.....

.....

5.7 GAME THEORY

Von Neumann developed the theory of game in 1928. However, it was only after 1944 Von Neumann and Oskar Morgenstern published their now well known "Theory of Games and Economic Behaviour" that the theory received the proper attention. A more recent and in-depth presentation of the game theory with economic applications to economic problems is found in the works of M.J. Osborne and A. Rubinstein (1994). The other important contributors to the game theory were John Nash and Thomas Schelling. In general, game theory is concerned with the choice of an optimal strategy in conflicting situations. It deals with the mathematical analysis of competitive problems and is based on the minimax strategy put forwarded by Von Neumann, which implies that each competitor will act so as to minimize his maximum loss (or maximize his minimum gain). In economics game theory can help a doupolist or an oligopolist to choose the course of action that maximizes the benefit or profit after considering all the possible actions of its rival. In both form of markets it is very difficult to arrive at a determinate solution as the interests and strategies of the participants are conflicting. Thus game theory, for example, can help an oligopolistic firm to decide: (1) the conditions under which lowering of its price will not result in price war, (2) whether the firm should build excess capacity to discourage the entry of new firm into the industry, even though it may incur lose in the short run; (3) why cheating leads to break up of cartels. Game theory can be of great use in the analysis of conflicting situations like these. *In short, game theory can be defined as the modeling of economic decisions by games to gain competitive advantage over the rival or to minimize the potential harm from a strategic move by the rival, whose outcomes depends on the decisions taken by the two or more agents or players, each having to make decisions without knowing what strategies each of them are following.*

Assumptions of the Game Theory:

The theory of game is based on certain assumptions, which are mentioned below:

i) There is finite number of participants called players.

ii) A list of finite or infinite number of possible courses of action is available to each player. The list need not be same for each player. Such a game is said to be in normal form.

iii) A play is played when each player chooses one of his courses of action, the choices are assumed to be made simultaneously, so that no player knows his opponents choice until he has decided his own course of action.

iv) Every play is associated with an outcome, known as payoff, which determines a set of gains, one to each player. Here a loss is considered as a negative gain. Thus, after each play of the game, one player pays to others an amount determined by the courses of actions chosen.

Besides above, few more assumptions of game theory are as under:

i) All players act rationally and intelligently. It means that each player has a Consistent ranking over the all the possible outcomes and calculates the strategy that serves his interest best. Thus they are perfect calculators and flawless followers of best strategies.

ii) Each player attempts to maximize his gain or minimize loss.

iii) Each player makes individual decision without direct communication.

iv) Each player knows complete relevant information.

Before discussing about game theory in detail, let us first throw light on the some basic concepts of game theory:

Every game theory involves players, strategies, and payoffs. Let us first define this three:

Player: Each of the participants in a game is called a player. For example in a duopoly there are two players who can participate in the game.

Play: In game theory, a play results when each player has chosen a course of action.

Strategy: The decision rule by which a player determines his course of action is called a strategy. Simply strategies are the choices available to the players. It clearly defines specific course of action in value terms for the policy variable. For example, a strategy may consist of setting a price of Rs. 5.00, spending Rs 3000 on advertising, making a change in the packaging of the product, and selling the product in the super markets. Another strategy may involve keeping price unchanged, spending Rs. 1000 on advertisement and spending Rs. 3000.00 on research and

development activities for a new product and so on. On the other hand the rivals will take their own course of action as against each of these strategies separately. They may take same course of action or may not take same course of action. However, to reach the decision regarding which strategy is to use, neither player needs to know the other's strategy.

Strategy can be of two types: (a) Pure Strategy and (b) Mixed Strategy. A strategy is called *pure strategy* if a player decides to use only one particular course of action during every play. A pure strategy is usually represented by a number with which course of action is associated. No randomness is associated with this strategy. On the other hand, if a player decided in advanced to use all the courses of action or some of available courses of action in some fixed proportion, then the player is said to use *mixed strategy*. A mixed strategy is a selection among pure strategies with some fixed proportion.

Pay-off : The payoff is the outcome or consequence of each strategy. While taking any strategy by a firm some alternative strategies are available to the competitive firms and payoff is the result of the each of the combination of strategies by the firms. The payoff is usually expressed in terms of the profits or loses.

On the other hand the payoff matrix is a table showing the amounts received by the player named at the left hand side after all possible plays of the game. The player named at the top of the table makes the payment.

For example, if player A has m-courses of action and the player B has n-courses of actions, then a payoff matrix can be constructed as given below:

i) Row designations for each matrix are the courses of action available to player A.

ii) Column designations for matrix are the courses of action available to B.

iii) With a two-person zero-sum game, the cell entries in B's payoff matrix will be the negative of corresponding entries in A's payoff matrix and the matrices will appear as follows:

Table-1

Payoff Matrix of Player A

		Player B					
		1	2	3	…	j….	n
Player A	1	a_{11}	a_{12}	a_{13}	…	a_{1j} …	a_{1n}
	2	a_{21}	a_{22}	a_{23}	…	a_{2j} …	a_{2n}
	3	a_{31}	a_{32}	a_{33}	…	a_{3j} …	a_{3n}
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	i	a_{i1}	a_{i2}	a_{i3}	..	a_{ij}	a_{in}
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	m	a_{m1}	a_{m2}	a_{m3}	…	a_{nj}	a_{mn}

Table-2

Payoff Matrix of Player B

		Player B					
		1	2	3	…	j….	n
Player A	1	$-a_{11}$	$-a_{12}$	$-a_{13}$	…	$-a_{1j}$ …	$-a_{1n}$
	2	$-a_{21}$	$-a_{22}$	$-a_{23}$	…	$-a_{2j}$ …	$-a_{2n}$
	3	$-a_{31}$	$-a_{32}$	$-a_{33}$	…	$-a_{3j}$ …	$-a_{3n}$
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	i	$-a_{i1}$	$-a_{i2}$	$-a_{i3}$	..	$-a_{ij}$	$-a_{in}$
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	m	$-a_{m1}$	$-a_{m2}$	$-a_{m3}$	…	$-a_{nj}$	$-a_{mn}$

Thus the sum of the payoff matrices of the two players A and B are zero. Here, the objective of the players is to determine the optimum strategies that result optimum payoff to each, irrespective of the strategy used by the other. In the above payoff matrices constructed for A and B reflects that whatever is the strategy followed by the Player A, the player B follows its own counter strategy out of the several alternatives available to it so

as to minimize his lose. Here negative payoff of the player B suggest that it is just the negative of the payoff matrix A. Usually, we do not construct the payoff matrix for B, as it is just negative of A.

5.7.1 SOME OTHER IMPORTANT CONCEPTS:

Competitive Game: A competitive situation is called a competitive game if it possesses the following six properties:

1. There are finite number of participants. The number of participants is $n \geq 2$. if n=2, it is called a two person game; and if $n > 2$, then it is called n person game.
2. Each person has a finite number of possible courses of action.
3. Each participants must know the possible courses of action available to others but must not know which of these will be chosen.
4. A play of the game occurs when each player chooses one of his courses of action. The choices are assumed to be made simultaneously, so that no participant no the choice of other until he has decided his own.
5. After all participants have chosen a course of action, their respective gains are finite.
6. The gain of the participant depends upon his actions as well as those of others.

Two-Person Zero Sum Game: A game with two players, where a gain of one player equals the loss to the other player is known as the *two person zero sum game*. In such type of game, interests of the two players are opposed so that the sum of their gains is zero. For example, if two players of chess agree that at the end of the game the loser would pay Rs 100 to the winner of the game, then it would mean a zero sum game since the gain of one is equal to the loss of the other. On the other hand, if there are n players and sum of the game is zero then it is called *n person zero sum game*.

Two person zero sum game is also called rectangular games because their payoff matrix is in the rectangular form. A two person zero sum game exhibits certain characteristics as follows:

1. Only two players participate.
2. Each player has finite number of possible courses of action.
3. Each specific strategy results in a payoff.
4. Total payoff to the two players at the end of each play is zero.

Non Zero Sum Game: A game is called a non zero sum game if the gains or loses of one firm do not come at the expenses of or provide equal benefit to the other firm. If for example increased advertisement results in the higher profits to both the firms, it is a positive sum game. On the other hand if increased advertisement results in rise in cost than revenue suggesting a declining profit. This is a case of negative sum of game. Both these are the example of non-zero sum game.

Cooperative Games: Games in which joint-action agreements are enforceable are called cooperative games.

Non-cooperative Games: Games in which enforcement of joint-action games are not possible and individual must be allowed to act in their own interest are called non-cooperative games.

Maximin and Minimax Strategy:

Consider a two person zero sum game involving the set of pure strategies $\alpha = \{A_1, A_2, A_3\}$ for player A and $\beta = \{B_1, B_2\}$ for player B and having the following payoff matrix for the player A.

		Player B		
		B_1	B_2	Row Minima r
Player A	A_1	9	2	2
	A_2	8	6	6
	A_3	6	4	4
Column Maxima c		9	6	

Suppose player A starts the game knowing fully well that whatever strategy he adopts, B will select that particular counter strategy which will minimize the payoff to A. Now if player A selects the strategy A_1, player B will reply by selecting B_2 , as this corresponds to the minimum payoff to A in the first row corresponding to A_2 . Similarly if A chooses the strategy A_2 , he may gain 8 or 6 depending upon the strategy chosen by B. However, A can guarantee a gain of at least minimum {8,6}=6 regardless of the strategy chosen by B. thus whatever strategy A may adopt, he can guarantee only minimum of the corresponding row payoffs. These corresponding to each $A_i \in \infty$ are indicated by forming a column vector r = {2,6,4} of the row minima. Naturally A would like to maximize his gain, which is just the largest component of r. in the above example, the selection of A_2 will give the maximum

{2,6,4}=6 of the minimum gains to A. we can call this gain as the maximin value of the game and the corresponding strategy is called as ***maximin strategy.***

On the other hand, player B wishes to minimize his loses. If he plays strategy B_1, his loss is at most max {9,8,6}=9 regardless of what strategy A has followed. He can lose no more than max {2,6,4}=6 if he plays B_2. These maximum loses corresponding to each B$\in \beta$ are indicated by forming a row vector C={9,6} of the column maxima. The smallest component of C represents the minimum possible loss to B whatever strategy he adopts. This minimum of maximum loses will be called the mimimax value of the game and the corresponding strategy is called as the ***minimax strategy.***

Thus from the above example it is seen that the maximum of row minimum is equal to the minimum of the column maxima. In symbols,

$$\text{Max} \ \{r_i\} = 6 = \min \ \{c_j\}$$

Or, $\max_i \left[\min_j \{a_{ij}\}\right] = 6 = \min_j \left[\max_i \{a_{ij}\}\right]$

The selection of maximin and minimax strategies by the players A and B was based upon the so called ***maximin-minimax principle***, which guarantees the best of the worst results. The corresponding pure strategies where both maximin and minimax value of the game are equal, called as the ***optimum strategies.***

SADDLE POINT:

A saddle point of a payoff matrix is that position in the matrix where the maximum of row minima coincides with the minimum of the column maxima. The corresponding payoff at the saddle point is called the *value of the game* and is obviously equal to the maximin and minimax values of the game. In general, the following rules are followed for determining saddle point:

a) Select the minimum element of each row of the payoff matrix and mark them.

b) Select the greatest elements of each column of the payoff matrix and mark them.

c) If there appears an element in the payoff matrix marked, the position of that element is a saddle point of the payoff matrix.

On the other hand, sometimes it is not possible to get saddle point by using pure strategy. Under such situation one has to use mixed strategies by mixing some or all of their possible courses of action to get the best possible strategy. However, in our discussion we will be restricting to the game of pure strategy only.

Example1:

Player A	Player B		
	B_1	B_2	B_3
A_1	1	3	1
A_2	0	-4	-3
A_3	1	5	-1

Solution:

From the given pay-off matrix we have to find out minimum payoff in each row and minimum payoff for each column. The row minima vector 'r' is obtained by writing the minimum payoff of each row. The largest component of 'r' represents the minimum value '$\underline{v}$'. The column maxima vector c is obtained by writing the maximum payoff of each column. The smallest component of c represents the minimax value $\underline{v}$.

Player A		Player B			
		B_1	B_2	B_3	Row Minima (r)
	A_1	1	3	1	1
	A_2	0	-4	-3	-4
	A_3	1	5	-1	-1
	Column Maxima (c)	1	5	1	

The matrix has two saddle points at positions (1,1) and (1,3). Thus the solution to the game is given by:

i) the optimum strategy for player A is A_1 .

ii) The optimum strategy for player B is B_1 either B_1 or B_2, i.e, B can use either of the two strategy.

iii) The value of the game is 1 for A and -1 for B.

Example2:

Player A		Player B		
		B_1	B_2	B_3
	A_1	6	8	6
	A_2	4	12	2

Solution:

From the given pay-off matrix we have to find out minimum payoff in each row and minimum payoff for each column. The row minima vector 'r' is obtained by writing the minimum payoff of each row. The largest component of 'r' represents the minimum value '$\underline{v}$'. The column maxima vector c is obtained by writing the maximum payoff of each column. The smallest component of c represents the minimax value $\underline{v}$.

Player A	Player B				
		B_1	B_2	B_3	Row Minima r
	A_1	6	8	6	6
	A_2	4	12	2	2
	Column Maxima c	6	12	6	

The matrix has two saddle points at positions (1,1) and (1,3), that is, saddle point exists at A_1B_1 and at A_1B_3 . Thus the solution to the game is given by:

iv) the optimum strategy for player A is A_1 .

v) The optimum strategy for player B is B_3 .

vi) The value of the game is 6 for A and -6 for B.

On the other hand in some game no saddle point occurs and under such circumstances it is not possible to find out its solution in terms of the maximin-minimax strategy. Games without saddle points are not strictly determined. The solution of such problems can be obtained by using mixed strategy. A mixed strategy refers to a combination of two or more strategies that are selected one at a time, according to pre-determined probabilities, that is, at the time of using the mixed strategy a player has to make decision to mix his choices among several alternatives in a certain ratio.

5.7.2 Dominant Strategy and Nash Equilibrium:

The dominant strategy is the optimal choice of the player, no mater

what his/her opponent does. It is a strategy that yields higher payoffs to a player for every strategy of his/her opponent. In other words if both the firms have dominant strategy, each of them can choose their own optimal strategy regardless what their opponents are doing. On the other hand, not all games have a dominant strategy for each player. Under such situation, Nash Equilibrium exists. According to Nash, firms reach their equilibrium state when they are doing their best, given what its competitors are doing. In terms of price theory, doing their best means maximizing profits and what others are doing means what rate of output their opponents are producing or what price they are charging or what advertising expenditure they are incurring to promote the sales of their products. When each firm is doing its best, given what others are doing, no one has any incentives to change its behaviour and hence equilibrium exists. Nash Equilibrium describes a set of strategies where each player believes that it is doing the best it can, given the strategy of the other player. More specifically, the Nash Equilibrium is a situation in which each player chooses an optimal strategy, given the strategy chosen by the other player. Cournot solution is an example of Nash Equilibrium. Not all games have Nash Equilibrium and some games have more than one.

5.7.3 ZERO-SUM GAME: CERTAINTY MODEL

We have already discussed about the meaning of zero sum game. In zero sum game, a firm's gain means the loss of the other firm. Thus any gain of one rival is offset by the loss of the other, and the net gain sums up to zero. Hence the name zero sum game. The certainty model is based on the following assumptions:

(1) the firms have a given well defined goal.

(2) each firm knows the strategies open to it and to its rival.

(3) each firm knows with certainty the payoffs of all combinations of the strategies being considered which implies the firm knows its total revenue, total costs and total profit from each combination of strategies.

(4) the actions chooses by the duopolists do not affect the total size of the market.

(5) each firm chooses its strategy expecting the worst from the rival

(6) in zero sum game there is no incentive for collusion

Now on the basis of the above assumptions how equilibrium

in case of a duopoly model the can be reached be find out. However, first we need information on the payoff matrix of the two firms. Adopting the minimax strategy as we have discussed above can do the solution. The solution is called as the saddle point.

5.7.4 ZERO-SUM GAME: UNCERTAINTY MODEL

The assumption that each firm knows with certainty the exact value of the payoff of each strategy is unrealistic. The most probable situation in real world is that when a firm adopts a particular strategy, may expect a range of results for each counter-strategy followed by its rival, each with an associated probabilities. Therefore the payoff matrix is constructed so as to include the expected value of each payoff. The expected value is the sum of the products of the possible outcomes of a pair of strategies (adopted by the two firms) each multiplied by its probability. Mathematically,

$$E(G_{IJ}) = g_{1i}P_1 + g_{2i}P_2 + ... + g_{ni}p_n$$

$$= \sum_{s=1}^{n} g_{si}P_s$$

where g_{si} the sth of the n possible outcomes of strategy i of firm I (given that Firm II has chosen strategy j)

P_s = the probability of the sth outcome of the strategy i.

5.7.5 CHARACTERISTICS OF THE GAME THEORY:

The game theory possess certain characteristics which are mentioned below:

1) *Chance of strategy:* A game may be game of strategy or game of chance. If in a game activities are determined by skill, it is said to be a game of strategy; and if they are determined by chance, it is a game of chance.

2) *Number of persons:* A game is called an n-person game if the number of persons playing the game is n. the person means an individual or a group aiming at a particular objective.

3) *Number of activities:* The number of activities in a game may be finite or infinite.

4) *Number of alternatives:* Number of alternatives available to each person in a particu8lar activity may also be finite or infinite. A finite game has a finite number of activities, each involving a finite number of alternatives, otherwise the game is said to be infinite.

5) *Information to the players:* Information to the players about the past activities of other players is completely available, partly available, or not available at all.

6) *Payoff:* A quantitative measure of satisfaction a person gets at the end of each play is called a payoff. It is a real-valued function of variables in the game. Let be the payoff to the player , 1£i£n, is an n-person game. If , then the game is said to be a non-zero sum game. Payoff may be such that the gains of some players may and may not be direct losses of others players.

5.7.6 LIMITATIONS OF GAME THEORY:

Game theory, which was initially received in literature with great enthusiasm as holding promise, has been found to have a lot of limitations. Some important limitations of the game theory are mentioned bellow:

i) The assumption that the players have the knowledge about their own payoffs and payoffs of their opponents is rather unrealistic. He can only make a guess of his own and his rival strategies.

ii) As the number of players increases in the game, the analysis of the gaming strategies become increasingly difficult and complex. In practice, there are many firms in an oligopoly situation and game theory can not be very helpful in such situations.

iii) The assumptions of maximin and minimax show that the players are risk-averter and have complete knowledge of the strategies. These do not seem practical.

iv) Rather than each player in an oligopoly situation making under uncertain conditions, the players will allow each other to share the secrets of business in order to work out a collusion and under such situations mixed strategies are not very helpful.

5.7.7 IMPORTANCE OF THE GAME THEORY:

Game theory as applied to value theory possesses the following merits:

i) Game theory shows the importance to duopolists of finding some way to agree. It helps to explain why duopoly prices tend to be administered in a rigid way. If prices were to change often, tacit agreements would not be found and would be difficult to enforce.

ii) Game theory also highlights the importance of self-interest in the business world. In game theory, self-interest is routed through the mechanism of economic competition to bring the system to the saddle point. This shows the existence of perfectly competitive market.

iii) Game theory tries to explain how duopoly problem can not be determined

iv) Game theory has been used to explain the market equilibrium when more than two firms are involved. The solution lies in either collusion or non-collusion. These are known as cooperative non-constant sum game and non-cooperative game respectively.

v) "Prisoner's Dilemma" in game theory points towards collective decision-making and the need for cooperation and common rules of road.

vi) The importance of the pay-off values lies in predicting the outcomes of a series of alternative choices on the part of the player. Thus a perfect knowledge of the pay-off matrix to a player implies perfect predictions of all factors affecting the outcome of alternative strategies. Moreover, the minimax principle shows to the player the next course of action, which would minimize the losses if the worst possible situation arose.

vii) Game theory is also helpful in solving the problems of business, labour and management. As a matter of fact, a business always tries to guess the strategy of his opponents so as to implement his plans more effectively.

viii) Last but not the least, there are certain economic problems, which involve risk and technical relations. They can be handled with the help of the theory of games. Problems of linear programming and activity analysis can provide the main basis for economic application of the theory of games.

5.8 THEORY OF LIMIT PRICING:

Bain first of all formulated the limit pricing theory in an article published in 1949, several years before his major work Barriers to New Competition was published in 1956. Bain's aim in his early article was to explain why firms over a long period of time were keeping their price at a level of demand where the elasticity was below unity, that is, they did not charge the price, which would maximize their revenue. His conclusion was that the traditional theory was unable to explain this empirical fact due to the omission from the pricing decision of an important factor, name the threat of potential entry.

Bain regards limit price as the highest price above the competitive price charged by the established firms. Such a price acts as barriers to entry of the new firm to the industry. Barriers to entry may be defined as the advantage accruing by the established firms in an industry over the new entrants. He starts his model by defining the condition of entry. It is the premium or percentage by which established firms can raise the price above the competitive price without attracting entry of the new firm to the group. Symbolically, the condition of entry:

$$E = \frac{P_L - P_C}{P_C}$$

and $$P_L = P_C(1 + E)$$

where P_L is the limit price and P_C is the competitive price.

After Bain, there are several developments to the theory of limit pricing: (i) the model of Sylos-Labini; (ii) the model of Franco Modigliani; (iii) the model of Bhagawati and (iv) the model of Pashigian.

5.9 BAUMOL'S SALES REVENUE MAXIMIZATION MODEL

Prof. William J. Baumol has built a model of oligopoly in which the firm assumed to compromise between the money volume of sales and profits. It is assumed here that an entrepreneur ordinarily does not pursue the objective of profit maximizing alone rather they like to pursue and combine with minimum level of profit some other objectives i.e. the maximization of firms total revenue by maximizing sales.

According to Baumol's model, a common belief among entrepreneurs is that large sales mean large size and these in turn mean large profits. But they also realize after a point, increases in sales can not had only at the cost of profits.

Prof. Baumol's model can be discussed with the help of the following diagram. The X-axis shows total output and Y-axis shows total revenue obtained from sales. The curve TR shows total revenues, which the firm would get with different levels of sales by adjusting its price. Total costs are shown by the curve TC. The vertical difference between TR and TC total profit, which increases at first as sales increase but after point, starts diminishing. The level of profit obtained by different levels of output has been traced in the diagram by the curve TP that has a maximum at the output OA. The horizontal line KK shows the level of maximum acceptable profits it would produce sell an output OA. On the other hand, if the firm to maximize its sales volume in terms of money (TR), it would produce the output OC. But under consideration in the model wants to have the maximum sales consistent with minimum acceptable level of profit shown by KK. Such an output as gives it maximum revenue along with the minimum profit level is the output OB. At this output, total revenue is BR_1, which is slightly less than

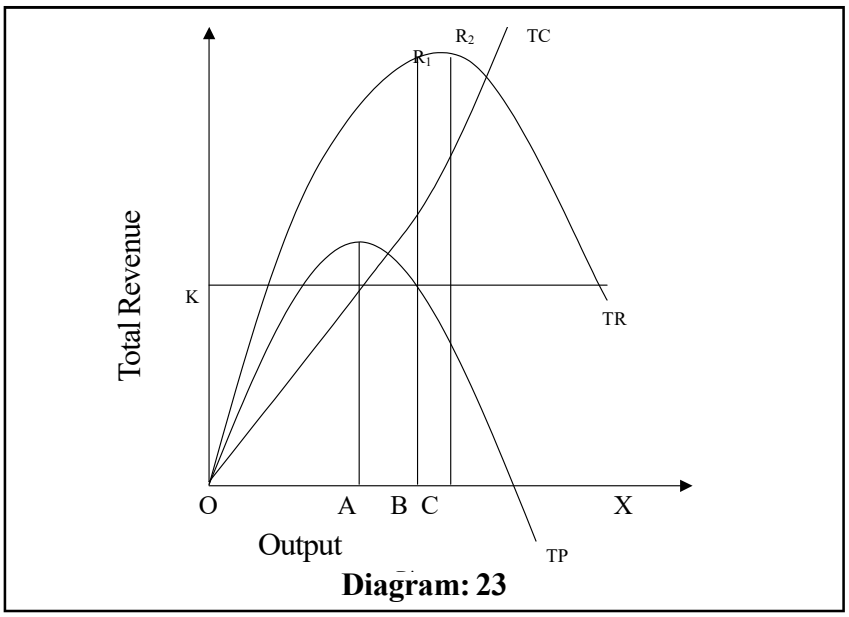


Diagram: 23

the maximum revenue CR_2, which the firm can earn by selling output OC. The result of the combination of the profit-maximization motive with the sales maximization objective is that output is more and price set less that what it would be if the firm were to maximize its sales revenue.

CHECK YOUR PROGRESS

1. What is a game?

.....
.....
.....
.....
.....

2. What is payoff?

.....
.....
.....
.....
.....
.....
.....

3. Write down the characteristics of the game theory.

.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....

4. What is meant by limit pricing?

.....
.....
.....
.....
.....

SUGGESTED READINGS

1. Modern Microeconomics, A. Koutsoyiannis, Macmillan
2. Advanced Economic Theory, H.L Ahuja, S. Chand
3. Microeconomics: Theory and Application, Salvatore, Oxford.
4. Microeconomics: Theory and Applications, Maddala & Miller, Tata Mcgraw-Hill.
5. Operation Research: Taha, PHI.
6. Operation Research, Hira and Gupta, S. Chand.

LET US SUM UP

In this particular unit we have discussed about oligopoly, game theory, limit pricing and Baumol's sales revenue maximization model. Oligopoly is a form of market structure wherein there are few sellers with either a homogeneous product or a differentiated product. Oligopoly is of two types. If the products are homogeneous, it is called a pure oligopoly and if the products are heterogeneous then we have a differentiated oligopoly. We have also discussed about the game theory. In general, game theory is concerned with the choice of an optimal strategy in conflicting situations. It deals with the mathematical analysis of competitive problems and is based on the minimax strategy put forwarded by Von Neumann, which implies that each competitor will act so as to minimize his maximum loss (or maximize his minimum gain). In economics game theory can help a doupolist or an oligopolist to choose the course of action that maximizes the benefit or profit after considering all the possible actions of its rival. In both form of markets it is very difficult to arrive at a determinate solution as the interests and strategies of the participants are conflicting. Lastly we have discussed about the limit pricing theory and the Baumol's sales revenue maximization theory.

KEY WORDS

Oligopoly: It is a form of market structure wherein there are few sellers with either a homogeneous product or a differentiated product.

Cartel: Cartel is a group of firms acting together to control output and price.

Game theory: It can be defined as the modeling of economic decisions by games to gain competitive advantage over the rival or

to minimize the potential harm from a strategic move by the rival, whose outcomes depends on the decisions taken by the two or more agents or players, each having to make decisions without knowing what strategies each of them are following.

Player: Each of the participants in a game is called a player.

Play: In game theory, a play results when each player has chosen a course of action.

Strategy: The decision rule by which a player determines his course of action is called a strategy.

Payoff: The payoff is the outcome or consequence of each strategy. While taking any strategy by a firm some alternative strategies are available to the competitive firms and payoff is the result of the each of the combination of strategies by the firms.

Two-Person Zero Sum Game: A game with two players, where a gain of one player equals the loss to the other player is known as the two person zero sum game.

Non Zero Sum Game: A game is called a non zero sum game if the gains or loses of one firm do not come at the expenses of or provide equal benefit to the other firm.

Saddle Point: A saddle point of a payoff matrix is that position in the matrix where the maximum of row minima coincides with the minimum of the column maxima.

SELF LEARNING MATERIAL

ECONOMICS

COURSE : ECO - 101

MICROECONOMIC THEORY

BLOCK - IV & V

Directorate of Distance Education
DIBRUGARH UNIVERSITY
DIBRUGARH - 786 004

ECONOMICS

COURSE : ECO - 101

MICROECONOMIC THEORY

Contributors :

Prof. J. K. Gogoi
P. P. Buragohain
J. Borgohain
Department of Economics
Dibrugarh University

Editor :

Prof. K. C. Borah
Department of Economics
Dibrugarh University

Reprint : August, 2015
Copies printed : 1000

Published on behalf of the Directorate of Distance Education, Dibrugarh University by the Director, DDE, D.U. and printed at Maliyata Offset Press, Mirza, Kamrup (Assam)

MICROECONOMIC THEORY

BLOCK - IV

THEORY OF DISTRIBUTION AND LINEAR PROGRAMMING

BLOCK - V

WELFARE ECONOMICS

UNIT 1 THEORY OF DISTRIBUTION

Structure

1.0 OBJECTIVES

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

- state and explain the marginal productivity theorem as well the Euler's theorem.
- explain the basic concepts and different theories of rent, wage, interest and profit.

1.1 INTRODUCTION

Production of a commodity or service is net outcome of combined efforts of different factors namely, land, labour, capital and entrepreneur. The distribution is an important part of the economic analysis. In distribution we study how the factors of production are rewarded for the services they rendered. Here an attempt has been made to explain the marginal productivity theory

of distribution, Euler's theorem and adding-up controversy, the various concepts of factors of production and the related theories with them.

1.2 Marginal Productivity Theory of Distribution

The marginal productivity theory is a bold attempt on the part of the economists to evolve a general theory of distribution, which will explain the determination of factor prices such as wages, rent, interests and profits. It serves as general theory in terms of which rewards of all factors of production could be explained. The concept of marginal productivity theory goes back to West and Ricardo. But they failed to develop a systematic theory of distribution. Apart from the Ricardian theory of rent and stray statements of the writers like Von Thunen and Mountfield Longfield, the theory did not gain much popularity until the last quarter of the 19th century when it was developed independently by Javon, Wicksteed and Marshal in England, Stuartwood and J. B. Clark in the U.S.A. and Warlas, Barone and others on the continent.

The marginal productivity theory is, in fact the neo-classical theory of distribution derived from the 'marginal principle' of the Ricardian theory of distribution.

Statement of the theory :

The marginal productivity theory states that 'in equilibrium each productive agent will be rewarded in accordance with its marginal productivity'.

Assumptions

The marginal productivity theory distribution is based on certain assumptions, which are mentioned as below:

(i) It assumes that all units of factor service are homogeneous.

(ii) They can be substituted for each other.

(iii) There is perfect mobility of factors as between different places and employment.

(iv) There is perfect competition both in the factor market and commodity market.

(v) There is full employment on factors and resources.

(vi) Technique of production is assumed to be remaining same.

(vii) The various units of factor services are divisible.

(viii) It is applicable in the long run.

Explanation:

According to the marginal productivity theory distribution, the reward of each of the factors of production in the long run is determined by its marginal productivity. By marginal productivity we meant the increment made to total product by employing an additional unit of factor, keeping all other factors constant. Now if we multiply this increase in output by the prevailing price of the product, we will get the marginal value of that product. Thus the Value of Marginal Product is equal to the Marginal Physical Product of a commodity multiplied by its Price. If the reward is paid at a rate lower than the marginal productivity, then it will not be advantageous for the factor to continue in the industry and it will move to another industry. But if reward is more than the marginal productivity, then it will replace the factor by another, whose marginal productivity is higher or whose cost is relatively lower. It is thus clear that the reward of each factor of production would tend to be equal to the marginal product.

But we are not much interested in the marginal physical product as in the money, which earns from the sale of this product. In other words, we are more interested in the money value of the marginal physical productivity of a factor. Thus the concept of marginal productivity leads to the marginal revenue productivity and it is better to measure marginal productivity in terms of marginal revenue productivity which is the addition made to total revenue resulting from the employment of an additional unit of factor, the other factor remaining the constant. In other words, marginal product implies the marginal physical product of the factor multiplied by the marginal revenue.

In equilibrium, the price of a factor service must be equal to its marginal productivity. If the marginal revenue product of a factor unit is more than its price (cost of employing it), it will be profitable for the firm to employ more units of this factor. As more units are employed, the marginal revenue product diminishes until it equalizes the price. This is the point of maximum profits for the firm. But, if more factor units are employed beyond this paint, the marginal revenue product will fall below the price and the firm will incur loss. As a result some firms will leave the industry.

In the ultimate analysis, however, the price of a factor service must equal its marginal as well as average revenue productivity. If at any time the price of a factor unit is higher than average revenue

productivity, the firm will be incurring losses. As a result, some of the firm will leave the industry thereby the price of the factor service will fall to the level of maximum average revenue productivity. On the other hand, if the price is less than the average revenue productivity, the firm will earn extra profits. Attracted by this profits, new firms will enter into the industry and compete for this factor service. This will tend to push the price upward to the level of average revenue productivity. These can be deviation from this equilibrium position in the short run, but in the long run the price of a factor service must equal marginal as well average revenue productivity. This can be explained with the help of the following diagram:

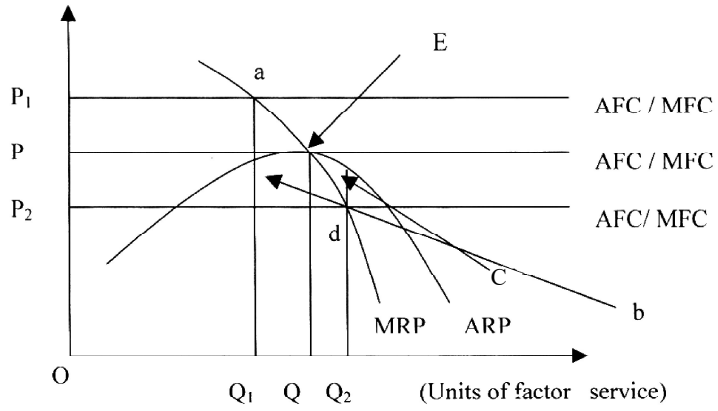


Diagram: 1

In the Diagram-1, at point E, ARP=MRP, and both are equal to average reward (average factor cost) and marginal reward (marginal factor cost) of the factor service. Thus each factor service will be paid OP price for OQ units. Suppose, the factor price rises to OP_1, at this price, firm will be incurring *ab* per unit of loss, as the price being paid to factor units is greater than the average revenue productivity (ARP). This will induce some firm to leave the industry and the factor price would again fall down to E. on the other hand, if factor price falls to OP_2, firms will be gaining *dc* per unit profit. When attracted by it some new firms enter in to the industry, price will again rise up to OP. These price variations are only possible in the short run. In long run, equilibrium position E will be stay on.

Criticisms:

The marginal productivity analysis suffers from the following drawbacks:

i) The theory assumes that all the factor units are divisible and therefore can be increased by small quantities. This is not

correct. Because, it is not possible to vary individual, large or lumpy factor.

ii) The theory assumes that all the units of a factor of production are homogeneous. But this is not correct. Because it is not possible to vary individual, large or lumpy factor.

iii) The theory is based on the assumption of perfect competition. In real life imperfect competition is the rule. Perfect competition is a myth and hence the application of the theory too in economic world is myth.

iv) The theory assumes that units of factors of production are perfectly mobile as between different regions and employments. However, factors are not mobile freely as between different regions and employment.

v) The marginal productivity theory also assumes constant returns to scale. However, the most economic activities in business and industrial world fall under either the increasing or decreasing returns. Constant returns to scale is only an exception.

vi) The theory neglects technical progress. But in real world, technical progress does take place.

The marginal productivity theory is not an adequate explanation of the determination of the pricing of factor services. It simply states the demand side of the factor pricing and therefore is one sided. "It (marginal productivity theory) is not a theory that explains wages, rents or interest, on the contrary, it simply explains how factors of production are hired by firms, once their prices are known."

Implication of the Theory:

The marginal productivity theory was not propounded for its own sake but was rather meant to support some type of conclusions, which had important implications:

i) Perhaps the most important theoretical conclusion out of it was the product exhaustion theorem built up by Philip Wickstead. This is also known as adding up controversy. It was the proposition that under perfect competition, where all the factors are rewarded according to their marginal products, the total product is exhausted by payments. There is nothing surplus left behind.

ii) Trade unions cannot do anything about wage rate because it is determined by a worker's marginal productivity to a firm.

1.3 EULER'S THEOREM AND ADDING UP CONTROVERSY :

Adding up theorem is one of the important issues in the theories of distribution. The proponents of the marginal productivity theory concluded that every factor is paid according to their marginal product. In other words, their conclusion was that if each factor is rewarded equal to their marginal product, the total product must be disposed of without any surplus or deficits. The problem of providing that the total product will be just exhausted if all factors are paid rewards according to their marginal productivity has been called the "adding up" problem. This is also known as the product exhaustion problem.

The adding up problem states that in a competitive factor market when every factor employed in production process is paid according to its marginal product, then payments to the factors exhaust the total product. It can be shown numerically as:

$$Q = (MP_L)L + (MP_K)K$$

Where Q= Total Output

MP= Marginal Product

K= Capital

L= Labour

Since marginal productivity can be expressed in terms of partial derivatives. So it can be written as:

$$Q = \frac{\partial Q}{\partial L}.L + \frac{\partial Q}{\partial K}.K$$

Where, $\frac{\partial Q}{\partial L}.L$ represents share of labour in the total product and $\frac{\partial Q}{\partial K}.K$ represents share of capital in the total product.

Philip Wickstead was one of the first economists who posed the adding up problem and provided a solution for it. Wickstead applied a mathematical formula (proposition) called Euler's theorem to prove the adding up theorem. Wickstead in his book, "The Coordination of the Laws of Distribution" demonstrated with the help of Euler's theorem that total product will be exhausted if every factor is paid according to its marginal product. Since Wickstead applied the Euler's theorem to prove the adding up theorem it is also called as Euler's theorem.

Euler's Theorem and the Wickstead's Solution to the Product Exhaustion Problem:

Wickstead pointed out that if production function is homogeneous of degree one, then Euler's theorem can easily demonstrate that payment in accordance with marginal productivity would exhaust the total product. This can be explained as below:

Let K and L be the quantities of two factors of production, capital and labour respectively. Q = Q (K, L) be the production function. Now if the production function is homogeneous of degree one, then according to Euler's theorem:

$$Q = \frac{\partial Q}{\partial L}.L + \frac{\partial Q}{\partial K}.K \longrightarrow \qquad (1)$$

Where, Q is the total product; $\frac{\partial Q}{\partial L}$ is the marginal productivity of labour; $\frac{\partial Q}{\partial K}$ is the marginal productivity of capital; $\frac{\partial Q}{\partial L}.L$ represents share of labour in the total product and $\frac{\partial Q}{\partial K}.K$ represents share of capital in the total product.

Now to prove the Euler's theorem, let us take the Cobb-Douglus production function:

Q = Q (K, L) = $AK^{\alpha}L^{\beta}$; $\alpha + \beta = 1$; and A, α, β are positive parameters.

Now partially differentiating with with respect to K and L we get,

$$\frac{\partial Q}{\partial K} = \frac{\partial}{\partial K}(AK^{\alpha}L^{\beta})$$
$$= A\alpha K^{\alpha-1}L^{\beta}$$
$$= \frac{\alpha AK^{\alpha}L^{\beta}}{K}$$
$$= \frac{\alpha Q}{K} \longrightarrow \qquad (2)$$

$$\frac{\partial Q}{\partial L} = \frac{\partial}{\partial L}(AK^{\alpha}L^{\beta})$$
$$= A\alpha K^{\alpha}\beta L^{\beta-1}$$
$$= \frac{\beta AK^{\alpha}L^{\beta}}{L}$$
$$= \frac{\beta Q}{L} \longrightarrow \qquad (3)$$

Now substituting the value of $\frac{\partial Q}{\partial K}$ and $\frac{\partial Q}{\partial L}$ in the right hand side of the equation (1), we get,

$$\begin{aligned}\frac{\partial Q}{\partial K}.K + \frac{\partial Q}{\partial L}.L &= K.\frac{\alpha Q}{K} + L\frac{\beta Q}{L} \\ &= \alpha Q + \beta Q \\ &= Q(\alpha + \beta) \\ &= Q.1 \qquad (\alpha + \beta = 1) \\ &= Q \\ \therefore Q &= \frac{\partial Q}{\partial L}.L + \frac{\partial Q}{\partial K}.K\end{aligned}$$

Hence the Euler's theorem is proved and we may conclude that the marginal product of capital multiplied by the unit of capital employed plus the marginal product of labour multiplied by the unit of labour employed, exactly equals the total product, Q and hence the total factor payments exhaust the total value of the product.

Assumptions:

Euler's theorem or the adding-up theorem is based on the following assumptions:

(I) It assumes a linearly homogeneous production function of first order, which implies constant returns to scale.

(II) It assumes that factors are complementary.

(III) It assumes that factors of production are perfectly divisible.

(IV) The relative shares of the factors of the factors of production are constant and independent of the level of output.

(V) There is stationary, risk less economy where there are no profits

(VI) There is perfect competition and it is applicable only in the long run.

Criticisms:

Wickstead's exposition of the marginal productivity theory was the target of trenchant criticism of Walras, Barene, Pareto and Edgeworth. The central point of criticism was the form of production function he assumed, i.e., the production function with constant returns to scale. It is pointed out that the assumption of constant returns to scale is incompatible with another assumption of the theory- the assumption of perfect competition. If there are increasing or decreasing or constant returns to scale, firms under perfect competition will have a tendency to expand to a point where diminishing returns to scale take place or where the firms in the

industry are no more in perfect competition. The typical production function, the critics maintained, is one showing varying returns to scale.

Another point on which Wickstead was later criticized that he treated laws of increasing returns to scale, constant returns to scale and diminishing returns to scale as mutually exclusive alternatives. Modern economists have made it clear that the returns to scale represent the different stages of the long run cost curve of a firm.

Importance of the theory :

The product exhaustion theory is of great importance in the theory of distribution on the following grounds:

(I) It tells us that if market results were fully predictable, rents, wages and interest would exhaust the entire net output of the economy.

(II) It also suggests how a firm should employ the various inputs. It tells us that firm should employ its inputs to that extent at which the reward to the factor equals its marginal revenue product.

CHECK YOUR PROGRESS- 1.1

1. State marginal productivity theory of distribution.

.....
.....
.....
.....
.....
.....

2. State Euler's theorem.

.....
.....
.....
.....
.....

3. What is adding-up controversy?

.....
.....
.....
.....
.....

1.4 RENT:

In popular language, the word rent is used to denote payments for the use of land, a house or a shop, etc. But in economics, the term rent is used not only in the sense of rewards for the use of land but also in the sense of earnings of the factors over their transfer earnings. The payment for the use of land is regarded as land rent and surplus over transfer earning as economic rent.

1.4.1 QUASI RENT AND TRANSFER EARNINGS:

The concept of quasi rent was first given by Alfred Marshal. Marshall defined quasi rent as the surplus earned by instruments of production other than land. *The term rent is applied to the income earned from the land and the other free gift of the nature. On the other hand the term quasi rent is applied to the income earned from the appliances and machines resulting from the effort of the mankind.* Quasi rent stands for whole of the income, which some agents of production yield when demand for them has suddenly increased. It is earned during the period that their supply cannot be increased in response to increase in demand for them. For example, in the short run supply of some durable goods such as machines, ships, houses, etc., cannot be increased. But in the long run supply of all these can be increased easily. So, due to scarcity of the products they earn a surplus that is like rent, but not rent as their supply can be increased in the long run. In Marshall's words, quasi rent refers to "income derived from machines and other appliances made by man." It can also be defined as the excess of total revenues earned in the short run over and above the total variable costs.

The concept of transfer earnings can also be applied to explain quasi rent. It is the amount that any factor of production could expect to earn in its best alternative use. In the short run, specialized machinery must remain in its present use; it cannot be transferred to other use. This means that their transfer earnings are zero. Hence the whole of earnings of machinery and capital equipments in the short run over transfer earnings is the rent. But this is not rent, as it is temporary and hence called as quasi rent. The supply of scarce factors cannot be increased in the short run whatsoever is the demand may be, as they are fixed in supply. But their supply can be increased in the long run and so the surplus earnings occur as a result of scarcity in supply will be disappeared. But it cannot be happened in case of land as its supply is perfectly inelastic and hence the rent of land will persist in the long run.

1.4.2 RICARDIAN THEORY OF RENT:

David Ricardo, a famous 19th century English economist developed a systematic theory of rent. He defined rent as *"that portion of the produced of the earth which is paid to the landlord for the use of the original and indestructible powers of the soil."* In fact he wanted to distinctly separate the payment for powers of the soil from the payment made for the improvement of the soil. Therefore according to his definition, land rent is a payment for the use of only land and is different from contractual rent which includes return on capital investment made by the landlord in the form of hedges, drains, well and the like.

Assumptions:

Ricardian theory of rent is based on the following assumptions:

1. Rent is peculiar to land alone. Rent arises because of the peculiar characteristics of land, namely, that its supply is inelastic and it differs in fertility. Rent arises because of the differences in the fertility of land. Rent is a differential surplus. All lands are not equally fertile. Only these lands, which are more fertile than others, will get rant.
2. Land has more original and indestructible powers.
3. Land is subject to the law of diminishing returns.
4. There is perfect competition.

Explanation:

Under the above assumptions, Ricardo explained his theory. He considered rent as a surplus, which arises due to the differences in fertility and situation. For explaining the origin of rent, he took farming technique, viz., intensive cultivation and extensive cultivation; and the assumption of marginal land, which is also termed as no rent land is a piece of land which is just covering the expenses with the produced output. In other words, marginal land is that grade of land after which no grade of land is used. The land, which has higher productivity than this marginal land, is called "intra marginal land." Hence all the plots of land which produce more than the marginal land even rent. So, Ricardo believed rent to be differential surplus covered by intra- marginal lands over the earnings of marginal lands.

Ricardo explained the process of formation of rent under both techniques with the aid of an example of cultivation. Imagining that a new island is discovered and there are three groups of land, A being the superior most and C the poorest, B grade of land lies between A and C.

When people first come to island they will take up the best category of land A for the production of say, rice. So as long as the first grade of land is available for cultivation, there will be no rent. Now suppose there is an increase in population and this leads to an increase in the demand as well as price for rice and make it necessary to bring the second category of land in to cultivation. Now rent on the best quality of land will arise. In this way, each increase in population necessitates recourse to land of progressively inferior quality, grade c and so on. The last cultivated land, i.e, C grade land does not yield any rent and the other grades of land A and B earns rent and above the produce of this C grade land.

Rent under Extensive Cultivation:

Extensive cultivation is the type of farming under which production is increased by using more of land. To illustrate the emergence of rent under extensive cultivation as discussed above, let us suppose that same dose of capital and labour to produce 25 quintals of rice on grade A land, 20 on B and 15 on the C grade land. So long as A grade land is cultivated, no rent arises. But when B grade land is brought under cultivation, land A will get a surplus, i.e., rent which is equal to their difference (25-20=5 quintals). Land A is called intra-marginal land. Again, when grade C land is brought under cultivation, there emerges a rent of 5 quintals on land B and 10 quintals on A while C becomes the no rent land. This fact is shown in Table-1:

Table: 1

Grade of Land	Production (qty)	Surplus, i.e., rent
A	25	25-15=10
B	20	20-15=5
C	15	15-15=0

From the table, we see that grade A and B of land earn rent, which are regarded as intra-marginal land; while grade C land earns no rent and it is called no rent or marginal land. This can be explained with the help of the diagram-2

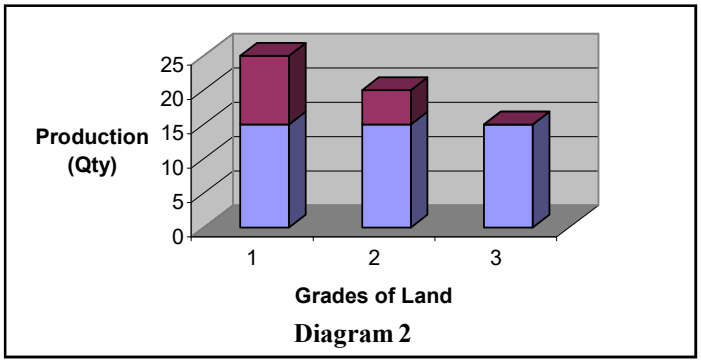


Diagram 2

In the above Diagram 2, the amount of rent earned by A and the

differently shaded areas show B grades of land. C grade land just covers its cost and hence produces no rent.

Rent under intensive cultivation:

Intensive cultivation is the type of farming where on the same piece of land, more units of labour and capital are used to increase production. When more and more units of labour and capital are put to work on the same plot of land, product will increase but at a diminishing rate. Hence, marginal product will go on diminishing. Let us suppose that three such units i.e., 1st, 2nd and 3rd are put to work and they produce 25, 20 and 15 quintals of marginal product respectively and rent in each case will be respectively 10, 5 and 0. This is shown in the following table-2:

Table-2

Units of Labour-Capital	Production (Qty)	Surplus, i.e., rent
1st	25	25-15=10
2nd	20	20-15=5
3rd	15	15-15=0

The above table shows that 1st and 2nd units earn rent, not due to the differences of the fertility of the land, but due to the diminishing returns on the same piece of land. This is shown in the following figure-2:

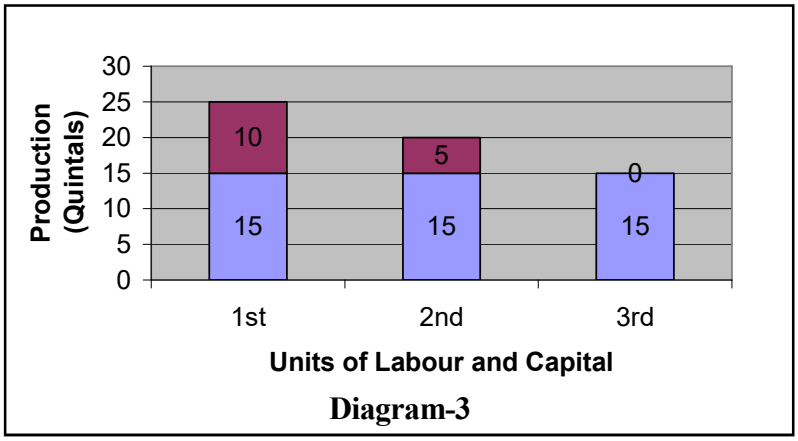


Diagram-3

In the above Diagram-3, units 1st and 2nd earn rents, while 3rd units earn no rents. In this way Ricardo shows that rent is a differential surplus that some plots of land earn over and above the least fertile land under cultivation. Thus the Ricardian theory of rent made a clear conclusion that rent of land is an unearned surplus.

Criticisms:

Ricardian theory of rent has been the subject of criticisms ever since it was propounded. It has been criticized both for its

assumptions as well as conclusions. The main points of criticisms are as under:

1. Ricardo assumed that land has some "original and indestructible powers of the soil ". But critics have argued that there are no such original powers of the soil and its powers are not indestructible. For the fertility of land may decrease in the course of time by continuous cultivation.
2. Again Ricardo assumed that best lands are cultivated first. But, there is no historical proof for this. Best lands are not always cultivated first.
3. Ricardo assumes the existence of marginal or no rent land. But it is difficult to find such land in practice.
4. Ricardo explained the origin of rent as a differential surplus of superior lands over inferior lands. Thus, Ricardo ignored the origin of rent due to scarcity of land.
5. Rent is not peculiar to land alone. Modern economists feel that the rent aspect can be seen in other factor incomes as well. According to the modern view the term rent is applied to "payments made for factors of production which are in imperfectly elastic supply." Thus the term rent includes besides payments for the use of land, other payments for labour as well as capital.
6. Ricardian theory of rent is based on the assumption of perfect competition. But in real world, perfect competition is a myth.

Conclusion:

Notwithstanding these criticisms of the theory, Ricardo's insights in to the nature and origin of rent stand as a singular contribution to the attempt at understanding of the determination of factor shares and the explanation of differences between them. Much of what has been done in the theory of rent is by way of the extension and the modification of his ideas. Rather his idea of the differential surplus has been used in explanation of other factor shares also, say in profits.

CHECK YOUR PROGRESS-2

1. Define rent.

.....

.....

.....

.....

2. What is quasi rent?

.....

.....

.....

.....

3. State Ricardian theory of rent.

.....

.....

.....

.....

1.5 WAGES:

Wages can be defined as the payment made to the workers for their participation in the productive activities. Wages are paid to the workers either for physical labour or for mental labour. Wages are paid usually on daily, weekly and on monthly basis. The term wages in economics may refer to time wages, money wages, and real wages etc. But in economics, the term money wages is most often used. It refers to the payment made to the workers for their mental and physical services considered either in per hour or per day or per week or per month. On the real wages refer to the net worth in terms goods and services of the worker's money remuneration, i.e., the amount of necessaries, comforts and luxuries of life that the worker can command in return for his services.

1.5.1 CLASSICAL THEORY OF WAGES:

J.S. Mill developed the classical theory of wages, also known as the wages fund theory. The theory states "wages depend upon the proportion between population and capital" or rather between "the number of the labouring classes who work for hire and the aggregate of what may be called the wages fund, which consists of that part of circulating capital, which is expended on the direct hire of labour." According to Mill every employer keeps a certain amount of money for the payment of wage that is known as the wages fund. It is called fund, as it is fixed and constant. Since the wages fund is constant, wages will depend directly on the size of the labour force. If the wage fund is unchanged and the size of labour force increases due to growing population, wages will fall. The opposite will happen if there is a contraction in the labour force due fall in the growth rate of population. According to this theory, wages depend on two quantities: (i) the wages fund keep by the employer for the payment

of wages; and (ii) the number of labourers seeking employment. Thus it follows from the theory that a rise in wages is possible only under two conditions. It can be increased either by increasing wages fund by an increase in savings or there must be a fall in the supply of labour. The advocates of the theory believe that increase in wages is possible at the expenses of the employer. Thus the only way to increase wages is to reduce the number of workers.

1.5.2 MODERN THEORY OF WAGES

According to the modern theory of wages the wage rate in a perfectly competitive market is determined by the demand for and the supply of labour. This is same as the general theory of value.

Demand for Labour:

Demand for labour is a derived demand. Labour is demanded for its service by the employers in helping them to produce goods and services. The greater the consumer demands for the product, the greater the producer's demand for the labour required in making it. Thus their demands rise or fall according to the rise or fall in the demand for the product produced by them. In fact, it is not the demand for labour that matters but the elasticity of demand for labour, which depends on elasticity of demand for its product. The more elastic is the demand for labour; more is the demand for labour. Demand for labour also depends on the prices and quantities of the co-operating factors. Suppose if the machines are costly as in the case of India, more labour will be demanded and hence, the demand for labour will be high.

Another factor that influences the demand for labour is the technological progress. Thus the demand for labour is determined by: (a) the nature of the demand for the product; (b) the proportion that the cost of labour bears to the total cost of the product; (c) the substitutability by other factors; and (d) the supply of capital as determined by the willingness to save and invest by the producers.

After considering all the relevant factors, e.g., demand for the products, technical conditions, etc., the producer is guided by another factor, called as the marginal productivity of labour. Marginal revenue productivity of labour is the addition made to total revenue as a result of the addition of labour by one unit. The wage rate of labour at any time is equal to the marginal productivity. So long as the productivity of labour as determined by the marginal revenue productivity of labour is greater than the wage rate, it is profitable to employ more labour for the employer as it adds more revenue than to costs. But after a point the marginal productivity of

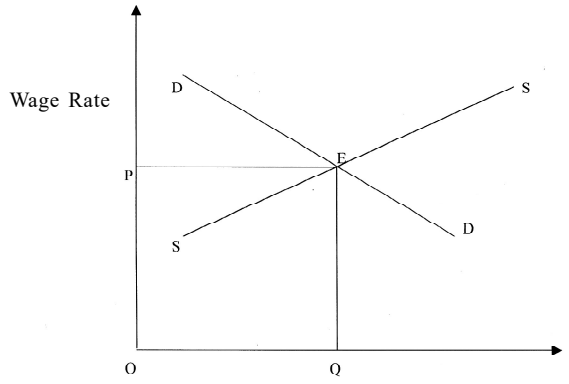
labour starts falling due to law of variable proportion. The employer will stop employing additional labourer at the point where cost of employment is just equal to the return made by him.

Supply of Labour:

Supply of labour means the number of workers that would offer themselves for employment at each wage rate. The relation between wage and the supply of labour is a direct one. Greater the wage rate more the supply of labour and vice versa. The supply of labour depends on a number of factors like the rate of growth of population, the working hours, the normal period of education and training, labour laws, attitude towards work, the mobility of labour, attitude of the worker towards work and labour etc. The supply curve of labour for a firm is perfectly elastic and for the industry as a whole is not infinitely elastic.

Determination of Wage Rate:

In a perfectly competitive market, the equilibrium wage rate is determined by the demand for labour and the supply of the labour. This can be explained with the help of the diagram 4.



Employment of Labour

Diagram : 4

In the Diagram 4, the DD curve represents the demand for labour and the SS curve represents the supply of labour. Both the demand curve and the supply curve intersect at the point E. This is the point of equilibrium where wage rate OP and the supply of labour OQ is determined. Thus in a perfectly competitive form of market wage rate is determined by the demand for and the supply of labour. However, in real world it is grossly absent.

1.5.3 ROBINSON'S THEORY OF EXPLOITATION:

Joan Robinson, an American economist developed the concept of exploitation of labor independently besides Edward

Chamberlin. The concept of exploitation of labor as developed by Robinson is different from that of Chamberlin. Following A.C Pigou, Robinson defines 'exploitation' as a wage less than the marginal physical product of labour valued at its selling price. In other words, according to Robinson, labour is exploited when it is paid fewer wages than the Value of its Marginal Product (VMP). By value of marginal product we meant marginal physical product multiplied by the selling price of the product .To quote Robinson "what is actually meant by exploitation, is usually, that the wage is less than the marginal physical product of labor valued at its selling price."

The exploitation of labour occurs when there is imperfect competition (or monopsony) in the buying of labour (i.e., labour market) as well as when there is imperfect or monopolistic competition in the product market. Thus, exploitation of labour can't occur when there is perfect competition in both the labour and product market. According to Robinson, under imperfect competition in the labour market, labour invariably gets less than the value of the marginal product, since it is paid according to marginal product multiplied by marginal revenue i.e., marginal revenue product which is less than the marginal product multiplied by price. Before the development of imperfect and monopolistic competition theories by Robinson and Chamberlin it was believed that when there is imperfect competition in the labour market, labour would be exploited because in that case wages would be less than the value of marginal product, (VMP). With the development of the theories of imperfect and monopolistic competition, it became clear that even if perfect competition prevailed in the labour market, labour could be exploited on account of imperfections in the product market.

Robinson's concept of exploitation of labour can be explained with the help of following Diagrams 5, 6 and 7.

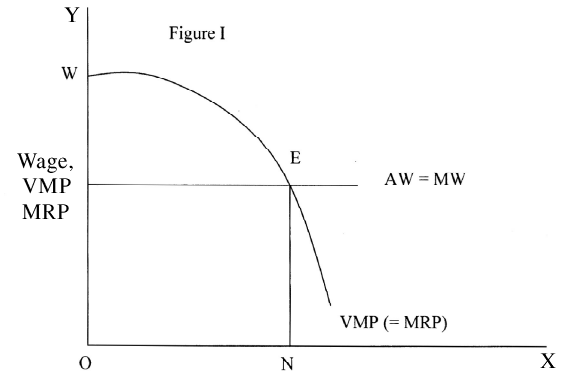


Diagram:5

In Diagram 5, we draw a situation when there is perfect competition in both the labour as well as in product market. Since, perfect competition prevails in the market for the product produced by the firm, price and marginal revenue would be the same and therefore value of the marginal product will be equal to the marginal revenue product (M R P) of the labour. Further, since there exists perfect competition in the labour market, the firm has no control over the wages and hence the average wage curve is perfectly elastic and coincides with marginal revenue curve. In the above diagram, we have seen that firms would be in equilibrium at ON employment of labour where they are equating the wage rate with the value of marginal product, which is equal to marginal revenue product (MRP). Thus, according to Robinson view no exploitation exists when perfect competition prevails in the labour market as well as in product market.

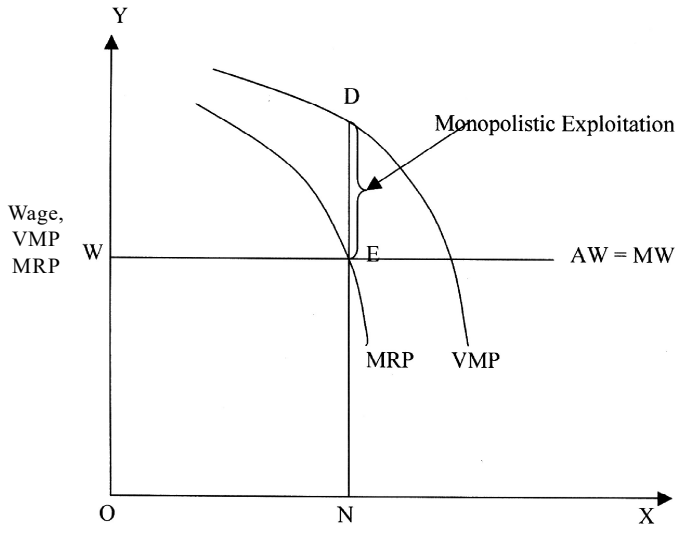


Diagram: 6

In the Diagram 6, there prevails imperfect competition in product market, but there is perfect competition in labour market. Robinson views that on this case exploitation prevails. Since there is perfect competition in labour the labour market, average and marginal wage curves coincide and are perfectly elastic at the current wage rate. On account of imperfect competition in the product market, marginal revenue is less than the value of the marginal product and hence the curves of these two diverse from each other; MRP curves lies below VMP curve. To be in equilibrium in this situation, a firm will equate wage with MRP. In the above diagram, we have seen that ON amount of labour is employed at OW

equilibrium level of wage. But, it is noticed in the diagram that value of the marginal product is ND while the labour is being paid equal to marginal revenue product NE (NE=OW) which is less than the VMP, ND. Thus, labour being paid less than the value of the marginal product by the amount ED. This is termed as exploitation of labour. Since this difference of ED between the wage paid to labour and the value of marginal product of labour has arisen on account of imperfect competition in product market, it has been termed as monopolistic exploitation by Joan Robinson.

The divergence between the wage rate and the value of marginal product (VMP) of labour can also arise on account of imperfect competition in labour market. When there is imperfect competition or monopsony in the labour market, supply curve of labour is not perfectly elastic but is upward sloping. As a result, marginal wage (MW) curve lies above the average wage (AW). How the exploitation of labour under this condition takes place can be explained with the help of the Diagram 7:

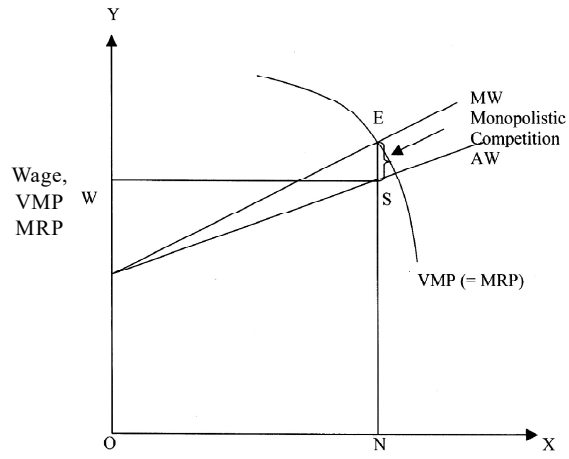


Diagram 7

In the above diagram 7, we have seen that there prevails perfect competition in the product market and VMP and MRP are same. But in the labour market, there exists imperfect competition. It is from the Diagram that marginal wage is above average wage and is upward sloping. The equilibrium level of wage NE is determined at point E when ON amount of labour is employed. Here the value of marginal product (VMP) is NE. On the other hand, labour is paid to the extent NS=(OW) and hence ES amount of labour is exploited. Since, exploitation ES has arisen on account of monopsony or imperfect competition in the labour market, it has

been termed as monopsonistic exploitation by Robinson. Monopsonistic exploitation arises because the supply curve of labour is not perfectly elastic and therefore marginal wage curve lies above the average wage curve.

Thus, when there is imperfect competition in both labour and product markets, labour would be subjected to double exploitation, monopolistic as well as monopsonistic.

Chamberlin's Critique of Robinson's Concept of Equilibrium:

Chamberlin has criticized Robinson's concept of exploitation. According to him, under imperfect competition in the product market, not only labour but also all factors receive less than the value of their marginal products. He said that under the situation of imperfect or monopolistic competition in the product market, all factors including entrepreneur get rewards according to the marginal revenue product (MRP), which is less than the value of marginal product (VMP). Thus he rejects the Robinson's view of exploitation, which suggests that labour is exploited when he gets wage less than its value of marginal products (VMP). On the other hand according to Chamberlin, labour is exploited only when he is paid less wage than its marginal revenue products (MRP).

How can labour exploitation be removed?

Monopolistic exploitation, which has arisen due to imperfect competition in the product market is concerned; it can't be removed by raising wages by trade union. This is because, in this situation if the trade union is succeed in raising wages, the employer will employ small amount of labour so as to compensate the raise in wages. But, the important point to note is that with lower employment and higher wage rate, labour would still be exploited for in this new wage position also, value of marginal products (VMP) would be greater than the marginal revenue product (MRP) with which new higher wage will be equated by the employer. Thus, the monopolistic exploitation as conceived by Robinson can't be removed by raising wages by trade union or government. Monopolistic exploitation can only be removed by creating the conditions of perfect competition in product market. State can take measures for removing monopolistic conditions or imperfections from the product market.

But so far as the monopolistic exploitation of labour is concerned, it can be removed by raising wages through trade unions or state.

CHECK YOUR PROGRESS -3

1. What is meant by wage?

.....

.....

.....
.....

2. What is real wage and money wage?

.....
.....
.....
.....

3. In which form market exploitation of labour takes place?

.....
.....
.....
.....

4. Who has developed the theory of monopolistic exploitation?

.....
.....
.....
.....

1.6 INTEREST:

Interest is a payment made by a borrower to the lender for the use of a sum of money for a period of time. It is one of the four types of income, besides rent, wages and profit. In economics, interest is regarded as the price paid for the use of money capital in production for a period of time. According to Saligman, *"interest is the return from the fund of capital."* To Wickshell interest is *"a payment made by the borrower of capital by virtue of its productivity, as a reward for his (capitalists) abstinence."* On the other hand treating money as a purely monetary phenomenon Keynes defined it as *"the premium, which has to be offered to induce people to hold their wealth in some form other than hoarded money."* Thus from all the definitions it is clear that the concept of interest is associated with money, being the price paid for its use.

There are a large number of theories to explain the nature, origin and determination of the rate of interest. In the following section we discuss three theories of interest, namely,

1. Classical Theory of Interest
2. Keynesian or Liquidity Preference Theory and
3. Modern Theory or the Neo-Keynesian Theory of Interest

.

1.6.1 CLASSICAL THEORY OF INTEREST:

The Classical theory of the interest seeks to explain the determination of interest rate by real factors like productivity and thrift. It treats interest rate as an equilibrating factor between the demand for and the supply of investible fund consisting of savings. This theory is also called as the real theory of interest as it explains the determination of the rate interest by real forces such as thrift, time preference and productivity of capital. Various classical writers differ a good deal from each other in respect of their views about interest. Some of them laid emphasis on the forces governing the supply of savings. Thus they considered interest as a price paid for abstinence or waiting or time preference. Some others like J. B. Clark and Knight thought that the marginal productivity of capital, which is a force that operates on the demand side of savings, determines the rate of interest. But, they agreed with the view that rate of interest is a payment for saving. The rate of interest, in this theory, is determined by the demand for savings to invest in capital goods and the supply of savings. Let us explain these demand and supply sides.

Demand for Savings :

The demand for capital goods comes from firms, which desire to invest, that is, to purchase or make new capital goods. Capital goods are demanded because they can be used to produce consumer goods - because they have revenue productivity like all other factors. For any given type of capital asset, e.g. a machine, it is possible to draw a marginal revenue productivity curve showing the addition made to total revenue by an additional unit of a machine at various levels of the stock of that machine.

We have that capital, like other factors of production, has marginal revenue productivity of capital is a more complex than that of other factors, because capital has a life of many years. A capital asset continues to yield returns for many years. Therefore, the entrepreneurs have to take into account the uncertainties of the future and estimate the percentage yield or returns from capital after making allowance for the maintenance and operating costs. In other words, they have to find out the net expected return of the marginal unit of capital expressed as percentage of the cost of the capital asset. The more capital assets of a given kind an entrepreneur has, the less revenue or income he will expect to earn by purchasing one more machine of the same kind. Therefore marginal revenue productivity of capital slopes downwards towards the right.

Now, under perfect competition, it is profitable for a firm to

purchase any factor up to the point at which the price of that factor equals its marginal revenue productivity. The price of savings required to purchase the capital goods is obviously the rate of interest. Hence, the entrepreneur will demand capital goods (or will demand savings) to purchase capital goods up to the point at which expected net rate of return on the capital goods equals the rate of interest. Since the marginal revenue of the productivity curve of capital good is downward sloping, it follows that as the rate of interest falls, more capital goods will be demanded and also more money will be required to purchase these capital goods.

The way in which the demand for capital goods rises as a result of fall in rate of interest is shown with the help of the Diagram 8 where MRP is the marginal revenue productivity curve. On the Y-axis, net rate of return on capital and the rate of interest are shown, while X-axis represent the amount of capital.

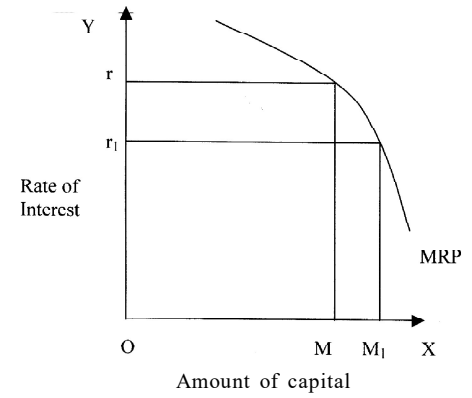


Diagram: 8

At Or rate of interest OM amount of capital is demanded. This is so because only at OM amount capital, the falling net rate of return on capital becomes equal to the prevailing rate of interest Or. Now if the rate of interest falls from Or to Or_1, then the amount of capital demanded will increase from OM to OM_1, since at OM the falling net rate of return equals the new interest rate Or_1. Thus, it is clear that the marginal revenue productivity curve of capital shows the demand for capital and further that the demand for capital slopes downwards towards the right. This is true for individual industries and for the community as a whole.

Thus, we conclude that demand for individual capital goods will increase as the rate of interest falls.

Supply of Savings:

According to this theory, those who save from their current income make the money, which is to be used for purchasing capital

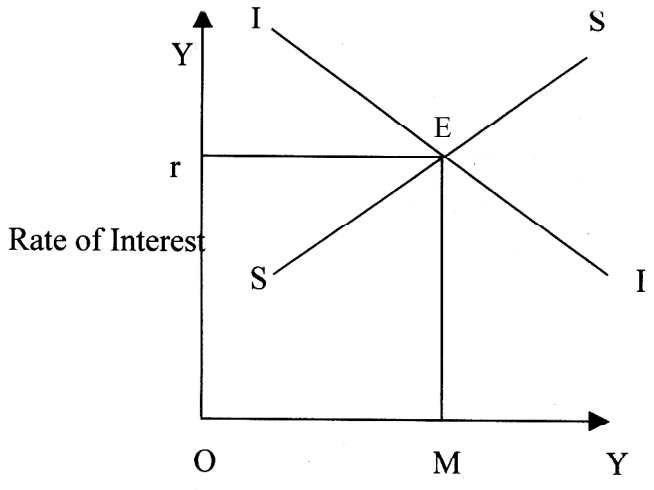
goods, available. By postponing consumption of a part of their current income, they release resources for productive purposes. Savings involve the element of waiting for the future enjoyment of savings. But the people prefer the present enjoyment of goods and services to the future enjoyment of them. Therefore, if people are to be persuaded to save money and to lend it to entrepreneurs, they must be offered some interest as reward. In other words, to make people overcome their time preference, inducement must be offered in the shape of interest. More savings the people will do, the more consumption they will have to postpone, the higher must be the rate of interest they will ask to make such a postponement worthwhile. Thus, to induce people to save more, higher rate interest must be offered.

Moreover, higher rates interest have also to be paid if savings have to come from those persons whose rates time preference are relatively more strongly weighed in favour of present satisfactions. The supply curve of capital will, therefore, slope upwards to the right.

Equilibrium between Demand and Supply:

The rate of interest is determined by the interacting of the forces of demand for capital and the supply of savings. The rate of interest at, which the demand for capital and the supply of savings are in equilibrium, will be determined in the market.

Now the interaction of demand for investment and supply savings determine the rate of interest can be explained with the help of the following Diagram 9:



Value of Saving & Investment
Demand & Supply of equilibrium

Diagram: 9

In Diagram 9, SS is the supply curve of savings and II is the demand curve of savings to invest in capital goods. The demand for investment and supply of savings are in equilibrium at Or rate of interest, at which the curves intersect each other. Hence Or is the equilibrium rate of interest. In this equilibrium position, OM amount of capital is lent, borrowed and invested. If any change in the demand for investment and supply of savings comes about, the curves will shift accordingly, and therefore, the equilibrium rate of interest will also change.

1.6.2 KEYNESIAN OR LIQUIDITY PREFERENCE THEORY:

Criticizing the Classical theory of interest, Keynes developed his famous liquidity preference theory of interest in 1936. According to Keynes, the rate of interest is purely a monetary phenomenon and is the reward for parting with liquidity. Keynes explained the level of interest rates in an economy in terms of the demand for money and supply of money. To Keynes supply of money is completely inelastic because it is fully determined by the Central Bank of the country and cannot be controlled by the individuals or firms. Hence, the supply of money does not influence the rate of interest in an economy.

Thus, the demand for money, according to Keynes, is the main determinant of interest rate. Keynes pointed out that there are three motives why people, firm and governments want to hold money despite the loss of interest keeping it in other forms of money. These motives are respectively transaction motive, precautionary motive and speculative motive.

Transaction motive:

The amount of money keeping for the purpose of day-to-day transaction by the individuals, business people and industrialists is known as the transaction motive. In case of individual, the amount of money held for transaction purposes depends on his/her level of income. As income of the people increases, the size of money held for transaction motive also increases. On the other hand, for the businessman and industrialists the amount of money for transaction motive depends on the size of the business.

Precautionary motive:

The common people as well as the businessman and the industrialists keep some amount of money to meet some unforeseen contingencies like sickness, accidents, etc. this motive was not included in the classical theory of interest rate. Keynes grouped transactions and precautionary balances together into a single sum, which varied directly with the level of income and the general price level.

Speculative motive:

Keynes first of all gave this motive. Speculative demand for money refers to the demand for holding certain amount of cash in reserve to make speculative gains out of the purchase and sale of bonds and securities through future changes in the rate of interest. Demand for speculative motive is related with the rate of interest and bond prices. However, there is an inverse relationship between the rate of interest and bond prices. For example, a bond with price of Rs. 100 yields a fixed amount of Rs. 3 at rate of interest 3%. If the rate interest increases to 4% the price of bond must fall to Rs. 75 to yield the same fixed income of Rs. 3.

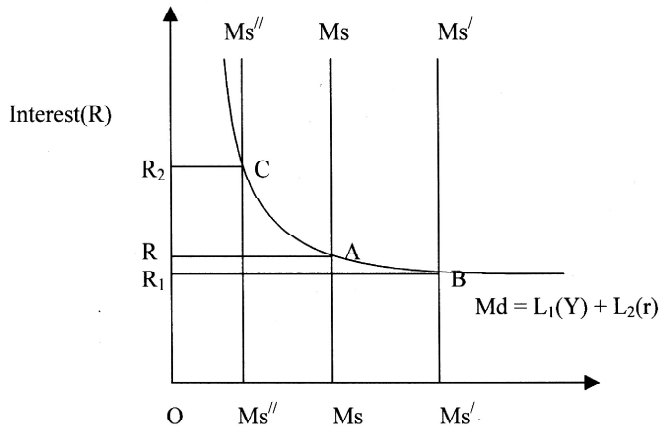
People desire to have money in order to take advantage from knowing better than others about the future changes in the rate of interest (bond prices). If people feel that the current rate of interest is low (bond prices are high) and it is expected to rise in future (or bond prices will fall in future), then they will borrow money at a lower rate of interest (or sell their already purchased bonds), keep cash in hand with a view to lend it in future at a higher interest rate (or to purchase bonds at cheaper prices in near future). Thus the demand for money will rise for speculative motive and vice versa.

Supply of Money:

Supply of money refers to the total quantity of money in the country for all purposes at a particular point of time. It is being fixed by the monetary authority of the country and is therefore independent of the rate of interest. The money supply curve is vertical in shape.

Determination of the rate of interest:

As we have mentioned that the rate of interest in the Keynesian theory of interest is determined by the demand for and supply of money (Ms). The total demand for money (Md) is the composite of the transaction demand for money, precautionary demand for money



Supply of and demand for money

Diagram: 10

(M_1) and the speculative demand for money (M_2). The equilibrium level of rate of interest is determined at that level where the total demand for money is equal to the total supply of money. This can be explained with the help of Diagram 10.

In Diagram 10, the rate of interest is measured on the vertical axis and supply and demand for money on the horizontal axis. The demand for money curve shows the amount of money, which is demanded at different interest rate. The rate of interest has been determined at the level OR where the two schedules: $Md = L_1(Y) + L_2(r)$ and MsMs intersects. When the supply of money increases from OMs to $OMs^{/}$, position of the liquidity preference schedule remaining unchanged, the rate of interest falls from OR to OR1. Conversely, when the money supply decreases from OMs to $OMs^{//}$, liquidity preference remaining unchanged, the rate of interest rises from OR to OR_2.

Liquidity Trap:

Liquidity trap is a situation in which real interest rate cannot be reduced by any action of the monetary authority. This is liable to arise if prices are expected to fall. If general price falls are expected, holding money will produce an expected real gain equal to minus the expected rate of inflation. The real interest rate cannot be lowered below this point at which the nominal interest rate falls below zero. Thus the monetary authority fails to promote investment by cutting real interest rates, even if investment would be responsive to a real interest rate cut if one should occur.

1.6.3 MODERN THEORY OF INTEREST:

The modern theory of interest or the neo-Keynesian theory of interest rate was developed by Hicks, Hansen and Lerner. In this theory, they have incorporated both the real variables and the monetary variables to determine the equilibrium rate of interest rate.

According to the modern theory of interest, a complete and determinate theory of interest must take in to account both the real variables as well as monetary variables influencing the interest rate. Thus to the modern theory there are four important determinants of the rate of interest: (a) the saving function, (b) the investment function, (c) the liquidity preference function and (d) the quantity of money. The equilibrium between these four variables together determines the rate of interest and the equilibrium level of income. According to Hensen, "An equilibrium condition is reached, when the desired volume of cash balances equals the quantity of money, when the marginal efficiency of capital is equal to the rate of interest and finally, when the volume of investment is equal to the normal

or desired volume of saving. And these factors are integrated." Thus according to the modern theory of interest, the equilibrium rate of interest rate and the level of income are determined at the point of intersection of IS curve and LM curve.

Derivation of IS curve:

The IS curve shows equilibrium in the product market represents those combinations of income and the rate of interest rate at which the aggregate savings S (+T) and aggregate investment I (+G) are equal. For derivation we make the following assumptions:

(I) Consumption is a function of disposable income, i.e., C=C (Yd), where Yd= Y-T.

(II) Investment is a function of rate of interest, i.e., I=I (i).

(III) Government purchases are exogenous, i.e., $G=G_0$.

(IV) Taxes are function of income, i.e., T=T (Y).

(V) For product market to be in equilibrium, investment plus Govt. expenditure must be equal to the saving plus taxes, i.e., I+G= S+T.

(VI) We further assume that investment is independent of output.

With these assumptions we can derive the IS curve that represents different equilibrium level of income at different interest rate. In the Diagram 11 drawn below, we have drawn I+G curve in the northwest quadrant. The I+G curve is obtained by summing investment and the government purchases at each rate of interest. The S+T curve is plotted in the southwest quadrant of the figure that shows that as income increases saving plus taxes also increases. In the southeast quadrant we have drawn a 45^0 line to translate income from vertical axis to the horizontal axis.

To derive IS curve in the northeast quadrant, we start with interest rate i_0 and investment plus government purchases equal to I_0+G_0. The equilibrium condition requires that I+G must be equal to S+T. This condition is being satisfied at Y_0 level of income. at any other income level, saving plus taxes is either greater than investment plus government purchases or less than investment plus government purchases. Consequently, for interest rate i_0, the equilibrium level of income is Y_0. This point, i.e., $A(Y_0,i_0)$ is plotted in the northeast quadrant.

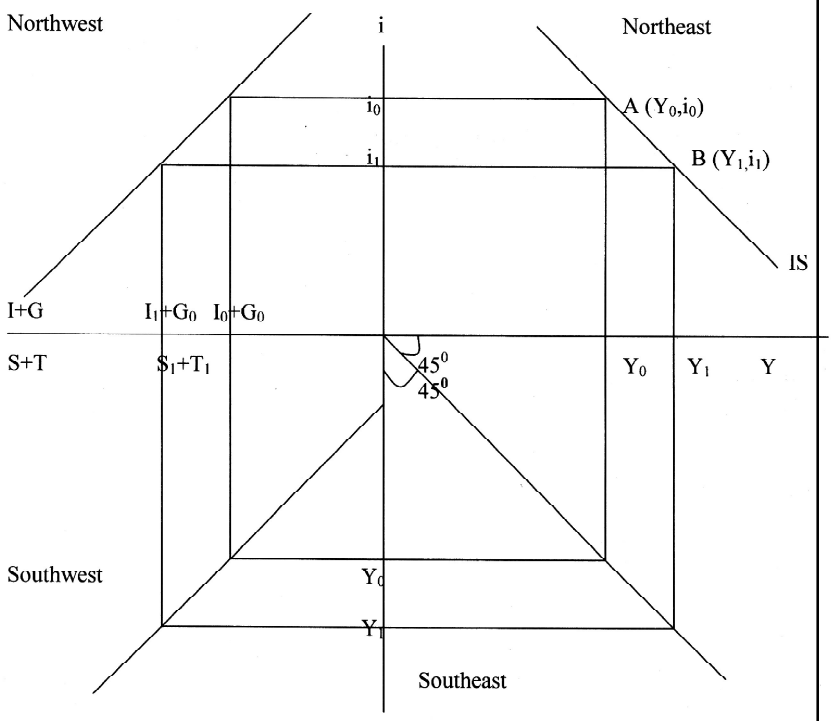


Diagram 11

Derivation of IS Curve

Now to obtain the other points on the IS curve, we suppose other interest rates and proceed in the same manner. Let us consider the market rate of interest be i. At this interest rate, an investment plus government purchase I_1+G_0 is higher than at interest rate i_0 as at a lower rate of interest investment is higher. This increase in investment shifts the income level. In this case the new equilibrium level of income is Y_1, since Y_1 is the only level of income at savings plus taxes is equal to the investment plus government purchases. Consequently, for interest rate i_1, the equilibrium level of income is Y_1. This combination of income and interest rate is plotted in the northeast quadrant of the figure as the point $B(Y_1,i_1)$. Similarly, if other interest rate are considered, the remainder of the IS curve will be traced out. Now by joining the points $A(Y_0,i_0)$ and $B(Y_1,i_1)$ we can draw the IS curve.

Thus the IS curve consists of the equilibrium combinations of income and interest rate for the product market and the market is in equilibrium at any combination of income and interest rate on the IS curve.

Derivation of the LM curve:

The LM curve consists of the equilibrium combinations of income and interest rate for money market at which demand for money is equal to the supply of money. To derive LM curve in the money market we have taken the following assumptions:

(i) Nominal money supply is exogenously determined by the monetary authority, i.e., $\frac{Ms}{P} = \frac{Ms_0}{P_0}$

(ii) The real demand for money is a function of income and the interest rate, i.e., $\frac{Md}{P} = L(Y, i)$

(iii) The equilibrium condition is $\frac{Ms}{P} = \frac{Md}{P}$.

Under the above assumption, we may derive the LM curve with the help of the Diagram 12.

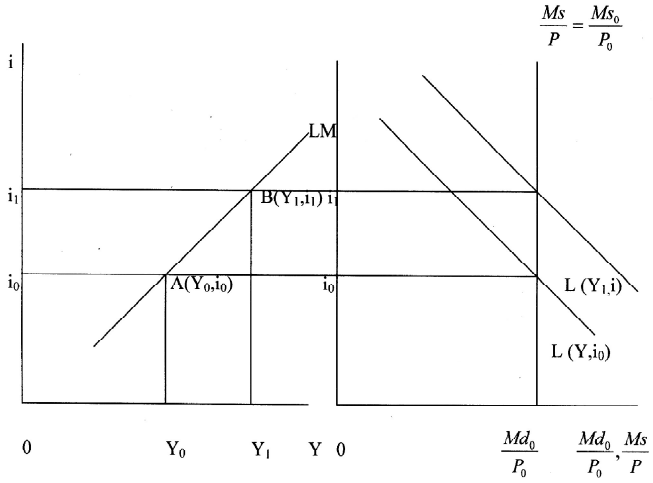


Diagram 12
Derivation of LM curve

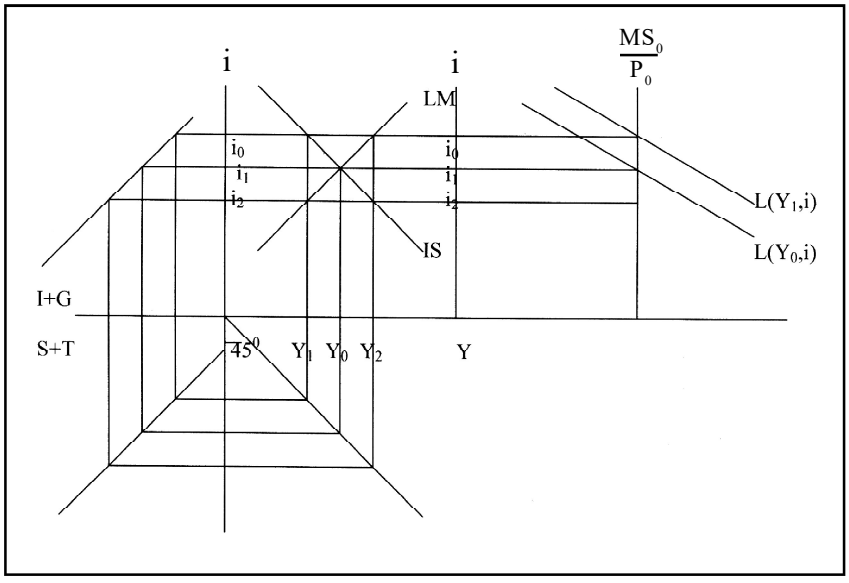
In the Diagram 12, the money supply and demand for money is plotted on the right panel of the diagram. On the other hand, we have plotted interest rate on the vertical axis and income on the horizontal axis. Suppose, the initial level of income is Y_0 the relevant demand for money curve is $L(Y_0,i)$. Given the demand for money and the supply of money, the equilibrium interest rate is i_0 and this is the rate, at which the amount of money demanded, $\frac{Md_0}{P_0}$ is equal to the amount of money supplied, $\frac{Ms_0}{P_0}$. Since i_0 is the equilibrium interest rate corresponding to the Y_0 level of income, the equilibrium combination A (Y_0,i_0) can be plotted in the left hand panel of the

diagram by plotting income level Y_0 on the horizontal axis and then extending a horizontal line from interest rate i_0 in the right panel to the left panel.

Again, suppose a higher level of income, say, Y_1. At the higher level of income, Y_1, the relevant demand for money function is $L(Y_1,i)$. The equilibrium interest rate is i_1 where $\frac{Ms}{P} = \frac{Md}{P}$, which is greater than the interest rate i_0. This equilibrium interest rate is corresponding to the income level Y_1 and hence this equilibrium combination B (Y_1, i_1) can be plotted in the left hand panel of the diagram by applying the same procedure. By considering other income levels and determining the corresponding equilibrium interest rates, other equilibrium points can be derived. If we join all the point, we will get an upward rising LM curve. This curve is called the LM curve as it is derived from the liquidity preference function and the money supply function. It is positively slopped because at higher level of income the amount of money demanded is greater, and with a constant money supply the interest rate must be higher to maintain equilibrium in the money market.

Determination of Interest Rate:

To determine the equilibrium rate of interest and the level of income we have to superimpose both the IS and LM curve. The equilibrium interest rate and the level of income lie on the both curve and then only they will be in equilibrium position. If these points are on different curve, then it will not be equilibrium interest rate and level of income. Let us superimpose both IS and LM curve as drawn in the Diagram 13.



From the Diagram 13 we have seen that at income level Y_0 and interest rate i_0, both the LM and IS curve intersect each other that means this is the equilibrium rate of interest rate and income level as modern theory of interest states that at equilibrium both the IS and LM curve intersects each other.

CHECK YOUR PROGRESS -4

1. What do you mean by interest?

.....

.....

.....

.....

2. State the modern theory of interest.

.....

.....

.....

.....

.....

3. Define IS curve.

.....

.....

.....

.....

4. Define LM curve.

.....

.....

.....

.....

.....

5. What are the important determinants of rate of interest according to the modern theory of interest?

.....

.....

.....

.....

.....

1.7 PROFIT:

Profit is the fourth category of income and it goes to the fourth factor of production, i.e., entrepreneur as remuneration. In ordinary sense, the entire return or income received by the businessman as result of his operations is called profit. Economists used the word profit to connote the surplus, which remains after implicit rent and implicit interest are deducted from the total income derived by the firm. Any surplus left after payment of implicit as well as explicit cost is the net profit. Thus we can distinguish the gross profit and net profit as below:

Gross profit: Total revenue minus total explicit cost. Total revenue refers to the total income earned by the firm by selling its products; and total cost refers to all the explicit costs, i.e., total expenses incurred by the organizer in hiring the factors of production.

Net profit: Total revenue minus total implicit and explicit costs, or gross profit minus implicit costs.

Economists have also made distinction between normal profit and super normal profit. Normal profit is the expected profit in any business without which the entrepreneur may find it not worthwhile to remain in the particular business. Such profit is the minimum to induce the entrepreneur to remain and work in an industry. On the other hand super normal profit is the return above the normal profit.

1.7.1 PROFIT: RISK AND UNCERTAINTY BEARING THEORY OF PROFIT:

Hawley gave the risk-bearing theory of profit. According to him, risk bearing is a special function of an entrepreneur that leads to the emergence of profit. Since business involves risk, it is therefore, necessary to involve some kind of profit for undertaking any new venture. The greater the risk involve in a venture, greater may be the expected gain to induce more an entrepreneur to start the business. All businesses are more or less speculative and unless the risk bearer is amply awarded, businesses would not be started. This view was completely upheld by Hawley, who has given the most complete exposition to his theory in his book "Enterprise and Productive Process" published in 1907. With regard to profit- the reward of an entrepreneur- Hawley states: "….the profit of an undertaking, or residue of the product after the claims of land, labour and capital are satisfied, is not the reward of the management or co-ordination but of the risk and responsibilities that the

undertaker.....subjects to himself to." Expressing the idea differently, he says "The final consumer is forced to include in the price he pays for any product not only enough to cover all the items of cost to the entrepreneur among which items is a sum sufficient to cover the actual or average losses incidental to various risks of all kinds necessarily assumed by the entrepreneur and his insurers- but the further sum, without which, as an inducement, the entrepreneur or enterprise and his insurers will not undergo or suffer the irksomeness of being exposed to risk. This surplus of consumer's cost over entrepreneur's cost, universally regarded as profit, and, from the nature of the case, an unpredetermined residue is the inducement for assumption by the entrepreneur, of all the risks, whatever their nature, necessitated by the process of production. "

Thus to Hawley, risk taking is an essential function of the entrepreneur and is the basis of profit. In his theory of distribution, entrepreneur is the only real productive factor while others - land, labour and capital- have been relegated to the subordinate position of mere means of production.

On the other hand Knight gives the uncertainty bearing theory of profit. According to Knight, it is uncertainty bearing rather than risk bearing, which is a special function of the entrepreneur and which occasions profit. If the future is perfectly known and certain, there would be no risk and consequently no profit. Profit is closely tied with uncertainty. According to Knight, risks inherent in any business are of two kinds: insurable and non-insurable. The risk that can be calculated statistically and thus can be insured are of two kinds: (a) risk to property or loss to property due to fire, earthquake and other calamities; and (b) risk of dishonesty such as loss due to theft, robbery, burglary, etc. These insurable risks are not the concern of the entrepreneur because by paying premium to the insurance company the problem can be solved. But there is certain non- insurable, which cannot be reduced to statistical measurement and thus, cannot be treated on actuarial basis, and yet the entrepreneur himself must undertake these risks if he is to carry on production at all. These non-insurable risks are:

1. Competition risks that arise from the possibility of other entrepreneurs entering the already populated industry or of the development of some new and competitive product.
2. Technical risks that arise from the possibility of newly installed machinery becoming obsolete.
3. Risks of Government policies.
4. Risks of business cycles.

5. Changes in Fashion.
6. Risks of marketability.

he above risks are non-insurable and hence the entrepreneur himself has to take all the risk if he wants to carry on the business. Knight calls these non-insurable risks as "uncertainties" as to him the term risk can be used to those risks, which are known and foreseen. Profit is said to arise for undertaking uncertainties of business.

1.7.2 SCHUMPETER'S THEORY OF PROFIT:

Schumpeter's theory of profit is known as the innovations theory. He resorts to innovations to explain the emergence of profit. An innovation is more than having a new idea. It is different from that of the invention. An inventor may invent something new but until and unless it has been owned by someone and put in to operation, there will be no original economic contribution. Money is needed for this and it must be risked if a new idea is to be put in to effect. It may often happen that one man may design a new machine, another may have the vision to see its commercial use and yet another may be required to advance money to realize these potentialities. All these man are the part and parcel of the innovation process and there fore all these three people should get profit.

The motive for introducing innovations that lead to economic progress is to obtain profit. Profit, therefore, is the cause of innovations. And if the innovations turn out to be successful, the result will also be profit. To Schumpeter, profit is the cause and effect of innovations and therefore, is the cause and effect of economic progress also.

Innovations, according to Schumpeter can be of several types, such as, introduction of new machine; enlargement of the size of an undertaking; exploitation of new source of raw material; change in the quality or grade of the commodity; discovery of new markets, etc. Whenever, any such change is introduced, it is necessary to go for a new combination of factors of production that would reduce the cost of production thereby resulting a profit. Profit accrues to the man who neither conceived the new idea of an innovation nor finances it but he is the man who introduces innovation. But here we will have to remember that if an innovation is no longer novel and patented, all the advantages and disadvantages of the idea or the process that was novel once become well known and consequently profit disappear. Thus the profit result of innovation is always temporary and would be competed away by rivals and

imitators. However, while one source of innovational profit is disappearing some new or cleverer innovation may be born. Consequently, innovational profits have a tendency to appear and disappear.

The theory suffers from two drawbacks: (a) it fails to take in to account that uncertainty as an essential dynamic element of the economic situation. (b) Schumpeter also denies the role of risk bearing in the determination of profit.

CHECK YOUR PROGRESS - 5

1. Define profit.

.....

.....

.....

.....

2. What do you mean by gross profit and net profit?

.....

.....

.....

.....

3. State risk bearing theory of profit.

.....

.....

.....

.....

4. What is innovation?

.....

.....

.....

.....

1.8 LET US SUM UP

This unit is an attempt to study the theories of distribution. Here we discuss about the marginal productivity theory of distribution, Euler's theorem and the adding-up controversy, the concepts and theories relating to the factors of production, viz, rent, wage, interest and profit. The marginal productivity theory is a bold

attempt on the part of the economists to evolve a general theory of distribution, which will explain the determination of factor prices such as wages, rent, interests and profits. The marginal productivity theory states that 'in equilibrium each productive agent will be rewarded in accordance with its marginal productivity'. Adding up theorem is one of the important issues in the theories of distribution. The problem of providing that the total product will be just exhausted if all factors are paid rewards according to their marginal productivity has been called the "adding up" problem. The adding up problem states that in a competitive factor market when every factor employed in production process is paid according to its marginal product, then payments to the factors exhaust the total product.

David Ricardo developed a systematic theory of rent and defined rent as "that portion of the produced of the earth which is paid to the landlord for the use of the original and indestructible powers of the soil." After discussing about rent we have turned our attention to discuss about the concept of wage and the different theories of wage. Here we have discussed about monopolistic exploitation. Following A.C Pigou, Robinson defines 'exploitation' as a wage less than the marginal physical product of labour valued at its selling price. In other words, according to Robinson, labour is exploited when it is paid fewer wages than the Value of its Marginal Product (VMP). Next we have discussed about interest and the different theories of interest. Interest is a payment made by a borrower to the lender for the use of a sum of money for a period of time. The Classical theory of the interest seeks to explain the determination of interest rate by real factors like productivity and thrift. It treats interest rate as an equilibrating factor between the demand for and the supply of investible fund consisting of savings. Criticizing the Classical theory of interest, Keynes, on the other hand, developed his famous liquidity preference theory of interest in 1936. According to him, the rate of interest is purely a monetary phenomenon and is the reward for parting with liquidity. Lastly we have discussed about profit and various theories of profit. Hawley gave the risk-bearing theory of profit. According to him, risk bearing is a special function of an entrepreneur that leads to the emergence of profit. On the other hand Knight gives the uncertainty bearing theory of profit. According to Knight, it is uncertainty bearing rather than risk bearing, which is a special function of the entrepreneur and which occasions profit. Schumpeter developed innovation theory of profit. He resorts to innovations to explain the emergence of profit.

1.9 KEY WORDS

Marginal Productivity: By marginal productivity we meant the increment made to total product by employing an additional unit of factor, keeping all other factors constant.

Marginal Revenue Productivity: Marginal revenue productivity refers to the addition made to total revenue resulting from the employment of an additional unit of factor, the other factor remaining the constant.

Quasi Rent: The term quasi rent is applied to the income earned from the appliances and machines resulting from the effort of the mankind.

Wages: Wages can be defined as the payment made to the workers for their participation in the productive activities.

Money Wages: Money wages refer to the payment made to the workers for their mental and physical services considered either in per hour or per day or per week or per month.

Real Wages: real wages refer to the net worth in terms goods and services of the worker's money remuneration, i.e., the amount of necessaries, comforts and luxuries of life that the worker can command in return for his services.

Interest: Interest is a payment made by a borrower to the lender for the use of a sum of money for a period of time.

Profit: Economists used the word profit to connote the surplus, which remains after implicit rent and implicit interest are deducted from the total income derived by the firm.

Net Profit: It is the total revenue minus total implicit and explicit costs, or gross profit minus implicit costs.

Normal Profit: Normal profit is the expected profit in any business without which the entrepreneur may find it not worthwhile to remain in the particular business.

Super Normal Profit: The return above the normal profit is known as the super normal profit.

1.10 SUGGESTED READINGS

1. Microeconomic Theory, Koutsoyianis, Macmillan
2. Advanced Economic Theory, Ahuja, S. Chand.
3. Advanced Economic Theory, Chopra, Kalyani.

Unit- 2: LINEAR PROGRAMMING

Structure

2.1 INTRODUCTION

The linear programming is a mathematical technique developed in recent origin that deals with the optimization of a function subject to a set of constraints. This technique has been used extensively in recent years in various problems of profit maximization, cost minimization, transportation, diet, blending etc.

2.2 DEFINITION OF LINEAR PROGRAMMING:

Linear programming is purely a mathematical technique for the analysis of optimum decisions subject to certain constraints in the form of linear inequalities. It deals with the optimization (maximization or minimization) of a function subject to a set of linear inequalities and/ or inequalities known as constraints. The objective function may be profit, cost, production capacity or any other measure of effectiveness, which is to be obtained in the best possible manner. The constraints may be imposed by different sources such as market demand, production processes and equipment, storage capacity, raw material availability, etc. By the term linearity in mathematics means an expression in which the

variables do not have powers. Simply a relationship is said to be linear if it gives straight line when plotted in graph. Linear programming technique is successfully used in the field of military, industry, economics, transportation system, health sector and even behavioral and social sciences.

L.V. Kantorovich, a Russian mathematician first formulated the linear programming technique. Several economists like Koopmans, Dorfman, Solow, Cooper have also contributed to the development of the technique. But it was mathematician George Dantzig who first developed the general computational technique, the simplex method (which is still considered as the most powerful and efficient technique), in 1947 while working on a project for U.S. Air Force is called as the father of linear programming technique.

2.3 REQUIREMENTS OF A LINEAR PROGRAMMING PROBLEM:

The basic objective of the use of linear programming technique is to make optimum use of limited resources. If resources were not limited perhaps the need of management technique like this would not arise. However, the application of linear programming technique rests on certain requirements:

1. *A set of non-negativity constraints:* This condition is required because we are not interested in getting negative solutions, e.g. negative output or negative price do not have any meaning.

2. *A set of linear constraints:* This represents the side conditions or the limitations or constraints involved in the solution of the problem. Such constraints are usually due to technological limitations, or resource limitations, or capacity limitations, or time limitations etc.

3. *An objective function:* A well-defined objective function is required either to maximize or to minimize subject to the above constraints.

2.4 APPLICATION OF LINEAR PROGRAMMING:

The linear programming technique can be successfully applied to a wide variety of problems. Some of these are discussed below:

1. *Military problems:* This technique is used in military planning problems.

2. *Manufacturing problems:* Linear programming technique is used to find out the number of items of each type that should be produced so as to maximize profit.
3. *Production problems:* It is used to decide the production schedule to satisfy the market demand and to minimize the labour cost, storage cost, etc.
4. *Purchasing problems:* Linear programming technique is also applied to minimize the cost of production in processing of goods purchased from outside and varying in quantity, quality, and prices.
5. *Diet problems:* This technique is used in the preparation of hospital diets, even at household levels to fix the minimum requirement of nutrients, subject to the availability of foods and their prices.
6. *Transportation problems:* Linear programming is also used to determine the optimum amount of goods to be transported from each warehouse to each of the retail stores in order to minimize the total costs of goods transportation. It can also be used in case of passenger transportation to minimize operation cost.
7. *Job assigning problems:* It is also used to assign jobs to workers for maximum effectiveness and optimum results subject to restrictions of wages and other costs.

2.5 BASIC CONCEPTS OF LINEAR PROGRAMMING:

In order to formulate a linear programming problem, we have to know certain basic concepts such as the objective function, constraints, non-negativity constraints, feasible region, feasible solution, iso-cost line etc. Let us discuss these concepts briefly below:

2.5.1 OBJECTIVE FUNCTION:

In every linear programming problem, there always exists a definite objective with a set of constraints. The objective of the linear programming problem may be maximization of profit, revenue, sales or output, or minimization of costs such as production, diet or in transportations, etc. It is generally expressed in terms of linear equations with the dependent variables like profit, revenue, output, sales, costs etc., on the left hand side of the equation and

other relevant independent variables on the right hand side of the equation. Such functional relationship with a definite objective of maximization or minimization is known as the "objective function" of the linear programming problem. The objective function should be expressed as a linear function of the decision variables. In short, objective function, which is also known as the criterion function, describes the "determinants of the quantity to be maximized or to be minimized". If for example, the objective of a firm is to maximize output or profit, then this is the objective function of the firm. An objective function has two parts: the primal and the dual. If the primal of the objective of the firm is to maximize output, then its dual will be the minimization of the cost.

2.5.2 CONSTRAINTS:

The maximization or minimization of the objective function is subjected to certain limitations, which are called constraints or restrictions. The constraints or restrictions are the limitations or bounds imposed on the solution of the problem. For example, if a firm possesses a maximum amount of Rs. 60,000 for investment, then the firm cannot spend/invest more than this amount, which is obviously a limitation on the part of the said firm. Such type of limitation is called as "investment constraint" of the firm. Similarly, the firm may have storage problem, say cannot store more than 60 units of its product. This is another limitation o the firm and we can call this as the "space constraint."

Constraints are also called as inequalities because they are generally expressed in terms of inequalities. In our example, investment constraint is expressed as less than or equal to Rs. 60,000 (i.e., $\leq$60,000) and the space constraint is expressed as less than or equal to 60 (i.e., $\leq$60).

2.5.3 ISO-COST AND ISO-PROFIT LINE:

The concept of iso-cost and iso-profit lines have been playing very important role in the solution of linear programming problem through graphical method. The term "iso" means equal and thus the iso-cost and iso-profit lines show equal amount of cost and profit respectively. . Such profit and cost functions are expressed in the form of linear equation of first degree. Therefore, they represent straight lines.

Such iso-cost and iso-profit lines are drawn through the corner points of the feasible region. Thus, we can get a set of parallel iso-profit or iso-cost lines in the graph paper. In case of iso-profit lines,

the outer most line from the origin "O" is selected for the optimum solution, as it will show the maximum profit. However, it must also touch one of the corner points of the feasible region. Similarly, in case of iso-cost lines the near most such line is selected for optimum solution as it shows the minimum costs.

2.5.4 FEASIBLE REGION:

Feasible region is that region where all the constraints are satisfied. All feasible solutions lie within the feasible region. However, there is large number of feasible solutions within the feasible region itself. Therefore the problem arises how to choose the optimum solution. In the graphical methods, the corner points of all feasible regions are considered to get optimum solution.

2.5.5 FEASIBLE SOLUTION:

The feasible solution can be defined as a point which specifies such values to all the variables involved in the objective function of the problem which would satisfy both types of constraints: structural and non-negativity.

2.5.6 OPTIMUM SOLUTION:

The best of all feasible solutions is the optimum solution for a linear programming problem. In other words, the optimal solution is the best of all feasible solutions. If a feasible solution maximizes or minimizes the objective function, then it is an optimum or optimal solution. For example, if the objective function of a businessman is to maximize profit by selling a combination of two goods, viz., radio and T.V., then the optimal solution will be that combination of radio and T.V. that maximizes the profit of the businessman. On the other hand if the objective of the businessman is to minimize the cost by the choice of a process or combination of processes, then the process or combination of processes, which actually minimizes the cost, will represent the optimum solution. The optimum solution will lie within the feasible solution.

2.5.7 NON-NEGATIVITY CONSTRAINTS:

In linear programming problem a set of non-negativity constraints are taken in to consideration. This non-negativity constraints express the necessity that the level production, price, cost of commodities or transportation cannot be negative, since in

economics, negative quantities do not carry any sense. These non-negative constraints can be expressed as:

$X \geq 0, Y \geq 0$ and so on.

2.6 GRAPHICAL SOLUTION OF THE LINEAR PROGRAMMING PROBLEMS:

Two methods are available for solving the linear programming problems. One is the simple graphical method and the other is the mathematical method, known as the simplex method. The graphical method is simple and is presented on a two dimensional diagram. It is suitable when one considers only two variables and when we have to consider more than two variables; the graphical solution becomes extremely difficult to draw any conclusion. In such cases, simplex method is extremely useful. The graphical solution involves two steps: (1) the graphical determination of the region of feasible solution and (2) the graphical presentation of the objective function.

Let us consider the following problem to have an idea about the graphical method of solving the linear programming problems:

A manufacturer produces nuts and bolts for some industrial machinery. It takes 1 hour of work on machine A and 3 hours on machine B to produce a package of nuts while it takes 3 hours on machine A and 1 hour on machine B to produce a package of bolts. He earns profit of Rs. 2.50 per package of nuts and Rs. 1.00 per package of bolts. How many package of each should be produced so as to maximize his profit if he operates his machine almost 12 hours a day. What is the value of maximum amount of profit?

If this is the nature of the problem, we have to translate first the problem in to mathematical form by finding the objective function and the other necessary constraints.

To solve the problem let us consider that the manufacturer produces X packages of nuts and Y packages of bolts to maximize his profit (π). Therefore our objective function becomes:

$\pi = 2.50X + 1.00Y$, Which is to be maximized subject to the following constraints:

$X+3Y \leq 12$

$3X+Y \leq 12$

Non-negativity Constraints:

$X \geq 0, Y \geq 0$

Now to represent the above inequalities in graph paper, we have to transform them in the following form:

X+3Y=12

3X+Y=12

The next step is to find out the coordinates for the both equation by assuming the value of X and Y as 0.

Therefore,

When X=0, Y=4; so A (0,4).

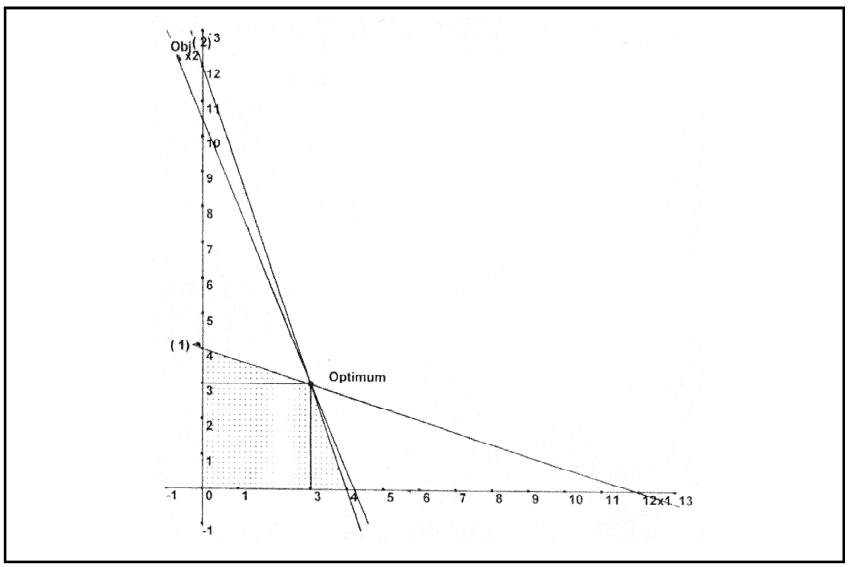
When Y=0, X=12; so B (12,0).

Similarly, when we take the 2nd equation: 3X+Y=12, we get the following coordinates:

When X=0, Y=12; C (0,12)

When Y=0, X=4; D (4,0).

Now we are in a position to plot the above points on the graph and by plotting it in the graph we get OADE is the feasible region, which will yield optimum solution. There are four corner points O (0,0), A (0,4), E (3,3) and D (4,0).



The next step is to draw the iso-profit line through these corner points to choose the outer most iso-profit line, which yields maximum profit. We can ignore the point O (0,0), which implies:

π= 2.50 X + 1.00 Y

Iso-profit line passes through A (0,4):

Therefore, by substituting X=0 and Y=4 in π= 2.50X+ 1.00 Y, we get:

π= 2.50 X 0+ 1.00 X 4

= 0+ 4.00

= 4.00

$\therefore$ 2.50X+ 1.00 Y=4.00

Now, when Y=0, substituting it in 2.50X+ 1.00 Y=4.00, we get X=1.6. Thus, when X= 0, Y=4; A (0,4) and when Y=0, X=1.6; F (1.6,0).

Iso-profit line passes through point E (3,3):

π= 2.50 X 3+ 1.00 X 3 = 10.50

$\therefore$ 2.50X+ 1.00 Y = 10.50

Thus when, X = 0, Y = 10.50; H (0,10.5) and when Y=0, X = 4.2; I (4.2,0).

Iso-profit line passes through point D (4,0):

π= 2.50 X + 1.00 Y

$\therefore$ 2.50 x 4 + 1.00 x 0 = 10

Thus when, X = 0, Y = 10; J (0,10) and when Y=0, X = 4; D (4,0).

Conclusion:

After plotting the iso-profit lines in the graph through the corner points of the feasible region, it is found that the iso-profit line: 2.50X + 1.00Y = 10.50 is the outer most iso-profit line from the origin "O" indicating maximum profit of Rs. 10.50. this particular iso-profit line passes through the corner point E (3,3). It indicates that the manufacturer will produce 3 packages of nuts and 3 packages of bolts to maximize his profit at Rs. 10.50.

Let us take another example where our objective is to minimize the cost of diet without compromising with the minimum requirements of food value. A dietician wishes to mix two types of food in such a way that the vitamin contents of the mixture contain at least 8 units of vitamin A and 10 units of vitamin C. food I contains 2 units of vitamin A and 1 unit of vitamin C, while food II contains 1 unit of vitamin A and 2 units of vitamin C. it costs Rs. 5.00 per unit to purchase food I and Rs. 7.00 per unit to purchase food II. Determine the minimum cost of such a mixture.

Solution:

Let us assume that the dietician mixes X units of Food I and Y units of food II to minimize cost C. With the information given in the problem, we can construct the following table:

Food	Vitamin A	Vitamin B
Food I	2	1
Food II	1	2

Also given, cost to purchase per unit of food I = Rs. 5.00 and cost to purchase per unit of food II = Rs. 7.00.

The objective function is:

5X+7Y=C is to be minimized subject to the following constraints:

$2X+Y \geq 8$

$X+2Y \geq 10$

Non-negativity constraints: $X \geq 0, Y \geq 0$

To represent the above inequalities in graph, we transform the equations into the following form:

$2X+Y = 8 \longrightarrow (1)$

$X+2Y = 10 \longrightarrow (2)$

For equation (1),

When $X = 0, Y = 8$, A (0,8)

& when $Y = 0, X = 4$, B (4,0)

For equation (2),

When $X = 0, Y = 5$, C (0,5)

& when $Y = 0, X = 10$, D (10,0)

Now we are in a position to plot the above points on the graph and by plotting it in the graph we get AED is the feasible region, which will yield optimum solution. There are four corner points A (0,8), E (2,4) and D (10,0). So, we are to determine the iso-cost lines passes through points A, E and D.

Iso-cost line passes through point A (0,8):

We know that, $C = 5X + 7Y$

When $X = 0, Y = 8$

$C = 5x\ 0 + 7x8 = 56 \therefore C = 56$

$\therefore 5X + 7Y = 56$

Thus, when $X = 0, Y = 8$; A (0,8) and when $Y=0, X = 11.2$; F (11.2,0)

Iso-cost line passes through point E (2,4):

Now substituting $X = 2$ and $Y = 4$ in $C = 5X + 7Y$, we get,

$C = 5 \times 2 + 7 \times 4 = 38$

$\therefore 5X + 7Y = 38$

Thus, when $X = 0, Y = 5.43$; G (0,5.43) and when $Y=0, X = 7.6$; H (7.6,0)

Iso-cost line passes through point D (10,0):

Again, substituting, $X = 10$ and $Y = 0$ in $C = 5X + 7Y$, we get,

$C = 5 \times 10 + 7 \times 0 = 50$

$\therefore$ 5X + 7Y = 50

So, when X =10, Y = 7.14; I (0,7.14) and when Y=0, X = 10; D (10,0)

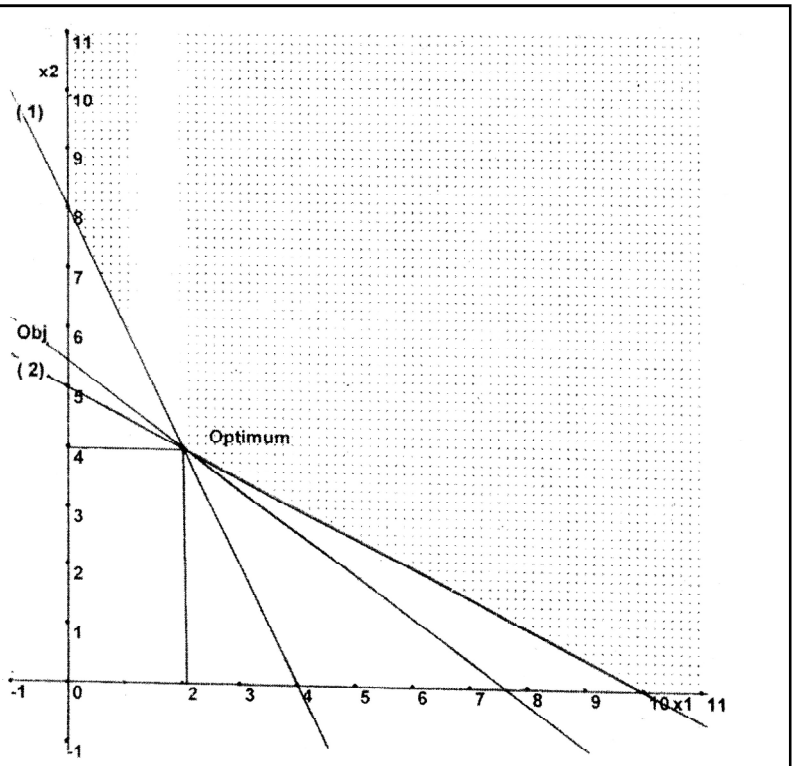


Diagram 3

Conclusion:

After plotting the iso-cost line in the graph, we have seen that the innermost line from the origin O indicating the minimum cost. This cost line passes through the corner point E (2,4) of the feasible region. Thus we may conclude that the dietician mix the vitamin contents in food I is 2 units and in food II is 4 units to minimize their cost 38.

Thus from the above discussion, it is to be noted that for solving any linear programming problem applying graphical methods involves the following steps:

Step I: Formulation of the linear programming problem stating the objective function and the structural as well as the non-negativity constraints.

Step II: Conversion of the inequalities into equality form to find out the coordinates for plotting graphs so that we may get the feasible region and the iso-cost line.

Step III: From the drawn graph, then one has to see the outer most

line for maximization problem and innermost line for minimization problem.

2.7 SIMPLEX METHOD:

The graphical method of solving linear programming problem is very crude. When we have to solve the problem of more than two variables in objective function and structural constraints, then the problem itself will not lead easily to graphical solution. Such type of problem can be dealt with the method called as simplex method. The simplex method involves certain steps and if we proceed accordingly, the optimal solution for the problem will exist.

CHECK YOUR PROGRESS -1

1. What is linear programming?

.....

.....

.....

.....

.....

2. Mention any two uses of linear programming technique.

.....

.....

.....

.....

.....

3. Define:

(i) Objective function

.....

.....

.....

.....

(ii) Constraints

.....

.....

.....

.....

(iii) Feasible region

.....

.....

.....

.....

(iv) Feasible solution

.....

.....

.....

.....

(v) Optimal solution

.....

.....

.....

.....

4. Solve the following problems:

(i) A businessman deals in only two items (simplified version of a real life situation), black and white TV and radio. He has Rs. 50,000 to invest and show room space of store almost 60 pcs. A TV costs him Rs. 2,500 and the radio costa him Rs. 500. He can sale a TV at a profit of Rs. 500 and a radio at a profit of Rs. 150. Assuming he can sell all the items that he buys, how should he invest his money in order he may maximize his profit. What is his maximum profit?

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(Ans. Maximum profit: Rs. 12,500.00, and he will sell 10 TV and 50 radios)

(ii) A dealer wishes to buy a number of washing machine and vacuum cleaners. He has only Rs. 57,600 to invest and has a maximum showroom space for keeping 20 items. A washing machine costs him Rs. 3600 and a vacuum cleaner Rs. 2400. His expectation is that he can sell a washing machine at a profit of Rs. 220 and a vacuum cleaner at a profit of Rs. 180. Assuming he can sell all the items that he can buy. How should he can invest his money to maximize profit? What is the value of maximum amount of profit?

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(Ans. Maximum profit is Rs. 3920 and number of washing machine and vacuum cleaner the businessman sells are respectively 8 and 12.)

(iii) Two carpenter A and B earn Rs. 150 and Rs. 200 per day respectively. Carpenter A can construct 6 tables and 4 chairs while carpenter B can construct 10 tables and 4 chairs per day. How many days shall each work if it is desired to produce at least 60 tables and 32 chairs at a minimum labour cost. What is the amount of such minimum cost?

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Ans. A will work for 5 days and B will work for 3 days and their labour cost will be minimized at Rs. 1350)

2.8 LET US SUM UP

In this unit we have discussed about linear programming technique. From the discussion it is observed that this technique can be applied for constrained optimization problems. There are two methods to solve the problems of linear programming: (a) the graphical method and (b) the simplex method.

2.9 KEY WORDS

Linear Programming: Linear programming is purely a mathematical technique for the analysis of optimum decisions

subject to certain constraints in the form of linear inequalities

Objective Function: Objective function, which is also known as the criterion function, describes the determinants of the quantity to be maximized or to be minimized.

Constrained: The constraints or restrictions are the limitations or bounds imposed on the solution of the problem.

Feasible solution: The feasible solution can be defined as a point which specifies such values to all the variables involved in the objective function of the problem which would satisfy both types of constraints: structural and non-negativity.

Optimal solution: The best of all feasible solutions is the optimum solution for a linear programming problem.

2.10 SELECTED READINGS

1. Operation Research- An Introduction, Taha, Prentice Hall India.
2. Operation Research, Gupta & Hira, S. Chand.
3. Advanced Economic Theory, Ahuja H.L., S. Chand
4. Microeconomic Theory, Koutsoyiannis, Macmillan.

BLOCK - V

WELFARE ECONOMICS

UNIT 1: New Welfare Economics

Structure

1.0 OBJECTIVES

The objective of this unit is to give you a brief idea about the meaning of welfare economics, idea of the Pigovian Welfare Economics and New Welfare Econmics as well concepts, conditions, merits and demerits of Paretian Social Welfare. More specifically after going through the unit, you will be able to:

1. understand what is welfare economics and the two major approaches- Pigovian and Paretian;
2. examine Pareto optimility; and
3. evaluate two fundamental theorems of welfare.

1.1 INTRODUCTION :

Welfare Economics is a branch of economics that uses microeconomic techniques to simultaneously determine the allocational efficiency of a macroeconomy and the income distribution consequences associated with it. It attempts to maximize the level of social welfare by examining the economic activities of the individuals that comprise society.

Welfare economics is concerned with the welfare of individuals, as opposed to groups, communities, or societies because it assumes that the individual is the basic unit of measurement. It also assumes that individuals are the best judges of their own welfare, that people prefer greater welfare to less welfare, and that welfare can be adequately measured either in monetary terms or as a relative preference.

Social welfare refers to the overall utilitarian state of society. It is often defined as the summation of the welfare of all the individuals in the society. Welfare can be measured either cardinally in terms of dollars or "utils", or measured ordinally in terms of relative utility. The cardinal method is seldom used today because of aggregation problems that make the accuracy of the method doubtful, as well as strong underlying assumptions.

There are two sides to welfare economics: economic efficiency and income distribution. Economic efficiency is largely positive and deals with the "size of the pie". Income distribution is much more normative and deals with "dividing up the pie".

1.2 PIGOVIAN WELFARE ECONOMICS:

Arthur Churchill Pigou is the first economist who has given a new look to the welfare economics. The first systematic attempt to discuss welfare economics was his book 'The Welfare Economics'. The current popularity of the branch of economics, the welfare economics can be attributed to the pioneering work done by Pigou. He is also called as the father of welfare economics.

Pigou has concentrated on the concept of economic welfare rather than the concept of social welfare. According to Pigou, economic welfare is that part of social welfare, which can be directly or indirectly related with the measuring rod of money. Pigou's concept of social optimum stands for a situation in which the economic welfare of the society is maximized. Pigou used national income to measure economic welfare. He stated that, 'other things being equal, an increase in the national income tends to improve the economic welfare of the people in the society and vise-versa'.

Pigou had linked his concept of social optimum with the maximization of economic welfare of the society. He further indicated economic welfare in terms of national income. In other words, he looked upon national income as the quantitative indicator of economic welfare. Given the same tastes and income distribution, an increase in the national income represents an increase in welfare. In real terms, an increase in the output other things remaining the

same in a country is similar to an increase in the economoic welfare of the people because they would now derive greater satisfaction than before from the increased output of goods and services. Hence the first condition for the maximization of economic welfare according to Pigou is to maximize real national output.

The second condition for welfare maximization is that the increased real national output should be properly and equitably distributed between the richer and poorer sections of the society. If national income remains constant, transfer of income from the rich to the poor would improve welfare. According to Pigou, such transfers mean less to the wealthy than to the poor, as a result the economic welfare of the poor will improve. This welfare condition is based on the dual Pigovian Postulates of "equal capacity for satisfaction " and "diminishing marginal utility of money". Pigou argued that different people derive the same satisfaction out of the same real incomes. He also assumed that the poor man had the same capacity for satisfaction from money income as rich man had.

To find out improvements in social welfare Pigou adopted dual criterion. First an increase in the real national income brought about either by increasing some goods without diminishing others is regarded as an improvement in welfare. Second, any reorganization of the economy, which increases the share of the poor without reducing the national income, is also considered as an improvement in social welfare.

Discussing his view on economic welfare Pigou made a fundamental distinction between Private Net Product (PNP) and Social Net Product (SNP). The PNP refers to the income, which accrues to the owner of a private business firm or enterprise. On the other hand SNP refers to the contribution made by the enterprise to the national dividend. It is very seldom that these two are equal to each other. Pigou was of the view that state should actively intervene to eliminate the divergence between the PNP and the SNP. If in a particular industry, the PNP is greater than the SNP, that industry should be subjected to taxation to bring about equality. If on the other hand, in an industry the PNP is smaller than SNP it should be subsidized by the State to equalize the two products. In either case, State action will promote the economic welfare of the society.

1.3 NEW WELFARE ECONOMICS :

Pareto criterion deals with only economic change that harms no one and makes someone better off indicates an

improvement in social welfare, but it does not mean that the economic changes which harm some and benefit others. Therefore, economist like Kaldor, Hicks and Scitovosky have sought to extend Pareto's criterion for changes in social welfare to cases in which some individuals better off and other worse off. This extension is popularly known as the compensation principle or the new welfare economics.

1.4 PIGOVIAN WELFARE ECONOMICS Vs NEW WELFARE ECONOMICS:

The New Welfare Economics is a departure from the Pigovian welfare economics. The basic difference between these two analyses is that the former is based on the cardinal utility analysis and thus in Pigovian analysis; welfare can be directly brought in to the measuring rod of money. On the other hand new welfare economics is based on the ordinal utility analysis. In this analysis, welfare is entirely a normative analysis.

1.5 PARETO OPTAMILITY :

Vilfredo Pareto, an Italian economist was the first man to find out an objective test of social welfare maximum, which is often called as Pareto Optimality or Pareto Criterion. According to this concept, a distribution of inputs among commodities and of commodities among consumers is Pareto Optimal or efficient if no reorganization of production and consumption is possible by which some individuals are made better off (in their own value judgment) without making someone worse off. Any reorganization that improves the well being of some individuals without reducing the well being of others clearly improves the welfare of the society as a whole.

Pareto optimality can be explained with the help of Samuelson's utility possibility curve in Diagram 1. Suppose, there are two individuals A and B sharing given bundle of goods X. A's utility is represented on the horizontal axis and B's utility on the vertical axis. Thus AB represents utility possibility curve of all combinations of the individual utilities. According to Pareto criterion, any change, which causes a movement from C to F on the utility possibility curve BA, is an improvement because it makes both individuals better off thereby maximizing their welfare. Similarly a movement from C to D or E on the BA curve is an

improvement for it makes at least one person better off without making the worse off. But any point outside the segment DE is not a Pareto improvement. For instance, any movement from C to Q increases B's welfare at the expense of A's welfare. In such circumstances, Pareto criterion cannot tell us as to whether social welfare increases or decreases.

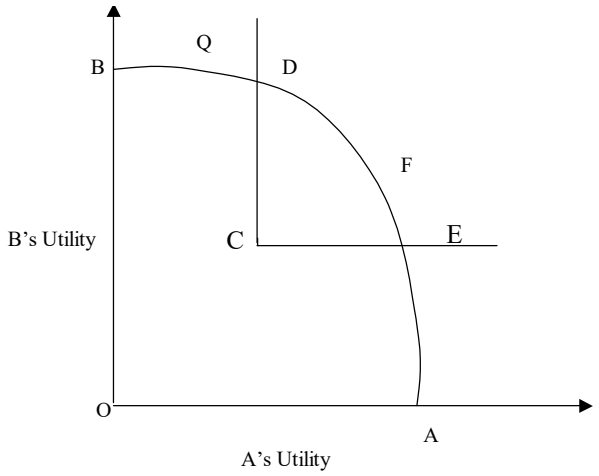


Diagram - 1D

1.5.1 MARGINAL CONDITIONS OF PARETO OPTIMILITY:

Pareto concluded from his criterion that competition leads the society to an optimum position but he did not give any mathematical proof of it, nor he derived the marginal conditions to be fulfilled to achieve the optimum position. Latter on, Lerner and Hicks derived the marginal conditions which must be fulfilled for the attainment of Pareto optimum. These marginal conditions are based on following assumptions: -

The production function of each producer is given i.e. the technical knowledge remains constant.

(1) Each individual has his own ordinal utility function and possesses a definite amount of each product and factor.

(2) Each product is divisible.

(3) Each individual tries to maximize his satisfaction.

(4) Each producer tries to maximize its profits and minimize its production cost.

(5) Factors of production are perfectly mobile.

Given these assumptions, the conditions of optimum welfare are discussed below:

(1) The optimum distribution of products among the consumers (efficiency in exchange): - This condition states that the marginal rate of substitution (MRS) between any two products must be the same for every individual who consumes both $MRS_{xy}^{A} = MRS_{xy}^{B}$. Where A and B refer to individuals and x and y to commodities.

This condition is illustrated in following Diagram 2.

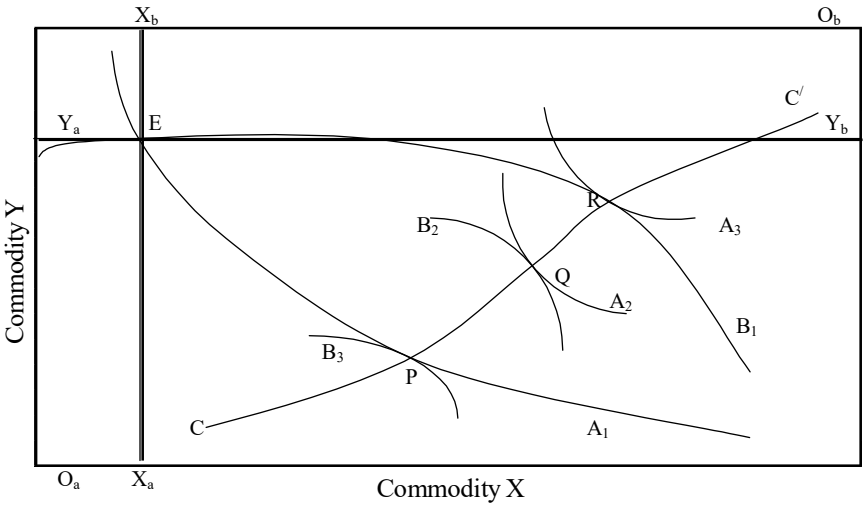


Diagram 2

In box diagram, let two individuals A and B who posses two goods x and y respectively shown on the horizontal and vertical axis. The indifference map of A with indifference curves A_1, A_2, and A_3 with origin Oa superimposed on map by B_1, B_2, and B_3 indifference curve with origin O_b. CC is the contract curve passing through various tangencies P, Q, R of indifference curves of A and B. The marginal rate of substitution (MRS) between two goods for A and B are equal on the various point of the contract curve. Any point outside the contract curve does represent the equality of MRS between the two goods for the two individuals. Say, point E, where two indifference curves A_1 and B_1 intersect instead of being tangential. Therefore, at point E, MRS between two goods of individual A is not equal to that of B. So, E is not the point of optimum exchange. But a movement from E to R increases the satisfaction of A without decreases the satisfaction of B. similarly a movement from E to P increases the satisfaction of B without decreases the satisfaction of A. the movement from E to Q increases the satisfaction of both because both move to their higher indifference curve. Thus, P, Q, R are the three conceivable points of exchange.

(2) The optimum allocation of factors (efficiency of production): - This condition states that the marginal rate of technical substitution (MRTS) between any two factors must be the same for any two firms using both to produce the same products.

i.e. $MRTS_{KL}^{x} = MRTS_{KL}^{y}$

Where K= capital and L = labour.

The diagrammatic representation of this condition is similar to that of the optimum exchange, where, A and B may be treated as firms producing two goods and their indifference curve as representing the corresponding iso-product curves. The locus of the tangency points of the two sets of iso-products curves will be called here the production contracts curves. Any point on it shows equality of MRTS between the factors for the two firms. Hence the conclusion is that if the two firms are on CC societies production will then be optimally organized.

This condition is illustrated in the Diagram 3.

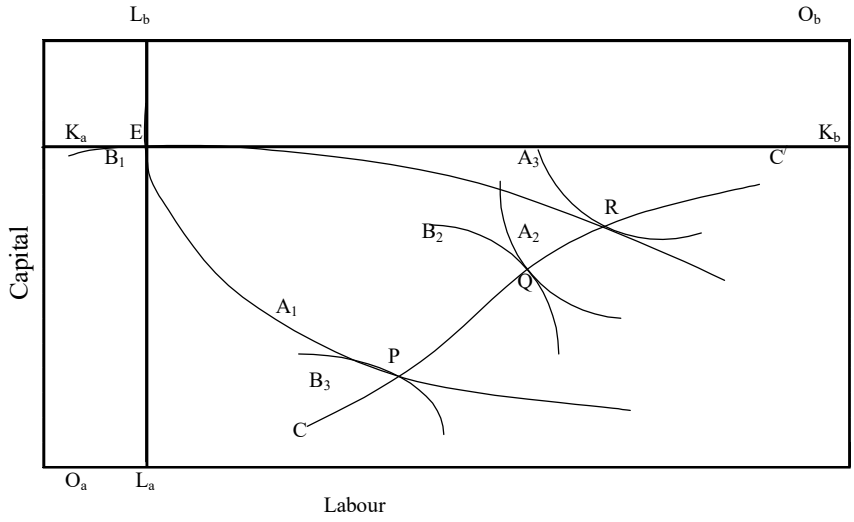


Diagram 3

In the box diagram, suppose two firms A and B producing two goods X and Y by using two factors labour and capital. The available quantities of labour and capital are represented on X and Y-axis respectively. A_1, A_2 and A_3 represent the isoquants of firm A and that of firm B are B_1, B_2 and B_3. The slope of an isoquant gives the MRTS between two factors. $CC^{/}$ is the contract curve passing through tangency points PQR of isoquants of firm A and B. The MRTS between firm A and B are equal on the various points of the contract curve. Any point outside the contract curve does not represent the equality of the MRTS between A and B. Say, point E,

the two isoquants A_1 and B_1 intersect, which is outside the contract curve. Therefore, at point E the MRTS of two firms is not equal and therefore, the point E is not the point of optimum exchange. But a movement from E to R or E to P raises the output of one firm without any decrease in the output of the other. At point Q, the movement from E raises the output of both the firms individually as well as collectively. So, at point P, Q, R the combination of two factors is optimally utilized.

(3) The Optimum Degree of Specialization: this condition requires that the marginal rate of transformation between any two products must be the same for any two firms that produce both, i.e.,

$$MRT_{xy} = MRT_{xy}^{A} = MRT_{xy}^{B}$$

This condition is satisfied when the two products are produced in such combinations that the slopes of transformation curves are equal. To prove this, let TA be the transformation curve of firm A and PB that of the firm B, in panel A and panel B of the Diagram 4 respectively. Each point on the transformation curve shows the greatest possible simultaneous quantities of the two products.

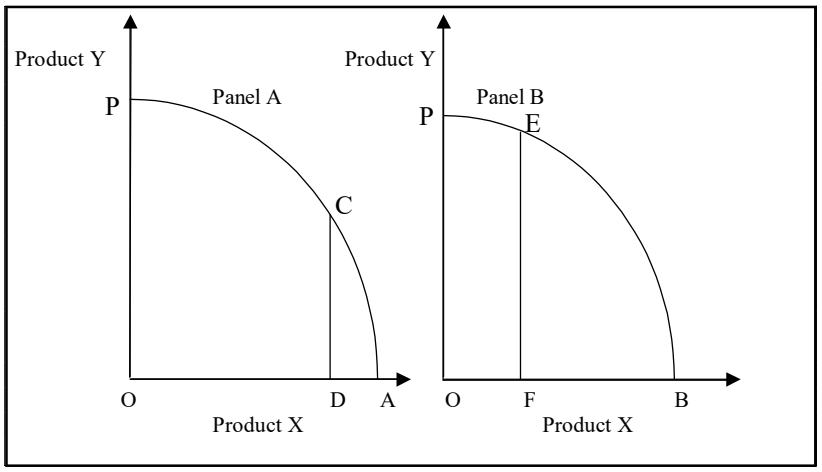


Diagram : 4

Suppose firm A is producing OD of X and Dc of Y and firm B is producing OF of X and EF of Y. Thus total output of both firms is OD+OF of X and DC+ EF of Y. Thus total amount of output of X and Y is shown in Diagram 5 by superimposing Panel A and Panel B of the Diagram 4. They are respectively GH and FD. Since the two transformation curves TA and PB intersect each other at L. But point L is not optimum situation because here marginal rate of transformation is not equal. If, the superimposed figure shifted a little upwards so that its transformation curve $P^{/}B^{/}$ becomes tangential to TA at R, the slopes of the two curves.

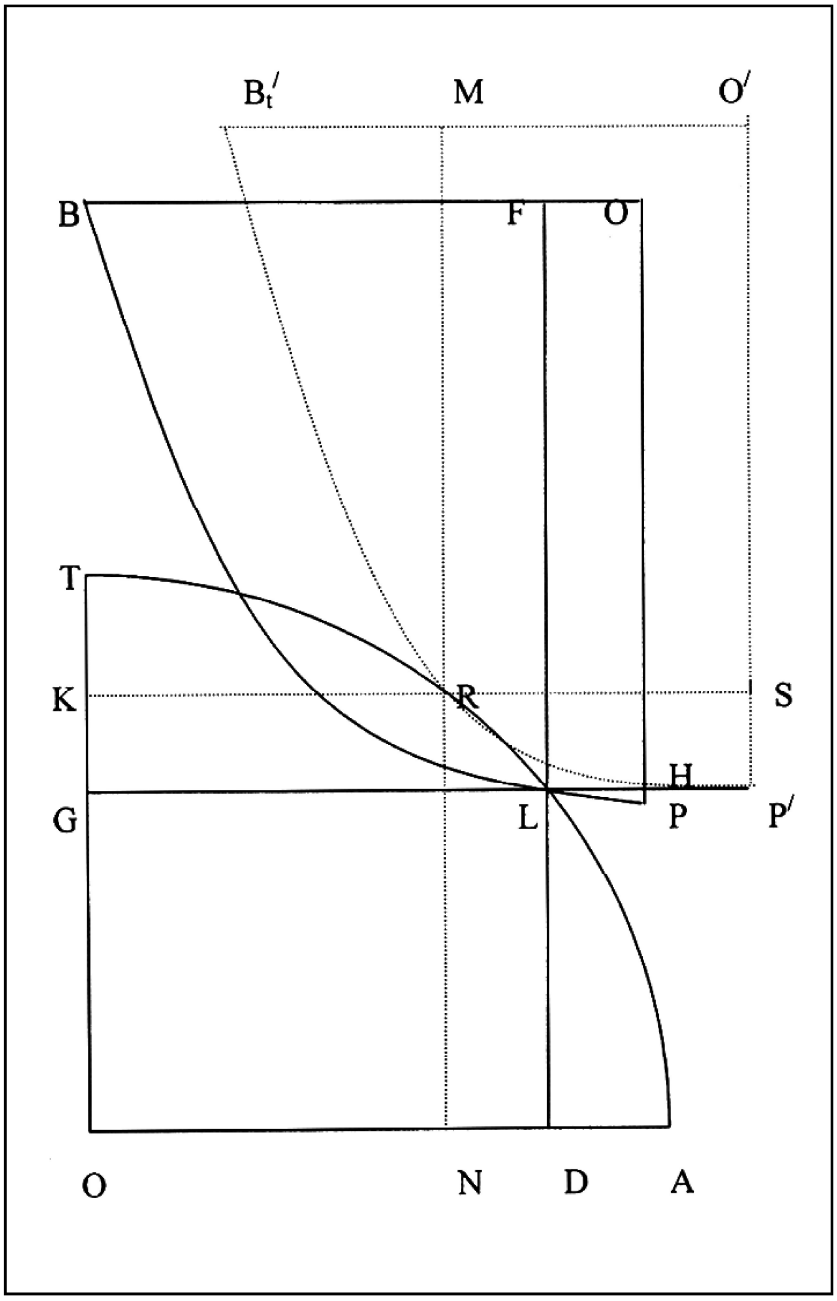


Diagram 5

coincide. Thus this condition is satisfied because at point R, the MRT between two goods are equal

(4) The Optimum Condition of Factor-Product Utilization : This condition states that the marginal rate of transformation between any factor and any product must be the same for any pair of firms using the factor and producing the product, i.e., the marginal productivity of any factor in producing a particular product must be same for all the firms. This condition can be illustrated with the help of a figure. In the figure, suppose OA be the transformation

curve of firm A and OB of firm B which is superimposed on the transformation curve OA, keeping the axes parallel to each other. The product Z being produced by the two firms is measured vertically and the amount of factor L used in its production is measured horizontally. The point of intersection F of the two transformation curves is not one of the optimum positions because they are not tangential to each other. In order to attain the optimum position,

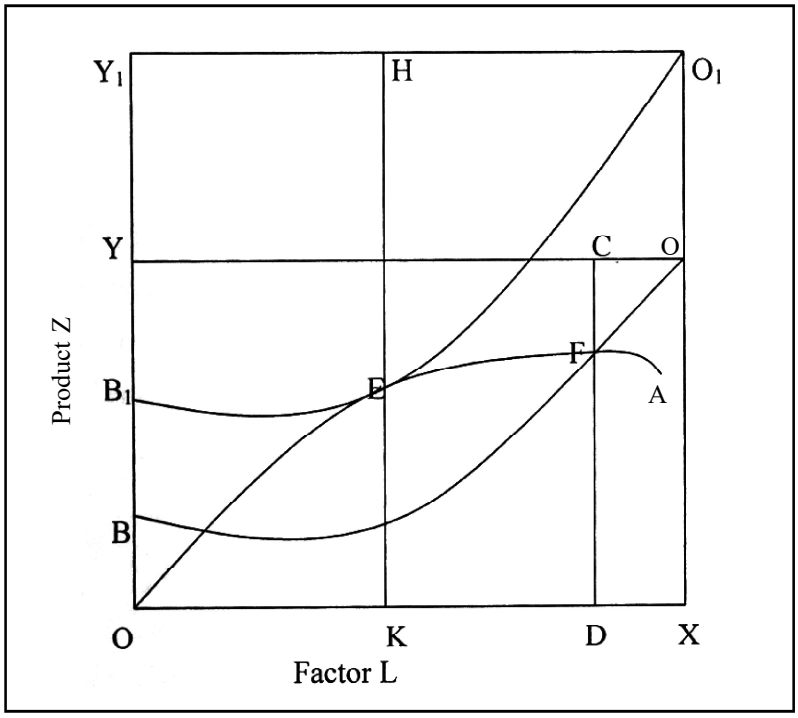


Diagram 6

the OB curve should be shifted upward so that it is tangential to the OA curve. Thus the OB curve becomes O1B1 which is tangential to the OA curve at E. This is the point of optimum factor product utilization because the slopes of the two transformation curves OA and O_1B_1 are equal and the product is now increased from DC to KH (Diagram 6).

(5) The Optimum Condition of the Product-Stabilization: This condition states that, the marginal rate of substitution between any pair of products for any person consuming both must be same as the marginal rate of transformation between them. This can also be explained with the help of Diagram 7

In Diagram 7, let AB be the community transformation curve between two products X and Y. The indifference curves I_1 and I_2 represented an individual consumer of the two goods with O_1X_1 and O_1Y_1 as axes.

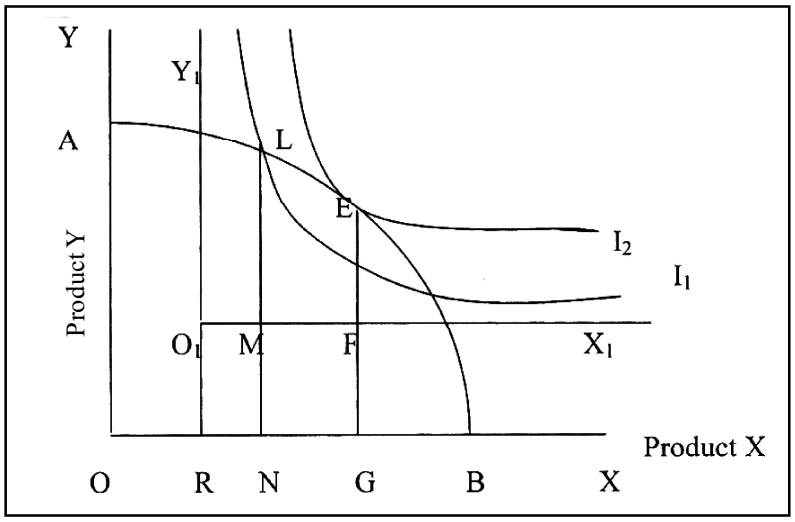


Diagram: 7

Suppose production occurs at L where the community produces ON of X and NL of Y and the consumer buys O_1M of X and ML of Y. But L is not the socially optimal point because the MRT does not equal to MRS. The two curves AB and I_1 are not tangential to each other. A change from L to E equalizes the two curves AB and I_2. Thus, point E represents an optimum position both for the producer and the consumer, MRS = MRT.

1.5.2 MERITS OF PARETIAN SOCIAL OPTIMALITY:

The following are some of the important merits of the Paretian social optimum:

(i) The Paretian concept of social optimum is more rigorous and objectives. It is truncated in scope because it does not take in to consideration income distribution.

(ii) Paretian concept sets forth explicitly the distinction between those changes in social variables, which can take place through "trading", i.e., through a mutual benefit of all parties and those changes, which involve conflict, the benefit of one party at the expense of another.

1.5.3 DEMERITS OF PARETIAN SOCIAL OPTIMALITY:

(i) Paretian social optimum does not define a 'point' but a 'range' of values of the economic universe.

(ii) Paretian social optimum does not include an optimum pattern of income distribution.

(iii) All the conditions of Pareto optimality are fully satisfied under perfect competition. But in real world, the concept of perfect competition is a myth.

CHECK YOUR PROGRESS - 1.1

1. What are the two major approaches of welfare economics?

Ans. ……

.....

.....

.....

.....

.....

2. What are the two sides of welfare economics?

Ans. ……

.....

.....

.....

.....

.....

LET US SUM UP

In this unit we have discussed about the Pigovian welfare economics and the new welfare economics. Thereafter we have concentrated on Paretian welfare criterion and conditions of Paretian optimality. Finally we discussed the merits and demerits of Paretian social optimum.

Key Words:

Welfare Economics: Welfare Economics is a branch of economics that uses microeconomic techniques to simultaneously determine the allocational efficiency of a macroeconomy and the income distribution consequences associated with it.

Pareto Optimal: A distribution of inputs among commodities and of commodities among consumers is Pareto Optimal or efficient if no reorganization of production and consumption is possible by which some individuals are made better off (in their own value judgment) without making someone worse off.

Suggested Readings:

1. Koutsoyiannis, A.; *Modern Microeconomics,* Macmillan
2. Ahuja, H.L; *Advanced Economic Theory,* S. Chand
3. Salvatore, D.; *Microeconomics : Theory and Application,* Oxford.
4. Chopra, P.N.; *Advanced Economic Theory*, Kalyani Publishers.

Terminal Questions:

1. What is Pigovian welfare economics? How it is different from the new welfare economics?
2. Discuss critically the Pareto's welfare criterion. Also discuss the Paretian conditions for optimality.

Answer to Check Your Progress:

Answer for question 1: (i) Pigovian Welfare Economics
(ii) New Welfare Economics.

Answer for question 2: (i) Economic efficiency and
(ii) Income distribution.

Unit- 2 COMPENSATION PRINCIPLE AND THE SOCIAL WELFARE FUNCTION:

Structure

2.0 OBJECTIVES

The basic objective of this unit is to introduce the learner with some of the development in welfare economics after Pareto's contribution. After going through the unit you will be able to know:

1. Compensation Principle of Kaldor-Hicks
2. Scitovosky's Paradox and Scitovosky's Double Criterion
3. Value judgement and Bergson-Samuelson Social Welfare Function
4. Arrow's theory of Social Choice and Impossibility Theorem and
5. Amartya Sen's Concept of Social Choice and Welfare

2.1 INTRODUCTION

It is very important to have knowledge about the recent developments in welfare economics. Therefore in this unit we are making a simple attempt to discuss briefly about the recent developments. At the very beginning of the unit we have discussed

the Kaldor-Hicks compensation principle that is came to be known as the new welfare economics. There after we have made some attempt to discuss some other development in recent past including the contributions made by our own Nobel Laureate Amartya Sen.

2.2 COMPENSATION PRINCIPLE OF KALDOR-HICKS:

Pareto optimality simply states that an economic change, which harms no one and makes some one better off, indicates an increase in social welfare. On the other hand social welfare will decrease if a certain change makes no individual better off while it makes at least one individual worse off. Thus economic situation is said to be Pareto optimum in which allocation of resources is such that by any arrangement of them, it is impossible to make any individual better off without making anyone worse off. In the Pareto optimum welfare position the welfare of an individual of the society can not be increased without decreasing the welfare of another member.

But Pareto criterion does not apply to those economic changes, which harm some and benefit others. Kaldor and Hicks have tried to evaluate the changes in social welfare resulting from any economic redistribution, which harms somebody, and benefit the others. This is known as the compensation principle. According to this new criterion if a certain change in economic policy makes some people better off and others worse off then that change will increase social welfare if those who gain compensate the losers and still be better off than before. Thus Kaldor-Hicks criterion is a departure from Pareto criterion. Kaldor-Hicks criterion can be explained with the help of the utility possibility curve as shown in Diagram 1.

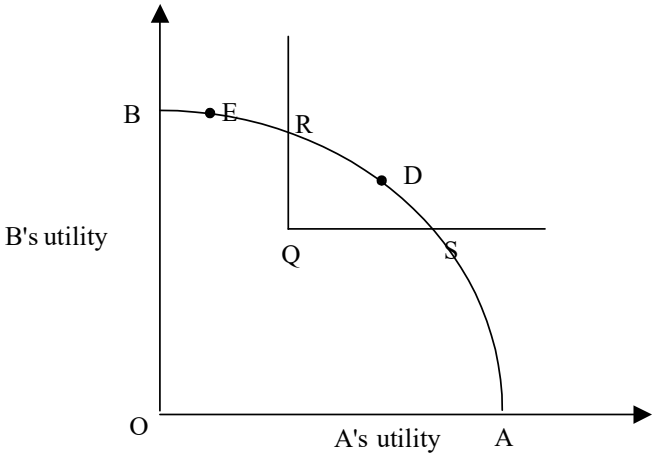


Diagram - 1

In figure utility of the two individuals A and B is shown on X and Y-axis. BA is the utility possibility curve that represents the various combinations of utility obtained by two individuals. As we move downward on the curve, utility of A increases while that of B falls. On the other hand if we move up on the curve, utility of B increases while that of A falls. According to Pareto criterion a movement from Q to R or Q to D or Q to S represents the increase in social welfare. In other words, a movement inside the triangle QRS is an improvement in social welfare. But it does not help us to evaluate the changes in welfare, if the movement as a result of redistribution from point Q to a point outside the segment RS such at point E. Kaldor-Hicks criterion says that such a movement from Q to E also increases social welfare if B compensate A and still be better off than before. This means that by redistribution of income if B gives some compensation to A for the loss suffered they can move to the portion R. At point R, person B is as well off as compared to Q. Therefore, according to Kaldor-Hicks criterion social welfare increases with the movement from position Q to E from where they can move to the position R through more redistribution of income.

Criticisms:

The compensation principle has been criticized on the basis of the following grounds:

(i) Little has pointed out that Kaldor did not formulate a new welfare criterion at all because he assumed welfare to be a function of increase in production or efficiency irrespective of the changes in distribution, i.e., it ignores income distribution.

(ii) In trying to separate production from distribution, the Kaldor-Hicks criterion confuses potential welfare with actual welfare.

(iii) This criterion does not take into consideration the payment of actual compensation.

2.3 SCITOVOSKY'S PARADOX:

Kaldor-Hicks criterion states that a change that benefits some people and harms others can be said to be an improvement if the gainers (Individual B) can bribe the losers (individual A) in to accepting the change while they still remaining better off. But professor Tibor Scitovosky has pointed out that Kaldor-Hicks criterion leads to a contradiction. He showed that if some situation position B is shown to be an improvement over position A on the

Kaldor-Hicks criterion, it may be possible that position A is also shown to be an improvement over B on the basis of same criteria. For getting consistent results when position B has been revealed to be preferred to position A on the basis of a welfare criterion, then position A must not be preferred to position B on the basis of the same criterion. According to Scitovosky, Kaldor-Hicks criterion involves such contradictory results. Since Scitovosky was the first to point out this paradoxical result in Kaldor-Hicks criteria, it is known as the "Scitovosky Paradox."

The "Scitovosky Paradox" is depicted in the Diagram 2. In the Diagram 2, JK and GH are the two utility possibility curves, which intersect each other. Now suppose that initially individuals A and B were on the position C on the JK curve. Further suppose that due to a certain policy change, they are brought to the position D on the utility possibility curve GH. Position D is superior to position C on the basis of Kaldor-Hicks criterion because from position D movement can be made through mere redistribution to position F at which individual B has been fully compensated but individual A is still better off as compared to the original position.

But Scitovosky shows that a reverse movement from D to C is also an improvement in welfare on the Kaldor-Hicks criteria. This is because from position C movement can be made by mere redistribution of income to position E on the utility curve JK. And it is also observed at position E that while A is as well off as at position D, the individual B is still better off than at D. Thus D is socially better than C on Kaldor-Hicks criterion and C is also socially better than D on the basis of the same criterion. So, Kaldor-Hicks criterion leads us to contradictory results.

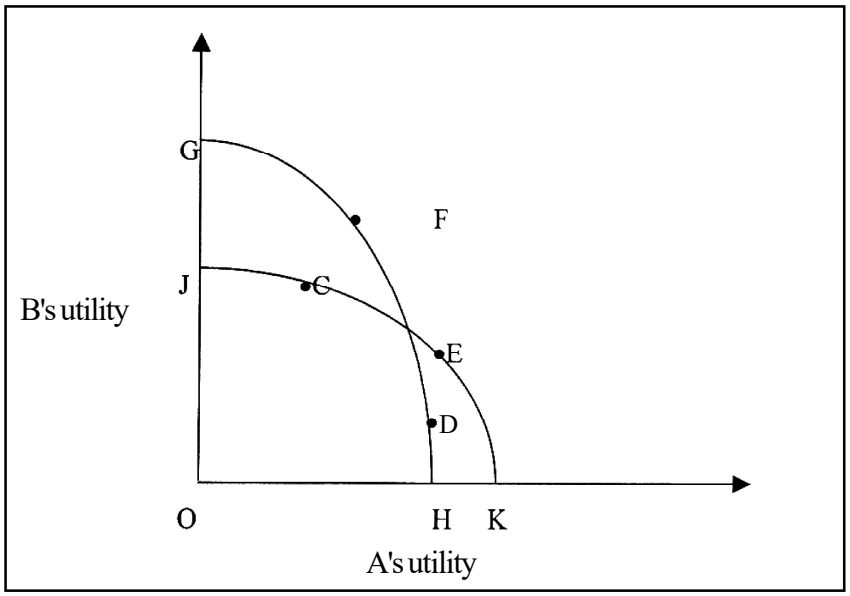


Diagram - 2

2.3.1 Scitovosky's Double Critetrion:

To rule out the possibility of contradictory results in Kaldor-Hicks test, Scitovsky formulated a criterion, which requires the fulfillment of Kaldor-Hicks test and also the fulfillment of reversal test. This criterion is generally known as the "Scitovosky's Double Criterion". According to this criterion, a change is an improvement if the gainer in the changed situation are able to persuade the losers to accept the change and simultaneously the losers are not able to persuade the gainers to remain in the original situations.

Scitovosky's double criterion is satisfied if the two utility possibility curves associated with the position before and after the policy change are non-intersecting. Scitovosky's double criterion can be explained with the help of the utility possibility curve.

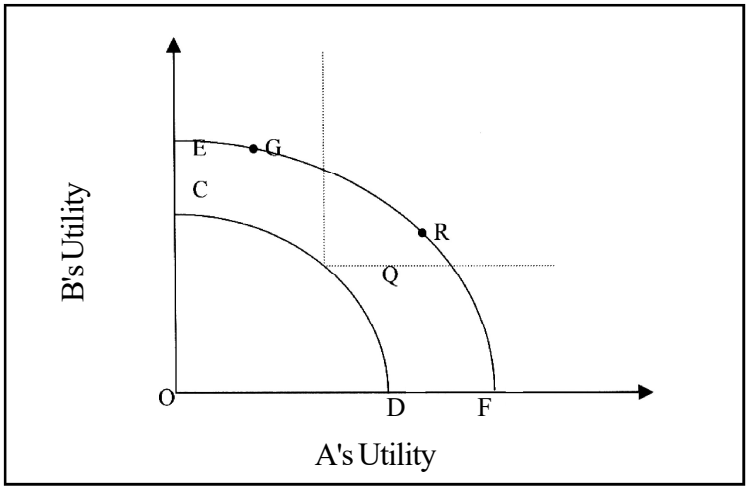


Diagram: 3

In Diagram 3, CD and EF are the two utility possibility curves, which do not intersect each other at any point. Suppose there is a change from Q on the utility possibility curve CD to the position G on the utility possibility EF as a result of the adoption of a new economic policy. Such a movement is an improvement on Kaldor-Hicks criterion because G lies on the higher utility possibility curve than that of Q. The points G and R are on the same curve EF, and it is possible to move from G to R simply by redistribution of income. R is better than Q because the utility of both the individual is greater at R as compared to the position Q. Thus, the Kaldor-Hicks criterion is satisfied and therefore change from Q to G will increase social welfare.

But a reverse movement from G to Q is not passed by the Kaldor-Hicks criterion and satisfied the Scitovosky's reversal test. It is clear from the Diagram 3 that it is possible to move any point on the utility possibility EF simply through the redistribution of

income, but if a change in policy lands the individuals on any point on the lower utility possibility curve CD, then it reduces the welfare of both the individuals. Thus we can conclude that the Scitovosky's double test is passed by a policy change, which moves the utility possibility curve as a whole.

2.4 VALUE JUDGEMENT :

Value judgement refers to an opinion about the relative merits of two or more states of the economy, which cannot be tested empirically. For example, any policy change can affect two individual differently. But it is impossible to test whether these two people are gainer or loser, by asking whether they accept or reject the change, if the choices left to them. Presence of value judgments is an important component of normative economics.

2.5 BERGSON-SAMUELSON SOCIAL WELFARE FUNCTION:

Professor Bergson introduced the concept of social welfare function and later on Samuelson and Tintner developed it. They are of the view that no meaningful propositions can be made in welfare economics without introducing value judgements. Value judgements determine the form of social welfare function with a different set of value judgements. The form of social welfare function will be different.

A social welfare function is simply a set of social indifference curves where an increase in one's individual's welfare, all other individual's welfare remaining the same, increases social welfare. Hence the social welfare function is an ordinal index of society's welfare and is also a function of all individual utility constituting the society. It is expressed as:

$$W = F (U_1, U_2, U_3, \ldots\ldots U_n)$$

Where W is the social welfare, F is the function and U_1, U_2,……..U_n are the levels of utilities of 1,2,3,….n individuals. W is an increasing function of these utilities.

The social welfare function of Bergon and Samuelson is based on the following assumptions:

(i) It assumes that social welfare depends on each individual's wealth and income and each individual's welfare depends in turn on his wealth and income and on the distribution of welfare among the members of the society.

(ii) It assumes that the presence of external economies and diseconomies with their consequent effects.

(iii) It is based on ordinal ranking of combinations of those variables, which influence individual welfare.

(iv) Interpersonal comparisons of utility involving value judgement are freely permissible.

The ordinal utility level of an individual is a function of his own consumption of goods and services and not of others. Moreover, the utility level of an individual depends on his own value judgement regarding the composition of different goods and services which in turn depends on his tastes. However, from the view point of the social welfare function the value judgements regarding the welfare of the society as a whole are relevant. The formulation of a welfare function for the society as a whole is a very difficult task because utility being a mental phenomenon cannot be measured accurately by any person or institutions to furnish value judgements regarding changes in social welfare. According to the advocate of social welfare function, the social welfare function and its form depends upon the value judgements of the person or institution whom the society has authorized to do so. For true value judgements regarding the social welfare he must be unbiased. Bergson and Samuelson had assumed a super man who provided value judgements about changes in the society. Superman can alone take decision about the solution of various problems of the economy. Question such as what goods and services are to be produced and supplied in the society, what should be the quality and type of goods, what should be the pattern of distribution of national income etc. All these can be answered by the superman alone in accordance with his views about the determinants of social welfare. The society could have to accept the solution of these questions provided by the superman under the assumption that he will give value judgements which aims to maximize social welfare rather than to maximize self-interest.

Diagrammatic Representation: We can explain the concept of social welfare function with the help of social indiffernce curves or welfare frontiers. In the following diagram, the utility of individuals A and B has been represented on the horizontal and vertical axes respectively. W, W_1 and W_2 are the social indifference curves representing successively higher levels of social welfare. A social indifference curve is a locus of the various combinations of utilities of A and B which results in an equal level of social welfare. Such curves help the policy makers to find out whether a particular policy brings an improvement or not. If a changes moves individuals

to a higher indifference curve, social welfare is said to have increased. In the diagram a movement from Q to T or from R to S represents an increase in social welfare and a movement from T to S or T to Q represents a decrease in social welfare.

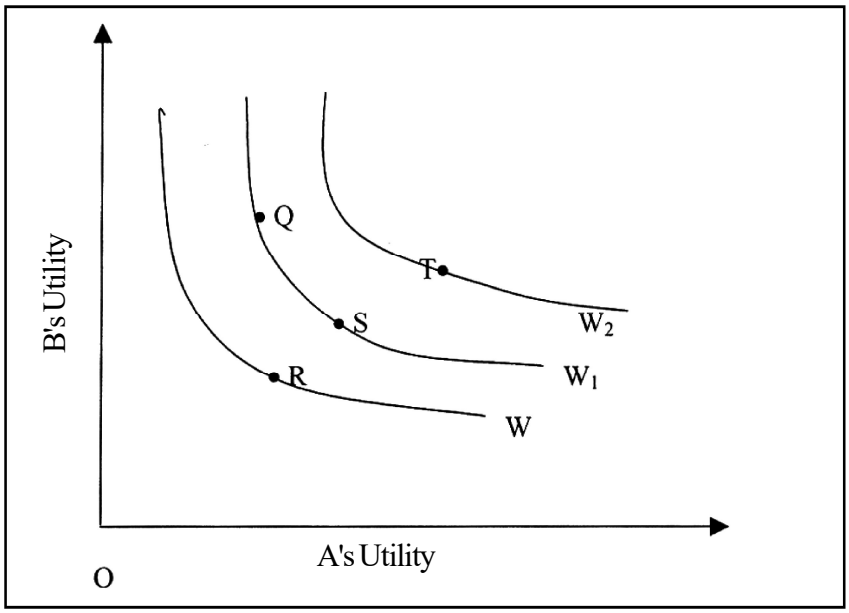


Diagram- 4

Maximization of Social Welfare:

Now we super impose the grand utility possibility curve on the social indifference curves representing social welfare function to find a unique optimum position of social welfare. In the diagram W, W_1 and W_2 the social indifference curves representing the social welfare function have been drawn along with the grand possibility curve VV/. Social indifference curve W_1 is tangent to the grand utility possibility curve VV/ at point Q. Thus point Q represents the maximum possible social welfare given the factor endowments, state of technology and preference scales of the individuals. Point Q is known as the point of constrained bliss, since, given the constraints regarding factor endowments and the state of technology, Q is the highest possible state of social welfare, which the society can attain. Point L is on the lower indifference curve W represents a lower level of social welfare; where as point C, on the W_2 curve is beyond the grand utility possibility curve VV/ of the society. Thus point Q represents maximum social welfare.

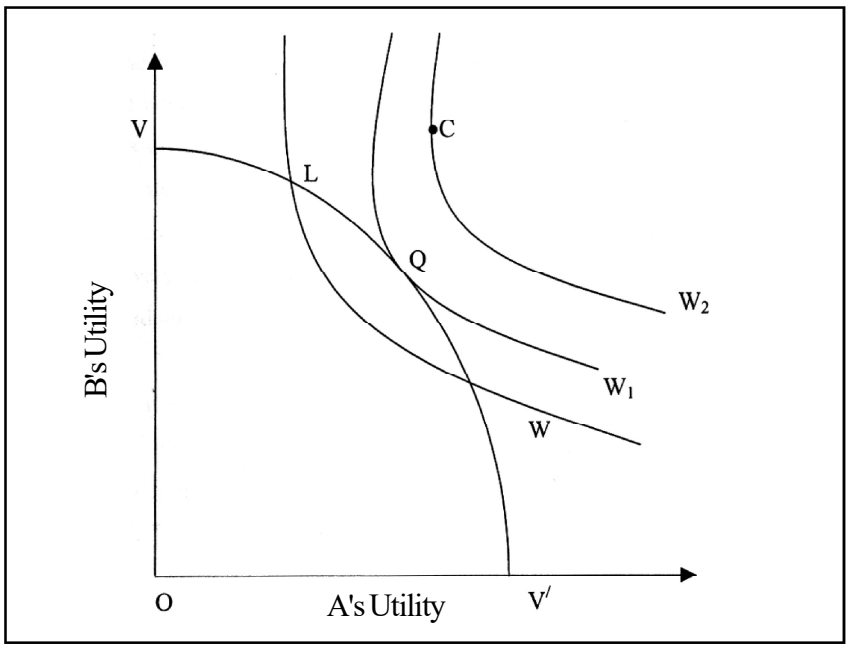


Diagram - 5

2.6 ARROW'S THEOREM OF SOCIAL CHOICE:

Professor Bergson and Samuelson made significant contribution to welfare economics by introducing explicit value judgments in the form of social welfare function. But they did not deal with the question as to how to get these value judgments or what this value judgment could be for constructing a social welfare function. Professor K.J. Arrow explored this question in his path breaking work "Social Choice and Individual Values". Arrow pointed out that the construction of social welfare function, which reflects the preferences of all individuals constituting a society, is an impossible task. In other words, it is very difficult to set up reasonable democratic procedures for the aggregation of individual preferences into a social preference for making a social welfare function.

There are many ways in which social choice can be derived. Choice may be made by a dictation or through custom of tradition or by some spiritual or religious head as was done in a traditional society or by individuals comprising a society through voting. All ways of translating individual choices into social choices may not be equally desirable or acceptable. The problem of social choice is easiest in a dictatorial rule in which the dictator makes all the social choices and all the individuals of the society are compelled to accept it. Similarly, in a traditional society various religious and spiritual rules and customs make the problem of social choice very easy. No individual can disregard the social choice made by a religious and

spiritual head. But the problem of making social choice based on individual orderings becomes difficult in a democratic society in which every individual is free to have his own individual ordering of various social states. From the point of view of desirability and acceptability of social choices, Arrow suggests five minimum conditions which social welfare must satisfy in order to reflect the preferences of individuals comprising a society.

(i) **Transitivity or consistency:** Social choices must be consistent or transitive. Transitivity of social choices implies that if an alternative A is socially preferred to alternative B and alternative B is preferred to alternative C, then the alternative C will not be preferred to alternative A.

(ii) **Responsiveness to individual preferences:** Social choice must be directly related to individual preferences. It implies that social choices must change in the same direction as the individual choices.

(iii) **Non-imposition:** Social choices must not be imposed by any one outside the community. It must be derived from individual preferences. For instance, if the majority individual does not prefer A to B, then society should not follow it.

(iv) **Non-dictatorship:** Social choices must not be dictated by any one individual in the community.

(v) **Independence of irrelevant alternatives:** Social choices must be independent of irrelevant alternatives. In other words, if any one alternative is excluded, it will not affect the ranking of other alternatives.

The above five conditions of Arrow reflect his value judgments and they seem to be quite reasonable set of conditions for making social choices in a free democratic society. However, Arrow demonstrated that it is not possible to satisfy all these five conditions and obtain a transitive social choice for each set of individual preferences without violating at least one condition. In other words, social choice is inconsistent or undemocratic because no voting system allows these five conditions to be satisfied. This has come to be known as the Arrow's impossibility theorem.

To illustrate Arrow's impossibility theorem we construct the following table

Individuals	Alternative Social Status		
	X	Y	Z
A	3	2	1
B	1	3	2
C	2	1	3

In the above table, there are three individuals, A, B and C who constitute the society have been shown to have voted for three alternative social states X, Y and Z, by writing 3 against the most preferred alternative, 2 for the next preferred alternative and 1 for the least preferred alternative.

The table shows that each individual has consistent preferences. A prefers X to Y, Y to Z and hence X to Z. B prefers Y to Z, Z to X and hence Y to X. C prefers Z to X , X to Y and hence Z to Y. It is clear that two individuals A and B prefer Y to Z and also two individuals A and C prefer Z to X. Thus the majority (two of the three individuals) prefers X to Y and also Y to Z and therefore Z to X. But majority prefers Z to M. Thus, we see that majority rule leads to inconsistent social choices because on the one hand X has been preferred to Z by the majority and on the other hand Z has also been preferred to X by the majority which is quite contradictory or inconsistent.

2.7 AMARTYA SEN'S CONCEPT ON POVERTY AND WELFARE:

Professor Amartya Kumar Sen has made some notable contributions to research on fundamental problems in welfare economics. His contributions range from "the axiomatic theory of social choice, over definition of poverty indices, to empirical study of famine".

Sen in his theory of social choice is concerned mainly with the link between individual values and collective choice. In early 1950's, Arrow in his "Social choice and individual values" shows that a voting system is necessary in a democracy to strike at the right social function. But the problem is that the majority rule leads to inconsistent and contradictory results. This situation leads to the impossibility theorem and welfare economics is at the dead end. Then Amartya Sen has found the way out. In his book "Collective Choice and Social Welfare", Sen pointed out that even in a democracy, by applying majority rule, we could promote social welfare. Sen firmly believed in individual right and to him, a prerequisite for collective decision-making rule is that it should be non-dictatorial.

In fact, Sen redefines welfare economics in terms of the concepts used by moral philosopher John Rawls. The minimax of Rawlsian welfare function is that social welfare of an allocation depends only on the welfare of the worst-off agent - the person with the minimal utility.

According to Sen, what create welfare is not goods as such, but the activity for which they are acquired. He believes that income is significant because of the actual opportunities it creates. But the actual opportunities (Sen calls them capabilities) also depend upon a number of factors such as basic education and health. These factors also should be taken into account while measuring welfare.

Professor Sen has also defined poverty index. As we know that a common measure of poverty in a society is the share of population below a tolerable standard of living. But the poverty line approach ignored the degree of poverty among the poor. Sen developed a new index to measure the inequality of people below the poverty line. Sen's poverty index is,

$$P = H [I + (1 - I) G]$$

Where, P is the poverty index, H is the head count of the number of people below the poverty line, I is the income gap which measures the additional income that would be needed to bring all the poor upto the level of the poverty line and G is the Gini co-efficient, the inequality measure of the distribution of income among the poor.

CHECK YOUR PROGRESS - 2.1

1. State the Compensation principle of Kaldor-Hicks.

Ans. ……..……

.....

.....

.....

.....

.....

2. Define Sen's poverty index.

Ans. ……..……

.....

.....

.....

.....

.....

2.8 Let Us Sum Up

In the second unit of the block we have discussed about the compensation principle of Kaldor-Hicks, Scitovosky's paradox, Scitovosky's double criterion. Next we have drawn our attention at discussing about value judgment; Bergon-Samuelson social welfare function; Arrow's impossibility theorem and finally Amartya sen's concept of poverty and welfare.

Key Words

Compensation principle of Kaldor-Hicks: According to this principle if a certain change in economic policy makes some people better off and others worse off then that change will increase social welfare if those who gain compensate the losers and still be better off than before.

Value Judgement : Value judgement refers to an opinion about the relative merits of two or more states of the economy, which cannot be tested empirically. For example, any policy change can affect two individual differently.

Suggested Readings:

1. Koutsoyiannis, A.; *Modern Microeconomics,* Macmillan
2. Ahuja, H.L; *Advanced Economic Theory*, S. Chand
3. Salvatore, D.; *Microeconomics : Theory and Application*, Oxford.

Terminal Questions:

1. Discuss the compensation principle of Kaldor-Hicks. What are Scitovosky's Paradox and his Double Criterion? Discuss.
2. What is Social Welfare Function? Discuss the social welfare function concept of Bergson-Samuelson.
3. Discuss critically the Arrow's theory of social choice and the impossibility theorem.

Answers to check your progress:

1. According to this principle if a certain change in economic policy makes some people better off and others worse off then that change will increase social welfare if those who gain compensate the losers and still be better off than before.

2. Sen's poverty index is given as:

$$P = H [I + (1 - I) G]$$

Where, P is the poverty index, H is the head count of the number of people below the poverty line, I is the income gap which measures the additional income that would be needed to bring all the poor upto the level of the poverty line and G is the Gini co-efficient, the inequality measure of the distribution of income among the poor.