M.A. ECONOMICS – I YEAR

DKN11 : MICRO ECONOMICS ANALYSIS

SYLLABUS

Unit 1: Demand Analysis

Meaning and Definition, Nature and scope of Micro Economics - Law of Demand - and Elasticity of Demand - Utility Analysis of Demand - Indifference Curve Analysis of Demand - Revealed Preference Theory - Recent Developments in Demand Theory.

Unit II : Theory of Production and Costs

Theory of Production - The law of variable proportions - The law of returns to scale.- soquants approach on producer's equilibrium - The law of Diminishing Marginal Rate of Technical Substitution - The Cobb - Douglas Production function - C.E.S. and VES production functions - Traditional and Modern Theories of costs

Unit III : Price and Output Determination

Equilibrium of the firm and industry under perfect competition - Price and output determination under monopoly - Price and output determination under monopolistic competition - meaning of concepts Monopsony, Bilateral monopoly, oligopoly and duopoly

Unit IV : Alternative Theories of Firm and Distribution

Bain's limit Pricing theory of recent development - Prof. Baurnol's Theory of Sales - Revenue maximization - Neo - classical Theory of Marginal Productivity theory Euler's Theorem - Factors Pricing under different market conditknis - **Modem** Theories of Rent, Wages, Interest and Profit (including Shakle's Theory of profit) .

Micro Theories of distribution.

a) Richardian b) Marxian c) Kalecki and d) Kaldor

Unit V : Welfare Economics

Nature of Welfare Economics - Pigovian of Welfare Economics and Externalites Pareto's Optimum Social Welfare Criterion - Kaldor - Hicks New Welfare compensations Principle - Scitovsky's Double Criterion of new welfare - Bergson's Social welfare Function - Arrow's impossibility Theorem

Reference Books

1. Koutsoyannis, A - "Modern Micro Economics", Macmillan Press, . London.
2. Dewett. K.K " - "Modern Economic Theory",S. Chand, New Delhi.
3. Ahuja . H. L - "Modem Economics", S. Chand , New Delhi.
4. Varian.H – "Modern Micro Economics Analysis" Macnnkan Press, London.
5. Jhingan, M. L - "Advanced Economic Theory" Brinda Publication - New Delhi.

UNIT : 1: DEMAND ANALYSIS

Meaning and Definition, Nature and scope of Micro Economics - Law of Demand - and Elasticity of Demand - Utility Analysis of Demand - Indifference Curve Analysis of Demand - Revealed Preference Theory - Recent Developments in Demand Theory.

Meaning:

The term "Micro Economics" is derived from the Greek word "Mikros" which means "Small". Therefore, Micro Economics is the study of small or individual units and not aggregates. In other words, it is the study of the economic action of individuals or small group of individuals. In brief, Micro Economics relates to the microscopic study of the economy. Micro Economics is also known as "Price Theory".

Definitions:

1. According to Prof. Boulding, Micro Economics is "the study of particular firms, particular households, individual prices, wages, incomes, individual industries, particular commodities".
2. To Maurice Dobb "Micro Economics is in fact, microscopic study of the economy"
3. According to Leftwitch, "Micro Economics is concerned with the economic activities of economic units as consumers, resource owners and business firms".

Merits or Advantages or Importance:

1. Micro Economics is helpful in understanding the working of a free-enterprise economy
2. Micro Economics provides the tools for evaluating the economic policies of the Governments.
3. Micro Economics deals with the economising of scarce resources with efficiency.
4. Micro Economics is also used to understand some of the problems of taxation and international trade.
5. Micro Economics can be used to examine the conditions of economic welfare.
6. Micro Economics provides a framework to analyse the various market structures, like Perfect Competition, Monopoly, Ologopoly, Duopoly etc.

Demerits or Disadvantages or Limitations

1. Micro Economics cannot give any idea of the functioning of the economy as a whole. What is true of small units need not be true of the economy as a whole.

2. Micro Economics does not give a correct picture of the working of the entire economy. Because, it explains the working of individual units.
3. The greatest defect of Micro Economics is that it assumes full employment. This assumption is an unrealistic one.

Scope of Economics

The Scope of a subject means the extent of a particular subject. In that aspect, Economics includes the following:

1. The Subject Matter of Economics
2. Economics is a Science or Art
3. Pure Science or Applied Science
4. Positive Science or Normative Science
5. Economic is a Social Science

Demand

Meaning of Demand

Before we study the Law of Demand, it is imperative to know the meaning, types and nature of demand. Demand in common parlance means the desire or want or need for a commodity. But in Economics, demand means something more than desire. That is, in Economics, demand means the desire backed up by a willingness to buy a commodity, at some price. In other words, demand for a commodity refers to quantity of the commodity which a consumer is willing to purchase at a particular price. Demand depends upon price. In brief, demand is backed up by the ability and the willingness to buy the commodity.

Definitions of demand

1. According to Stonier and Hague, "Demand in Economics means demand backed up by enough money to pay for the goods demanded".
2. According to Benham, "the demand for anything at a given price is the amount of it which will be bought per unit of time at that price".

Demand function

Demand depends upon price. This means demand for a commodity is a function of price. We represent demand function mathematically as,

$$D = f(p)$$

Types of Demand

Broadly speaking the demand is classified into Price Demand, income demand and Cross Demand. Let us see the meaning of each type

Price Demand

Price Demand means the change in the demand for a commodity due to change in its price alone

Income Demand

Income Demand means change in the demand for a commodity due to change in income of a consumer alone.

Cross Demand

Cross Demand means the change in the demand for a commodity, say 'x' due to a change in the price of other commodity, say 'y' other things remain the same.

Law of Demand

The Law of Demand was first stated by Augustin Cournot in 1838. Later it was refined and elaborated by Alfred Marshall.

Definitions

1. Marshall defines the Law of Demand as "the amount demanded increases with a fall in price and diminishes with a rise in price".
2. S.E. Thomas states the Law of Demand as "at any given time, the demand for a commodity or service at the prevailing price is greater than it would be at a higher price and less than it would be at a lower price".
3. To Samuelson, "the Law of Demand states that people will buy more at lower prices and buy less at higher prices, other things remaining the same".
4. In the words of Ferguson, "according to the Law of Demand, the quantity demanded varies inversely with price".

Explanation

The Law of Demand explains the relationship between the price of a commodity and the quantity demanded for it. This law states that quantity demanded of a commodity expands with a fall in price and contracts with a rise in price. In other words, a rise in price of a commodity is followed by a reduction of demand a fall in price is followed by contraction in demand. Therefore, the Law of Demand states that there is an inverse relationship between the price and the quantity demanded of a commodity.

Demand schedule

"Demand Schedule" is a tabular statement which shows the relationship between the price and the quantity demanded for a commodity.

The demand schedule is divided into two, namely, Individual Demand Schedule and Market Demand Schedule. The Individual Demand Schedule refers to the prices and the amount demanded for the commodity by an individual. Market Demand Schedule refers to the quantities of a given commodity which all the consumers will buy at all possible prices at a given moment of time.

The law of demand assumes that demand for a commodity is only the function of its price. The law of demand or the relationship between the price of a commodity and its amount demanded can be illustrated with an imaginary demand schedule of Table 1.1

Table 1.1 Mr. X's Demand Schedule for apples

Price in (Rs. Per kg)	Demand (in kgs)
1	5
2	4
3	3
4	2
5	1

Table 1.1 shows the amount of apples demanded at various prices by Mr. X A price increases from Re. 1 per kg to Rs. 2 per kg, the amount of apples purchased reduces from 5 kgs to 4 kgs. When price rises further from Rs. 2 to 3, Rs. 3 to 4 and Rs. 4 to 5, the amount demanded falls from 4 kgs to 3 kgs to 2 kgs and 1 kg respectively. This demand schedule is transformed into a demand curve in Fig 1.1. price is depicted on they axis and amount demanded on the Y axis. As price falls from Rs. 4 to Rs. 3, the quantity demanded increases from 2 kgs to 3 kgs.

The demand curve, normally sloped downwards from left to right on some cases, it is negatively sloped and is Iso-temporal in character.

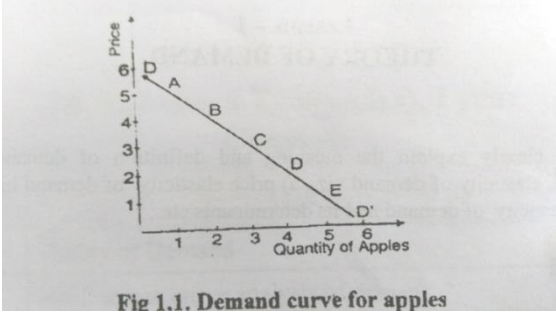


Fig 1.1. Demand curve for apples

From Individual Demand Curve to Market Demand Curve

Market demand is a horizontal summation of individual's demand. It means that the demand for all consumers in the market at a particular price is added to arrive at the market demand. The techniques of constructing the market demand schedule and the market demand curve are shown as follows:

The market demand schedule, (hypothetically) for apples is arrived at by assuming only three consumers A, S, and C in the market. The addition of the demand for these three consumers at each price gives the market demand schedule. For example, at price Rs.2, the market demand is 25 kgs (13+4+8). The figure shows the derivation of market demand from individual's demand. At price OP0 the demand fro A is OX_{A1} B is OX_B and C is OX_C. The addition of these three amounts ($OX_A + OX_B + OX_C = X_C$) gives the market demand as OXo. By a similar procedure, by adding the demand of various consumers in the market at various prices, the market demand schedule is drawn in the **Table 1.2** and explained in the **fig 1.2**

Table 1.2

Quantity Demanded of Apples				
Price (in Rs.)	**Consumer (A)**	**Consumer (B)**	**Consumer (C)**	**Consumer (D)**
1	15	5	10	30
2	13	4	8	25
3	12	2	6	20
4	10	1	4	15
5	8	0	2	10

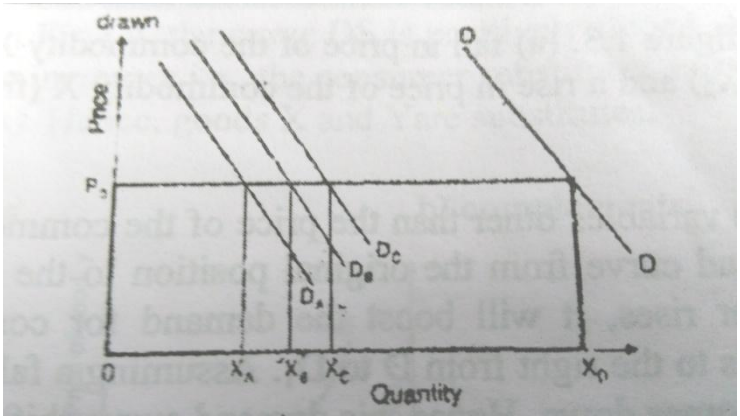


Fig.1.2 Derivation of market demand

Demand Function and Determinants of Demand

Demand is not only influenced by price but also by many other factors either simultaneously, or with a time lag i.e demand therefore is a multivariate relationship. The most important factors are consumer's income, prices of substitutes, prices of other commodities, tastes, income distribution, size of population, consumer's wealth, credit availability, government policy, past levels of demand and past levels of income

$D_X = f (Px, P_O, Y, Sp, W, T, G, D_{t-1}, D_{t-2}, D_{t-n}, Y_{t-1}, Y_{t-2}, Y_{t-n}\ldots)$

Where	Dx	= Demand for commodity
	Px	= Price of commodity x
	Po	= Price of other commodities
	Y	= income
	Sp	= Size of population
	G	= Government policy
	W	= Wealth
	T	= Taste
$D_{t-1}, D_{t-2}, D_{t-n}$		= Previous periods demand
$Y_{t-1}, Y_{t-2}, Y_{t-n}$		= Previous periods income

Shift in Demand and its Movement

The movement along a demand curve implies the changes in the quantity demanded of a commodity to the changes in its price alone. Demand expands when price falls and contracts when price rises.

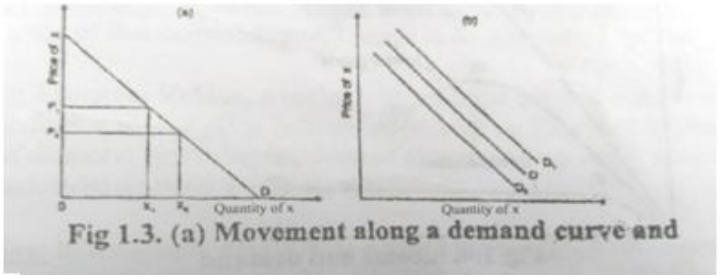


Fig 1.3. (a) Movement along a demand curve and

Fig.1.3

As shown in the figure 1.3 (a) fall in price of the commodity X (from P1 to P2) expands quantity demanded (from X1 to X2) and a rise in price of the commodity X (from P2 to P1) contracts the demand (from X2 to X1)

Any change in the variables other than the price of the commodity X1 like income, taste etc, leads to a shift in the demand curve from the original position to the right to left. For example, suppose income to the consumer rises, it will boost the demand for commodity X at each price. Therefore, the demand curve shifts to the right from D to D1. Assuming a fall in income, his capacity to buy the commodity at each price comes down. Hence, his demand curve shifts to the left form D to D2

Relation between Demand and income

The increase in income normally raises the demand for the commodity. These goods are called normal goods. In certain cases, increase income pushes the quantity demanded up to a certain level, then increase in income pushes the quantity demanded down. Goods of this nature are called inferior goods. In some other cases increases in income increases the demand for the commodity in the initial stages after that leaves the quantity demanded unaffected. All these three cases are depicted in **fig 1.3.** The curve A shows that as income increases the demand for the good also increases. Therefore, it depicts the case of normal goods. The curve B shows that the demand for the good moves positively with income up to a point 'm'. Beyond 'm', it turns negative signifying that the demand falls as income x increases. This is the case of inferior goods. The curve C depicts that as income increases the demand also increases but after point 'R' the quantity demanded remains constant as income increases. This is the case of necessities.

Relation between Demand and Price of other goods

The demand for a commodity and the prices of other goods are related in two ways.

1. A rise/fall in the price of other goods may increase/reduce the quantity demanded. If this happens they are substitutes
2. A rise/fall in the price of other goods may reduce/increase the quantity demanded. If this happens they are complements

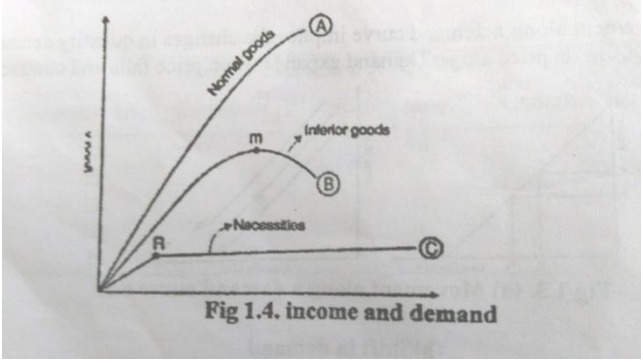


Fig.1.4

As shown in **fig 1.4** the curve DS is positively sloped. As the price of good Y increases the demand for good X also increases i.e., the consumer substitutes more of X in the place of costly Y (if price of X remains constant). Hence, goods X and Y are substitutes.

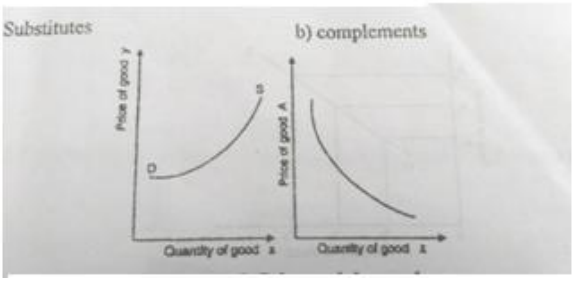


Fig.1.5

In the panel (b) as price falls, the demand for good Z rises i.e., a fall in price induces the demand for good Z (and A also) therefore goods A and Z are complements.

Effect of Taste on Demand

Taste is the introduction of new products initiated by producers. Alternatively, it gives changes in the lifestyle. A favourable change in taste and fashion for a commodity is so 'raises the demand for that commodity'. When the changes in taste and fashion are unfavourable to a commodity, then the demand for it declines.

Relation between size of population and Demand

The size of population is not subject to wide changes in a short-time span, increase in population proportionately increases the demand for food, clothing, shelter and many other necessities. The demand for most of the commodities in general is positively influenced by income and population.

Exception to the Law of Demand (Except Demand Curve)

There are some exception to the law of demand. They are as follows.

1. **Snob effect or Snob appeal**: When people want to be exclusive by purchasing selective items, then a rise in price of that commodity will increase its demand
2. **Veblan effect**: Thorestein Veblan, a social critic pointed out that sometimes people judge quality by price. Therefore, a higher price induces its purchases. Diamond is often cited as an example. As the price of diamond moves up the demand also moves up in the same direction. But this does not consider industrial demand for diamonds.
3. **Giffen paradox**: In Britain, Robert Giffen noticed the tendency of wage earners to increase their demand for bread as its price rose in the early 19th century. As shown in figure 1.6 as price increase from P0 to P1, the demand for bread also increases from

X0 to X1. However, it is not true that it happened entirely due to price rise. It is felt that wage earners brought more bread because they were unable to maintain their calorie intake. Since, the price of bread had increased, they substituted bread for mutton.

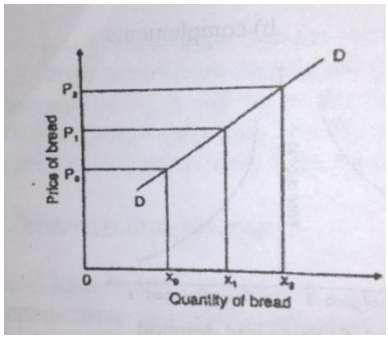


Fig 1.6

4. **Bandwagon effect:** individuals demand changes in the same direction as changes in others consumption, therefore sometime, a rise in the price of those commodities will not decrease the demand for them (Demonstration effect)
5. **Expectations effect:** When prices are falling, consumers may expect a further fall in price and wait for it. Thus, fall in price may fail to bring an expansion in demand. But these are occurrences of trade cycle and cannot be regarded as an exception to the demand law. (Price war between two local petrol sellers is cited as an example)
6. **Attachment to products:** Consumers are attached to some brand names. Even if their prices rise most of them continue their purchases.

Elasticity of Demand:

The concept 'elasticity' was borrowed from physics. The name elasticity of demand was given by Marshall though Cournot used it earlier as "degree of responsiveness". In general, the elasticity of demand is defined as the measure of the degree of responsiveness in quantity demanded to a change in price.

There are as many elasticities of demand as its determinants. The most important elasticities of demand are, 1. Price elasticity of demand. 2. Income elasticity of demand, and 3. Cross elasticity of demand.

Price Elasticity of Demand

Price elasticity of demand considers the responsiveness of changes in quantity demanded to a change in price. There are three methods of measuring price elasticity of demand

They are: (a) Point elasticity, (b) Are elasticity, and (c) Total outlay method

Point elasticity of demand is defined as the proportionate change in quantity demanded resulting from a very small change in price. It is written as

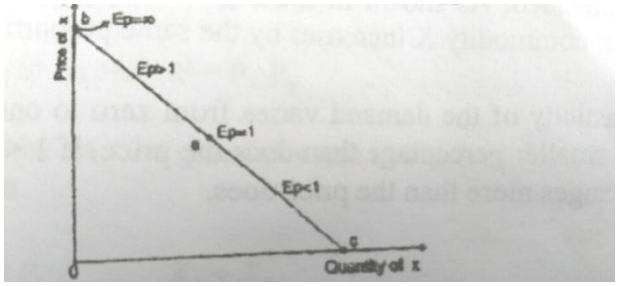


Fig.1.7 (a)

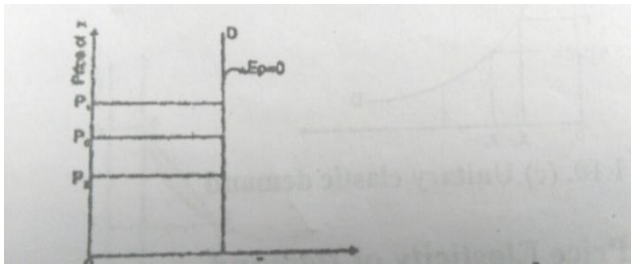


Fig.1.7(b)

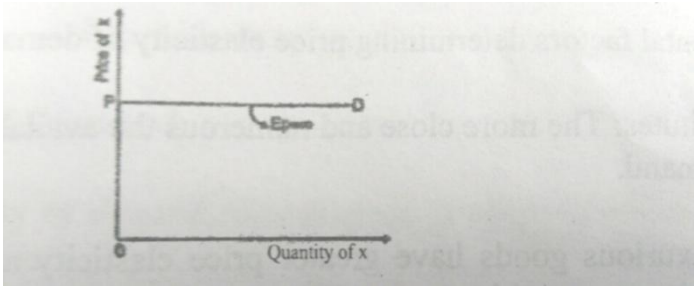


Fig. 1.7(c)

$$E_p = \frac{\text{Percentage change in quantity demanded}}{\text{Percentage change in price}}$$

$$\text{symbolically, } EP = \frac{\frac{\Delta q}{q} \times 100}{\frac{\Delta}{q} \times 100}$$

$$= \frac{\Delta q}{q} \quad \frac{x \Delta p}{p}$$

$$.EP = \frac{\Delta q}{p} \times \frac{p}{q}$$

If Ep = 0: The demand is said to be perfectly elastic as shown in the **Fig.1.7 (a).** At price OP the demand is infinite. It implies that the purchasers are willing to buy whatever is offered at OP and none at a price even slightly higher than OP, as shown in the Fig. 1.7

If ep = 1, it is called unitary elastic demand. The changes in price bring with it (equal) proportionate changes in quantity demanded. As shown in the **Fig.1.7 (c)** when price falls from P to P_1 , the quantity demanded of commodity X increases by the same proportion from X_1 to X_2

If 0 < Ep < 1: The elasticity of the demand varies from zero to one i.e., it is less elastic. The quantity demanded changes by a smaller percentage than does the price. If 1 < Ep <: The demand is elastic, i.e., the quantity demanded changes more than the price does.

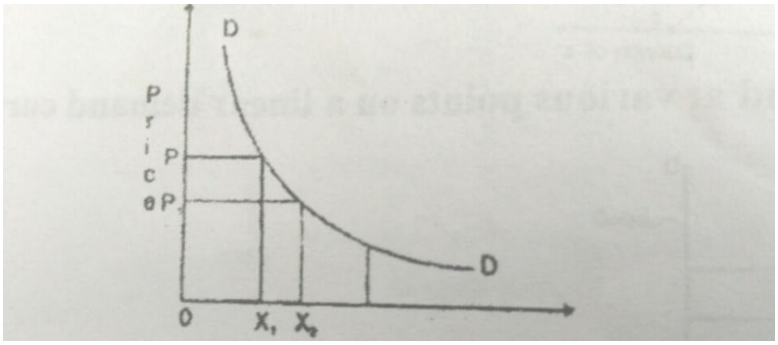


Fig 1.8

Determinants of Price Elasticity of Demand

The fundamental factors determining price elasticity of demand are

1. Availability of substitutes. The more close and numerous the availability of substitutes the more is price elasticity of demand.
2. Nature of goods: Luxurious goods have greater price elasticity and necessities have less price elasticity of demand since necessities cannot be reduced/increases mud for changes in price
3. Time period: The longer the time period the greater the elasticity of demand and the smaller the time horizon the smaller the elasticity of demand

4. Number of alternative uses: The more the alternative uses for a commodity greater the Ep for it since people tend to maximise utilisation
5. Proportion of income spent: The traditional argument is that if a commodity accounts for a greater portion of income spent, then its elasticity will be greater

Arc Elasticity Method

The point elasticity method can measure only the change in demand for small or minute changes in price. It becomes ineffective in cases of large changes in price. Arc elasticity is an average reaction or the average responsiveness.

$$\text{We know that, } Ep = \frac{\Delta q}{q} \times \frac{\Delta p}{p}$$

In the figure the price changes from P0 to P1 and the quantity demanded from X0 to X1. Therefore,

$$\Delta q = X_1 \text{ to } X_0 \text{ and } \Delta P = P_1 - P_0$$

Since there are two quantities and two prices, the price and the quantity are arrived at by the following procedure

$$P = \frac{P_1 + P_0}{2} \text{ and } q = \frac{X_1 + X_0}{2}$$

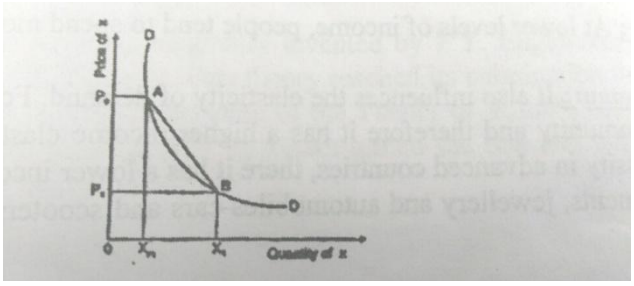


Fig.1.9

Income Elasticity of Demand

Income elasticity of demand is a measure of responsiveness in quantity demanded to a change in income. It is defined as

$$Ey = \frac{\text{Percentage change in quantity demanded}}{\text{Percentage change in income}}$$

Income elasticity is positive for most of the goods i.e., and increase in income increases the quantity demanded and vice versa. Some writers define goods as normal, superior and inferior on the basis of income elasticity of demand.

If 1 < Ey < the good is superior

If 0 < Ey < the good is normal

If Ey < 0 the good is inferior

The curve OY shows that the demand for good increases as income increases from o to point 'A'. Therefore, income elasticity of demand is positive in this range. From point 'P to 'B' income does not have any effect on demand for Y the income elasticity of demand is zero or perfectly inelastic. After Yb level of income the demand for good X decreases showing negative elasticity of demand. Therefore, the same good may be superior or normal (inferior depending upon the level of income). For example, a bicycle may be superior go to a peon, necessity to a clerk and inferior to a manager of a firm.

Determinants of income elasticity if Demand

1. **Level of income**: At lower levels of income, people tend to spend more of the income on food.
2. **Level of development**: It also influences the elasticity of demand. For example T.V. is a luxury in a low income community and therefore it has a higher income elasticity of demand. For luxuries like ornaments, jewellery and automobiles-cars and scooters, income elasticity would be high.

Cross Elasticity of Demand

Cross elasticity of demand measures the responsiveness of the change in quantity demanded of one commodity due to a change in the price of another commodity. Let A and B be two commodities, PA and PB be their prices, the cross elasticity of demand (Ec) can be defined as

$$Ey = \frac{\text{Percentage change in quantity demanded A}}{\text{Percentage change in income B}}$$

Summary

From the above text summarized the meaning and definition of demand, Law of demand determinants of demand, elasticity of demand viz. a) price elasticity of demand b) income elasticity of demand and c) cross elasticity of demand and its determinants etc.

INDIFFERENCE CURVE ANALYSIS OF DEMAND

Introduction

The ordinalists believe that the measurement of subjective utility on an absolute scale is neither possible nor necessary. They maintain that "all consumer behaviour can be described in terms of preferences or rankings in which the consumer need only state which of the two collections of good he prefers without reporting on the magnitude of any numerical index of preference". They constructed an 'indifference curve' technique to deal with "optimising satisfaction".

Indifference curves were originally invented by F.Y. Edgeworth and later developed by Vilferdo Pareto, Johnson and E.E. Sluktsky. This theory reached its culmination in the hands of J.R. Hicks and R.G.D. Allen. The article "A Reconsideration of the Theory of Value" (1934) in Economica and the successive famous work "Value and Capital" by Hicks turned the tide and tempo of modern economic theory.

Assumptions

1. **Rationality:** The consumer acts in a rational manner and he aims at maximization of his utility subject to the constraints imposed by his given income
2. **Non-**satiety: Consumer is not over supplied with either good that is he prefers to have more of commodity X or Y.
3. **Ordinal utility:** consumers need not know the absolute amount of satisfactions as maintained by cardinalists. It is sufficient to say that a bundle of good gives more or less or equal satisfaction compared to another bundle. i.e., they are required to order the ranks of different alternative combinations.
4. **Diminishing marginal rate of substitution:** Marginal rate of substitution has been defined as the rate at which one good is substituted for another.
5. **Consistency and transitivity:** When two combinations A and B are available if the consumer prefers A to B in no other circumstance, B will be preferred to A.

Construction of indifference Curve

Imagine a household starting with combination 'A' consisting of 20 units of clothing and 4 units of food. This household will definitely prefer a combination 'N' containing 50 units of clothing and 10 units of food, since it involves more of both goods. Similarly, the household will not prefer a combination consisting 5 units of clothing and 2 units of food. But there may be some combinations which are equally preferable to A or indifferent to A. suppose a combination 'B' involving 14 units of 'cloth' and 6 units of food is equally

preferable, then it can be said that the household is indifferent between A and B. similarly, some other combinations indifferent to A can be found and are listed in Table below

Table 1.3 Marginal Rate of Substitution

Combination	Food (X)	Clothing (Y)	Y	X	MRS xy AY/AX
A	4	20	-	-	-
B	6	14	6	2	3
C	8	10	4	2	2
D	10	8	2	2	1
E	12	7	1	2	½
F	14	7	0	2	0

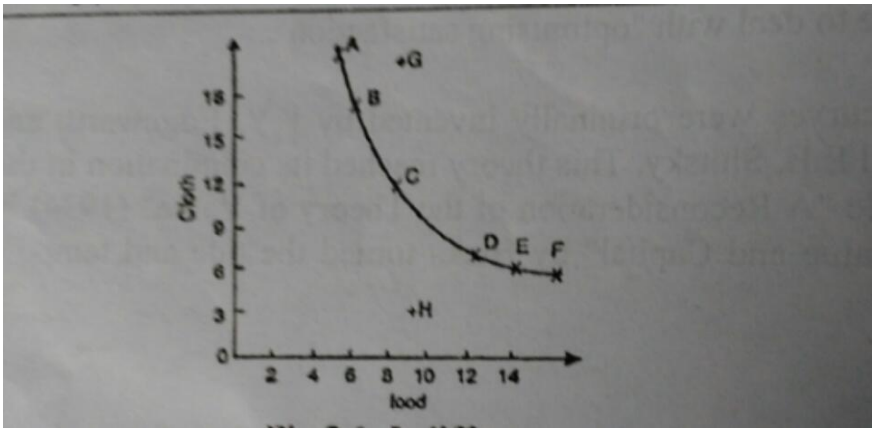


Fig.1.10

If these combinations are represented in a figure, we obtain the indifference curve. All the combinations A,B,C,D and E on the indifference curve are equally preferable i.e., an indifference curve is a no difference curve. It is a locus of combinations of two commodities which gives the consumer the same level of satisfaction. Indifference curves assume continuity i.e., goods are divisible. Therefore, there are some combinations in between A and B or B and C. they are also indifferent to A. Any combination above IC such as G is a preferable combination over B since G contains more of clothing with the same food as B. By transitivity all other combinations on IC become inferior to G. Similarly any combination below 'B' such as H is not preferred to B because it contains less cloth. By transitivity all combinations on IC becomes preferable to H.

Properties of indifference curve

Property 1 – Indifference curve is negatively sloped: It implies that the quantity of one commodity must decrease if the other commodity is increased to keep the level of satisfaction

intact. The slope of an indifference curve at any point is the ratio between the good replaced (Y) divided by the good added (X). The negative slope of the indifference curve as the marginal rate of substitution between the two commodities X and Y.

$$\text{Slope of IC} \quad \frac{\triangle Y}{\triangle X} = MRS_{xy}$$

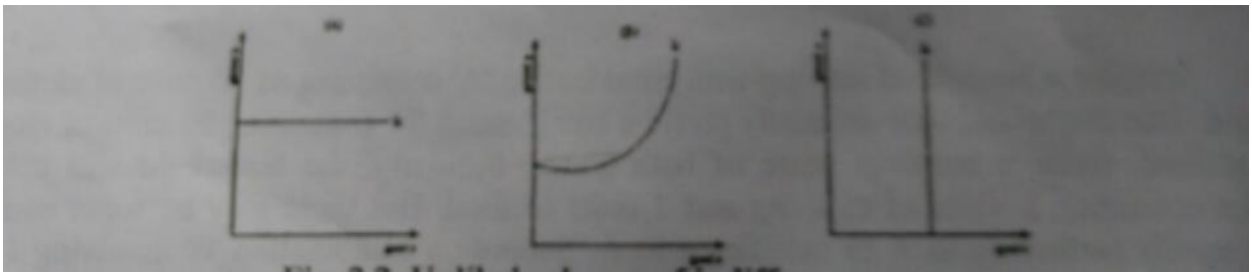


Fig.1.11

Marginal Rate of Substitution is defined by Hicks as "Marginal rate of substitution of X for Y as the quantity of Y which would just compensate the consumer for the loss of the marginal unit X". Since the slope of indifference curve is negative, it cannot be upward sloping or a vertical straight line or a horizontal straight line as shown in Fig.1.11

Property 2. Indifference Curves are convex to the origin: (Diminishing Marginal Rate of Substitution). The assumption of diminishing marginal rate of substitution dismisses a straight line negative shape, concave shape and also the types depicted above as likely shapes of indifference curve.

In Fig.1.11(a) for each unit of increment in X, a constant amount of Y is given up (YV1 = Y1 Y2). Therefore, the MRSxy remains constant. It is possible only if the two Goods are perfect substitutes. It means that consumer can replace the entire X with Y fully or Y with X fully.

In Fig.1.11 (b) the indifference curve is 'L'shaped. No substitution is possible in such case. It is a case of complements. In Fig.1.11 (c)for each additional increment in X more and more of Y is given up i.e., Y is increasing for a given X. Therefore, the MRSxy will increase. It is another abnormal case of indifference curve.

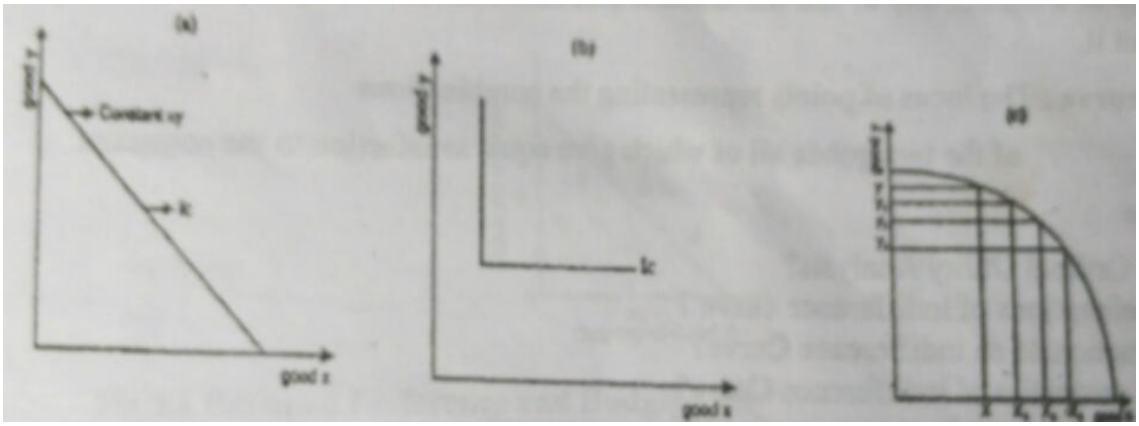


Fig.1.12

Property 3 – Higher indifference curve denotes higher level of satisfaction:

Thus, a higher indifference curve denotes a higher level of satisfaction, an a lower indifference curve denotes a lower level of satisfaction.

Property 4 – Indifference curves do not intersect or touch each other:

If indifference curves intersect with one another they will show two levels of satisfaction which is an incompatibility by definition. It also violates the assumptions of consistency and transitivity.

Summary

The text summarized indifference curve analysis has abandoned the concept of cardinal utility and it was invented by Edgeworth and R.G.D. Allen later J.R.Hicks. An indifference schedule explained a list of combination of to commodities being arranged that a consumer is indifferent to the combinations preferring nones of them to any to the others. It explained assumptions, properties of indifference curve and applications of indifference curves. Finally, it critically analysed indifference curves.

Revealed Preference Theory

Introduction:

Marshallian Cardinal Utility Approach and Hicksian Cardinal Utility Approach explained the consumer behaviour on the basis of introspection i.e., psychological explanation of consumer's demand. But these two approaches may not be applicable in practical life. Hence, the Noble Prize winner Prof. Paul A. Samuelson of Americal introduced the consumer's demand on the basis of actual behaviour of the consumer in 1938. Therefore, Revealed Preference Theory is behaviouristic explanation of consumer's demand.

Assumptions:

Revealed Preference Theory is based on the following assumptions:

1. Consumer's choice for a combination reveals his preference
2. The tastes of the consumer do not change
3. The consumer chooses only one combination of two goods at a given price – income line – strong ordering.
4. The consumer prefers to have a combination of more goods to less goods
5. This theory also assumes that the consumer is consistent in his behaviour (i.e., strong ordering). This means that if a consumer chooses one combination say x rather than y in one situation, he cannot prefer y to x in another situation, when both x and y are present.
6. This theory is based on the assumption of transitivity. It states that, if the different combinations of 'n' goods say x,y,z, etc., are arranged in order such that x is revealed preferred to y and y is preferred to z. But z must never be revealed to be preferred to x. thus, the ordering has always to be 'uni-directional' and never 'circular'
7. The income-elasticity of demand of the consumer must always be positive.

Demand theorem and Revealed Preference Hypothesis

Prof. Samuelson states his "Fundamental Theorem of Consumption Theory" on the basis of the above assumptions. His theorem is also known as "Demand Theorem". His Theorem is as follows:

"Any good (simple or composite) that is known always to increase in demand when money income alone rises must definitely shrink in demand when its price alone rises".

This means that there is inverse relationship between price and quantity demanded (i.e., price elasticity of demand is negative) assuming that income elasticity of demand is positive.

Figure1.13 illustrates the Demand Theorem on the basis of Revealed Preference Hypothesis.

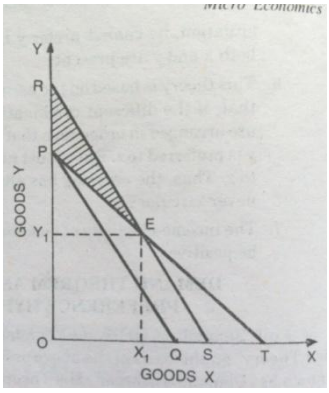


Fig.1.13

Demand Theorem when price of one commodity rises

We can explain this Theorem with an example. Suppose a consumer spends his entire income on the purchase of only two goods x and y. Further, suppose that this income in terms of good x in OT and in terms of y is OP. Hence PT is the original price-income line of the consumer. All the combinations of goods x and y lying within or on the triangle OPT are available to the consumer, from which he can choose any combination. Suppose the consumer chooses the combination E lying on PT. This means the consumer is revealed to be preferred this combination E (OX_1 +OY_1) to all other combinations lie in or on the triangle OPT.

Suppose the price of good x rises while the price of y remains the same. With the rise in price of x, the new price-income line will shift to PQ. Now, we want to know what is the effect of this rise in price of good x on its quantity demanded, assuming that demand varies directly with income i.e., income elasticity of demand is positive. It is obvious that in Figure -----that combination E is not available to the consumer in the price-income situation PQ. The consumer will buy less quantity of x i.e., he will buy less than OX_1 of good x, because of the rise in its price. Let us compensate the consumer for the loss in his real income as a result of rise in price of x, by granting him PR amount of money income in terms of the good y, so that he can buy the same combination E even when the price of good x rises. Draw a line RS, the new price-income line of the consumer which is parallel to PQ so that it passes through E. This is called by Samuelson as the "Over Compensation Effect", which is called as "Cost – Difference" by Prof. J.R. Hicks. This over compensation of the consumer leads to the formation of ORS as the new triangle of his choice.

Recent Developments in Demand Theory

A. The Pragmatic Approach to Demand Analysis:

There has been an increasing awareness that although the various approaches to utility are theoretically impressive, there is very little an applied economist can use to explain the complexity of the real world.

They accepted the fundamental 'law of demand' on trust, and formulated demand functions directly on the basis of market data without reference to the theory of utility and the behaviour of the individual consumer. Demand is expressed as a multivariate function, and is estimated with various econometric methods.

Such demand functions refer obviously to the market behaviour of the consumers, that is, to the behaviour of all consumers as a group, and not to the behaviour of single

individuals. Furthermore, in most cases the demand functions refer to a group of commodities, e.g. demand for food, demand for consumer durables, etc.

Serious difficulties arise in estimating demand functions. The aggregation of demand over individuals and over commodities makes the use of index numbers inevitable, but the problems associated with such indexes are numerous. Furthermore, there are various other estimation problems which impair the reliability of the statistically-estimated demand functions.

The most important of these difficulties arise from the simultaneous change of all the determinants, which makes it extremely difficult to assess the influence of each individual factor separately. However, there has been a continuous improvement in the econometric techniques and currently demand functions are fairly easy to estimate statistically.

The constant-elasticity demand function:

The most commonly used form of demand function in applied research has been the 'constant-elasticity' type:

$Q_x = b_0 - P^{b}_{x}{}^{1}. P^{b}_{0}{}^{2}. Y^{b3}. e^{b4t}$

Where Q_x = quantity demanded of commodity x

P_x = price of x

P_0 = prices of other commodities

Y = consumers' aggregate income

e^{b4t} = a trend factor for 'tastes' (e = base of natural logarithms)

b_1 = price elasticity of demand

b_2 = cross-elasticity of demand

b_3 = income elasticity of demand

The term 'constant elasticity demand function' is due to the fact that in this form the coefficients b_1, b_2, b_3 are elasticity's of demand which are assumed to remain constant.

Usually the followers of the pragmatic approach, although not adhering to utility functions, express the demand function in such a way as to incorporate the assumption of 'no money illusion' postulated by the traditional theory of the consumer. In technical jargon they express the demand as a homogeneous function of degree zero. This has been (most commonly) effected by introducing real income and relative prices in the function, that is

$$Q_x = b_0\left(\frac{P_x}{P}\right)^{b_1} \cdot \left(\frac{P_0}{P}\right)^{b_2} \cdot \left(\frac{Y}{P}\right)^{b_3}$$

where P is a general price index. In this formulation it is obvious that if prices and income change by the same proportion, for example by k per cent, the quantity demanded of x will not change, because k will appear in both the numerator and the denominator of the relative prices and real income, and hence will cancel out. The new quantity demanded will be the same as the initial one: there is no money illusion in the behaviour of the consumer.

Dynamic versions of demand functions: Distributed-lag models of demand:

A recent development in demand studies is the expression of demand functions in dynamic form.

Dynamic demand functions include lagged values of the quantity demanded and of income as separate variables influencing the demand in any particular period. Dynamisation of the demand functions expresses the generally accepted idea that current purchasing decisions are influenced by past behaviour.

To express the idea that current decisions are influenced by past behaviour we must postulate a particular type of relationship between the past and the present. The most common assumption in this respect is that current behaviour depends on past levels of income and past levels of demand. If the commodity is a durable past purchases constitute a 'stock' of this commodity which clearly affects the current (and future) purchases of such durable.

If the commodity is non-durable (for example, tobacco, food, etc.), past purchases reflect a habit which is acquired by buying and consuming the commodity in the past, so that the level of purchases in previous periods influences the current (and future) patterns of demand. Incorporating the influence of past decisions and experiences in the demand function is a way for rendering it dynamic.

Another usual assumption concerning the way in which past behaviour affects the present is that the more recent of past levels of income or demand have a greater influence on present consumption patterns than the more remote ones (for example, we are more influenced by our income in the last year than by the income we earned five or ten years ago).

Models (functions), including lagged values of demand, of income (or of other variables) are called distributed-lag models'. In general form a distributed-lag model may be expressed as

$Q_{x(t)} = f\{P_{x(t)}, P_{x(t-1)} \ldots Q_{x(t-1)}, Qx_{(t-2)} \ldots\ldots Y_{(t)}, Y_{(t-1)} \ldots..\}$

The number of lags depends on the particular relationship being studied. The necessity of a dynamic approach has long been recognized for the study of the demand of certain commodities (consumer durables).

Stocks S, however, cannot be measured:

(i) The stock of durables is composed of heterogeneous items of various ages-the electrical equipment we have is not of the same age, some items may be very old and need scrapping and replacing, some others are new. Their heterogeneity also makes direct measurement difficult. What we ideally want for stocks is the sum of depreciated inventories of durables; but the appropriate depreciation rates are not known.

(ii) The 'stock of habits' is a psychological variable and cannot be quantified.

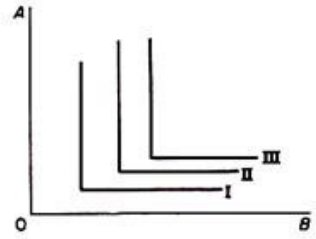
However, we can eliminate algebraically stocks, S_t, from the demand function and replace it with other measurable variables by making some 'reasonable' assumptions.

B. Linear Expenditure Systems:

These are models which deal with groups of commodities rather than individual commodities. Such groups, when added, yield total consumer expenditure. Linear expenditure systems are thus of great interest in aggregate econometric models, where they provide desirable disaggregation of the consumption function. One of the earliest linear expenditure models was suggested by R. Stone (Economic Journal, 1954).

The linear expenditure systems (LES) are usually formulated on the basis of a utility function, from which demand functions are derived in the normal way (by maximization of the utility function subject to a budget constraint). In this respect the approach of LES is the same as that of models based on indifference curves.

However, LES differ in that they are applied to 'groups of commodities' between which no substitution is possible, while the indifference-curves approach is basically designed for handling commodities which are substitutes. The very notion of an indifference curve is the substitutability of the commodities concerned. Actually the indifference map of a LES would appear as in figure 1.14, implying the non-substitutability of groups of commodities.



Indifference map for complementary goods

Fig.1.14

The utility function is additive, that is, total utility (U) is the sum of the utilities derived from the various groups of commodities.

For example, assume that all the commodities bought by the consumers are grouped in five categories:

A Food and beverages

B Clothing

C Consumers' durables

D Household-operation expenses

E Services (transport, entertainment, etc.).

The total utility is

$U = \sum U_i$

or

$U = U_{(A)} + U_{(B)} + U_{(C)} + U_{(D)} + U_{(E)}$

Additively implies that the utilities of the various groups are independent, that is, that there is no possibility of substitution (or complementarity) between the groups A, B, C, D and E. In linear expenditure systems the commodities bought by the consumers are grouped in broad categories, so as to be compatible with the additively postulate of the utility function. Thus each group must include all substitutes, and complements. In this way substitution between groups is ruled out, but substitution can occur within each group.

The consumers buy some minimum quantity from each group, irrespective of prices. The minimum quantities are called 'subsistence quantities' because they are the minimum requirements for keeping the consumer alive. The income left (after the expenditure on the minimum quantities is covered) is allocated among the various groups on the basis of prices.

The income of the consumer is, therefore, split into two parts the 'subsistence income', which is spent for the acquisition of the minimum quantities of the various commodities, and the 'supernumerary income', the income left after the minimum expenditures are covered.

UNIT II : THEORY OF PRODUCTION AND COSTS

Theory of Production - The law of variable proportions - The law of returns to scale.- soquants approach on producer's equilibrium - The law of Diminishing Marginal Rate of Technical Substitution - The Cobb - Douglas Production function - C.E.S. and VES production functions - Traditional and Modern Theories of costs

THEORY OF PRODUCTION

In economics, an effort to explain the principles by which a business firm decides how much of each commodity that it sells (its "outputs" or "products") it will produce, and how much of each kind of labour, raw material, fixed capital good, etc., that it employs (its "inputs" or "factors of production") it will use. The theory involves some of the most fundamental principles of economics. These include the relationship between the prices of commodities and the prices (or wages or rents) of the productive factors used to produce them and also the relationships between the prices of commodities and productive factors. On the one hand, the quantities of these commodities and productive factors that is produced or used, on the other.

The various decisions a business enterprise makes about its productive activities can be classified into three layers of increasing complexity. The first layer includes decisions about methods of producing a given quantity of the output in a plant of given size and equipment. It involves the problem of what is called short-run cost minimization. The second layer, including the determination of the most profitable quantities of products to produce in any given plant, deals with what is called short-run profit maximization. The third layer, concerning the determination of the most profitable size and equipment of plant, relates to what is called long-run profit maximization.

Minimization of short-run costs

The production function

However much of a commodity a business firm produces, it endeavours to produce it as cheaply as possible. Taking the quality of the product and the prices of the productive factors as given, which is the usual situation, the firm's task is to determine the cheapest combination of factors of production that can produce the desired output. This task is best understood in terms of what is called the production function, *i.e.*, an equation that expresses the relationship between the quantities of factors employed and the amount of product obtained. It states the amount of product that can be obtained from each and every combination of factors. This relationship can be written mathematically as $y = f(x_1, x_2, \ldots, x_n; k_1, k_2, \ldots, k_m)$. Here, y denotes the quantity of output. The firm is presumed to use n

variable factors of production; that is, factors like hourly paid production workers and raw materials, the quantities of which can be increased or decreased. In the formula the quantity of the first variable factor is denoted by x_1 and so on. The firm is also presumed to use m fixed factors, or factors like fixed machinery, salaried staff, etc., the quantities of which cannot be varied readily or habitually. The available quantity of the first fixed factor is indicated in the formal by k_1 and so on. The entire formula expresses the amount of output that results when specified quantities of factors are employed. It must be noted that though the quantities of the factors determine the quantity of output, the reverse is not true, and as a general rule there will be many combinations of productive factors that could be used to produce the same output. Finding the cheapest of these is the problem of cost minimization.

The cost of production is simply the sum of the costs of all of the various factors. It can be written:

$$C = p_1x_1 + \cdots + p_nx_n + r_1k_1 + \cdots + r_nk_n,$$

in which p_1 denotes the price of a unit of the first variable factor, r_1 denotes the annual cost of owning and maintaining the first fixed factor, and so on. Here again one group of terms, the first, covers variable cost (roughly "direct costs" in accounting terminology), which can be changed readily; another group, the second, covers fixed cost (accountants' "overhead costs"), which includes items not easily varied. The discussion will deal first with variable cost.

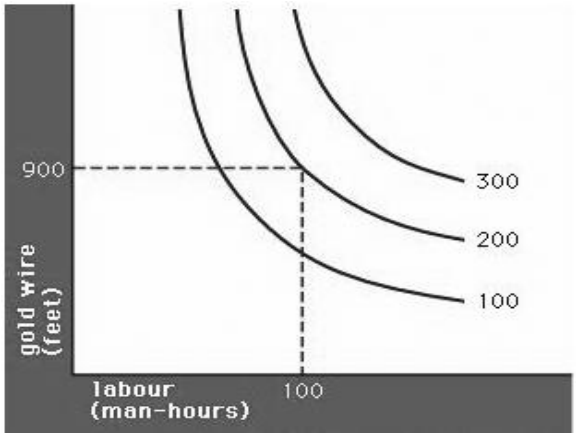


Fig.2.1

The principles involved in selecting the cheapest combination of variable factors can be seen in terms of a simple example. If a firm manufactures gold necklace chains in such a way that there are only two variable factors, labour (specifically, goldsmith-hours) and gold

wire, the production function for such a firm will be $y = f(x_1, x_2; k)$, in which the symbol k is included simply as a reminder that the number of chains producible by x_1 feet of gold wire and x_2 goldsmith-hours depends on the amount of machinery and other fixed capital available. Since there are only two variable factors, this production function can be portrayed graphically in a figure known as an isoquant diagram (Figure 1). In the graph, goldsmith-hours per month are plotted horizontally and the number of feet of gold wire used per month vertically. Each of the curved lines, called an isoquant, will then represent a certain number of necklace chains produced. The data displayed show that 100 goldsmith-hours plus 900 feet of gold wire can produce 200 necklace chains. But there are other combinations of variable inputs that could also produce 200 necklace chains per month. If the goldsmiths work more carefully and slowly, they can produce 200 chains from 850 feet of wire; but to produce so many chains more goldsmith-hours will be required, perhaps 130. The isoquant labelled "200" shows all the combinations of the variable inputs that will just suffice to produce 200 chains. The other two isoquants shown are interpreted similarly. It is obvious that many more isoquants, in principle an infinite number, could also be drawn. This diagram is a graphic display of the relationships expressed in the production function.

Substitution of factors

The isoquants also illustrate an important economic phenomenon: that of factor substitution. This means that one variable factor can be substituted for others; as a general rule a more lavish use of one variable factor will permit an unchanged amount of output to be produced with fewer units of some or all of the others. In the example above, labour was literally as good as gold and could be substituted for it. If it were not for factor substitution there would be no room for further decision after y, the number of chains to be produced, had been established.

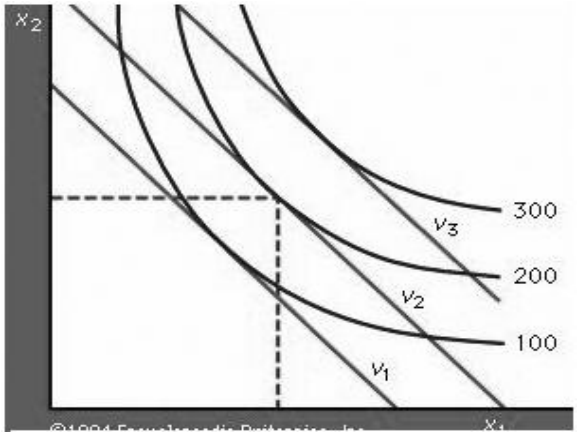


Fig.2.2

The shape of the isoquants shown, for which there is a good deal of empirical support, is very important. In moving along any one isoquant, the more of one factor that is employed, the less of the other will be needed to maintain the stated output; this is the graphic representation of factor substitutability. But there is a corollary: the more of one factor that is employed, the less it will be possible to reduce the use of the other by using more of the first. This is the property known as "diminishing marginal rates of substitution." The marginal rate of substitution of factor 1 for factor 2 is the number of units by which x_1 can be reduced per unit increase in x, output remaining unchanged. In the diagram, if feet of gold wire are indicated by x_1 and goldsmith-hours by x_2, then the marginal rate of substitution is shown by the steepness (the negative of the slope) of the isoquant; and it will be seen that it diminishes steadily as x_2 increases because it becomes harder and harder to economize on the use of gold simply by taking more care. The remainder of the analysis rests heavily on the assumption that diminishing marginal rates of substitution are characteristic of the production process generally.

The cost data and the technological data can now be brought together. The variable cost of using x_1, x_2 units of the factors of production is written $p_1x_1 + p_2x_2$, and this information can be added to the isoquant diagram (Figure 2.2). The straight line labelled v_2, called the v_2-isocost line, shows all the combinations of input that can be purchased for a specified variable cost, v_2. The other two isocost lines shown are interpreted similarly. The general formula for an isocost line is $p_1x_1 + p_2x_2 = v$, in which v is some particular variable cost. The slope of an isocost line is found by dividing p_2 by p_1 and depends only on the ratio of the prices of the two factors.

Three isocost lines are shown, corresponding to variable costs amounting to v_1, v_2, and v_3. If 200 units are to be produced, expenditure of v_1 on variable factors will not suffice since the v_1-isocost line never reaches the isoquant for 200 units. An expenditure of v_3 is more than sufficient; and v_2 is the lowest variable cost for which 200 units can be produced. Thus v_2 is found to be the minimum variable cost of producing 200 units (as v_3 is of 300 units) and the coordinates of the point where the v_2 isocost line touches the 200-unit isoquant are the quantities of the two factors that will be used when 200 units are to be produced and the prices of the two factors are in the ratio p_2/p_1. It may be noted that the cheapest combination for the production of any quantity will be found at the point at which the relevant isoquant is tangent to an isocost line. Thus, since the slope of an isoquant is given by the marginal rate of substitution, any firm trying to produce as cheaply as possible will always purchase or hire

factors in quantities such that the marginal rate of substitution will equal the ratio of their prices.

The isoquant–isocost diagram (or the corresponding solution by the alternative means of the calculus) solves the short-run cost minimization problem by determining the least-cost combination of variable factors that can produce a given output in a given plant. The variable cost incurred when the least-cost combination of inputs is used in conjunction with a given output of fixed equipment is called the variable cost of that quantity of output and denoted VC(y). The total cost incurred, variable plus fixed, is the short-run cost of that output, denoted SRC(y). Clearly SRC(y) = VC(y) + R(K), in which the second term symbolizes the sum of the annual costs of the fixed factors available.

Marginal cost

Two other concepts now become important. The average variable cost, written AVC(y), is the variable cost per unit of output. Algebraically, AVC(y) = VC(y)/y. The marginal variable cost, or simply marginal cost [MC(y)] is, roughly, the increase in variable cost incurred when output is increased by one unit; *i.e.*, MC(y) = VC(y + 1) - VC(y). Though for theoretical purposes a more precise definition can be obtained by regarding VC(y) as a continuous function of output, this is not necessary in the present case.

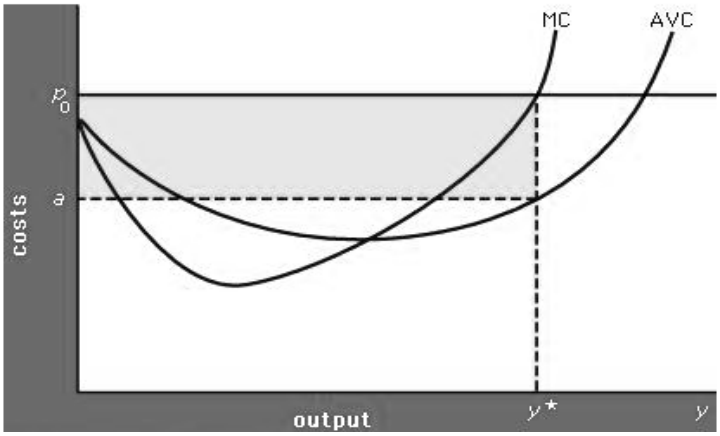


Fig.2.3

The usual behaviour of average and marginal variable costs in response to changes in the level of output from a given fixed plant is shown in Figure 2.3. In this figure costs (in dollars per unit) are measured vertically and output (in units per year) is shown horizontally. The figure is drawn for some particular fixed plant, and it can be seen that average costs are fairly high for very low levels of output relative to the size of the plant, largely because there is not enough work to keep a well-balanced work force fully occupied. People are either idle much of the time or shifting, expensively, from job to job. As output increases from a low

level, average costs decline to a low plateau. But as the capacity of the plant is approached, the inefficiencies incident on plant congestion force average costs up quite rapidly. Overtime may be incurred, outmoded equipment and inexperienced hands may be called into use, there may not be time to take machinery off the line for routine maintenance; or minor breakdowns and delays may disrupt schedules seriously because of inadequate slack and reserves. Thus the AVC curve has the flat-bottomed U-shape shown. The MC curve, as might be expected, falls faster and rises more rapidly than the AVC curve.

Maximization of short-run profits

The average and marginal cost curves just deduced are the keys to the solution of the second-level problem, the determination of the most profitable level of output to produce in a given plant. The only addition needed is the price of the product, say p_0.

The most profitable amount of output may be found by using these data. If the marginal cost of any given output (y) is less than the price, sales revenues will increase more than costs if output is increased by one unit (or even a few more); and profits will rise. Contrariwise, if the marginal cost is greater than the price, profits will be increased by cutting back output by at least one unit. It then follows that the output that maximizes profits is the one for which $MC(y) = p_0$. This is the second basic finding: in response to any price the profit-maximizing firm will produce and offer the quantity for which the marginal cost equals that price.

Such a conclusion is shown in Figure 2.3. In response to the price, p_0, shown, the firm will offer the quantity y^* given by the value of y for which the ordinate of the MC curve equals the price. If a denotes the corresponding average variable cost, net revenue per unit will be equal to $p_0 - a$, and the total excess of revenues over variable costs will be $y^*(p_0 - a)$, which is represented graphically by the shaded rectangle in the figure.

Marginal cost and price

The conclusion that marginal cost tends to equal price is important in that it shows how the quantity of output produced by a firm is influenced by the market price. If the market price is lower than the lowest point on the average variable cost curve, the firm will "cut its losses" by not producing anything. At any higher market price, the firm will produce the quantity for which marginal cost equals that price. Thus the quantity that the firm will produce in response to any price can be found in Figure 2.3 by reading the marginal cost curve, and for this reason the marginal cost curve is said to be the short-run supply curve for the firm.

The short-run supply curve for a product—that is, the total amount that all the firms producing it will produce in response to any market price—follows immediately, and is seen to be the sum of the short-run supply curves (or marginal cost curves, except when the price is below the bottoms of the average variable cost curves for some firms) of all the firms in the industry. This curve is of fundamental importance for economic analysis, for together with the demand curve for the product it determines the market price of the commodity and the amount that will be produced and purchased.

One drawback must, however, be noted. In the demonstration of the supply curves for the firms, and hence of the industry, it was assumed that factor prices were fixed. Though this is fair enough for a single firm, the fact is that if all firms together attempt to increase their outputs in response to an increase in the price of the product, they are likely to bid up the prices of some or all of the factors of production that they use. In that event the product supply curve as calculated will overstate the increase in output that will be elicited by an increase in price. A more sophisticated type of supply curve, incorporating induced changes in factor prices, is therefore necessary. Such curves are discussed in the standard literature of this subject.

Marginal product

It is now possible to derive the relationship between product prices and factor prices, which is the basis of the theory of income distribution. To this end, the marginal product of a factor is defined as the amount that output would be increased if one more unit of the factor were employed, all other circumstances remaining the same. Algebraically, it may be expressed as the difference between the product of a given amount of the factor and the product when that factor is increased by an additional unit. Thus if $MP_1(x_1)$ denotes the marginal product of factor 1 when x_1 units are employed, then $MP_1(x_1) = f(x_1 + 1, x_2, \ldots, x_n; k) - f(x_1, x_2 \ldots, x_n; k)$. The marginal products are closely related to the marginal rates of substitution previously defined. If an additional unit of factor 1 will increase output by f_1 units, for example, then one more unit of output can be obtained by employing $1/f_1$ more units of factor 1. Similarly, if the marginal product of factor 2 is f_2, then output will fall by one unit if the use of factor 2 is reduced by $1/f_2$ units. Thus output will remain unchanged, to a good approximation, if $1/f_1$ units of factor 1 are used to replace $1/f_2$ units of factor 2. The marginal rate of substitution is therefore f_2/f_1, or the ratio of the marginal products of the two factors. It has already been shown that the marginal rate of substitution also equals the ratio of the

prices of the factors, and it therefore follows that the prices (or wages) of the factors are proportional to their marginal products.

This is one of the most significant theoretical findings in economics. To restate it briefly: factors of production are paid in proportion to their marginal products. This is not a question of social equity but merely a consequence of the efforts of businessmen to produce as cheaply as possible.

Further, the marginal products of the factors are closely related to marginal costs and, therefore, to product prices. For if one more unit of factor 1 is employed, output will be increased by $MP_1(x_1)$ units and variable cost by p_1; so the marginal cost of additional units produced will be $p_1/MP_1(x_1)$. Similarly, if additional output is obtained by employing an additional unit of factor 2, the marginal cost will be $p_2/MP_2(x_2)$. But, as shown above, these two numbers are the same; whichever factor i is used to increase output, the marginal cost will be $p_i/MP_i(x_i)$ and, furthermore, the firm will choose its output level so that the marginal cost will be equal to the price, p_0.

Therefore it has been established that $p_1 = p_0MP_1(x_1)$, $p_2 = p_0MP_2(x_2)$, . . . , or the price of each factor is the price of the product multiplied by its marginal product, which is the value of its marginal product. This is also a fundamental theorem of income distribution and one of the most significant theorems in economics. Its logic can be perceived directly. If the equality is violated for any factor, the businessman can increase his profits either by hiring units of the factor or by laying them off until the equality is satisfied, and presumably the businessman will do so.

The theory of production decisions in the short run, as just outlined, leads to two conclusions (of fundamental importance throughout the field of economics) about the responses of business firms to the market prices of the commodities they produce and the factors of production they buy or hire: (1) the firm will produce the quantity of its product for which the marginal cost is equal to the market price and (2) it will purchase or hire factors of production in such quantities that the price of the commodity produced multiplied by the marginal product of the factor will be equal to the cost of a unit of the factor. The first explains the supply curves of the commodities produced in an economy. Though the conclusions were deduced within the context of a firm that uses two factors of production, they are clearly applicable in general.

Maximization of long-run profits

Relationship between the short run and the long run:

The theory of long-run profit-maximizing behaviour rests on the short-run theory that has just been presented but is considerably more complex because of two features: (1) long-run cost curves, to be defined below, are more varied in shape than the corresponding short-run cost curves, and (2) the long-run behaviour of an industry cannot be deduced simply from the long-run behaviour of the firms in it because the roster of firms is subject to change. It is of the essence of long-run adjustments that they take place by the addition or reduction of fixed productive capacity by both established firms and new or recently created firms.

At any one time an established firm with an existing plant will make its short-run decisions by comparing the ruling price of its commodity with cost curves corresponding to that plant. If the price is so high that the firm is operating on the rising leg of its short-run cost curve, its marginal costs will be high—higher than its average costs—and it will be enjoying operating profits, as shown in Figure 3. The firm will then consider whether it could increase its profits by enlarging its plant. The effect of plant enlargement is to reduce the variable cost of producing high levels of output by reducing the strain on limited production facilities, at the expense of increasing the level of fixed costs.

In response to any level of output that it expects to continue for some time, the firm will desire and eventually acquire the fixed plant for which the short-run costs of that level of output are as low as possible. This leads to the concept of the long-run cost curve: the long-run costs of any level of output are the short-run costs of producing that output in the plant that makes those short-run costs as low as possible. These result from balancing the fixed costs entailed by any plant against the short-run costs of producing in that plant. The long-run costs of producing y are denoted by LRC(y). The average long-run cost of y is the long-run cost per unit of y [algebraically LAC(y) = LRC(y)/y]. The marginal long-run cost is the increase in long-run cost resulting from an increase of one unit in the level of output. It represents a combination of short-run and long-run adjustments to a slight increase in the rate of output. It can be shown that the long-run marginal cost equals the marginal cost as previously defined when the cost-minimizing fixed plant is used.

Long-run cost curves

Cost curves appropriate for long-run analysis are more varied in shape than short-run cost curves and fall into three broad classes. In constant-cost industries, average cost is about the same at all levels of output except the very lowest. Constant costs prevail in

manufacturing industries in which capacity is expanded by replicating facilities without changing the technique of production, as a cotton mill expands by increasing the number of spindles. In decreasing-cost industries, average cost declines as the rate of output grows, at least until the plant is large enough to supply an appreciable fraction of its market. Decreasing costs are characteristic of manufacturing in which heavy, automated machinery is economical for large volumes of output. Automobile and steel manufacturing are leading examples. Decreasing costs are inconsistent with competitive conditions, since they permit a few large firms to drive all smaller competitors out of business. Finally, in increasing-cost industries average costs rise with the volume of output generally because the firm cannot obtain additional fixed capacity that is as efficient as the plant it already has. The most important examples are agriculture and extractive industries.

Criticisms of the theory

The theory of production has been subject to much criticism. One objection is that the concept of the production function is not derived from observation or practice. Even the most sophisticated firms do not know the direct functional relationship between their basic raw inputs and their ultimate outputs. This objection can be got around by applying the recently developed techniques of linear programming, which employ observable data without recourse to the production function and lead to practically the same conclusions.

On another level the theory has been charged with excessive simplification. It assumes that there are no changes in the rest of the economy while individual firms and industries are making the adjustments described in the theory; it neglects changes in the technique of production; and it pays no attention to the risks and uncertainties that becloud all business decisions. These criticisms are especially damaging to the theory of long-run profit maximization. On still another level, critics of the theory maintain that businessmen are not always concerned with maximizing profits or minimizing costs.

Though all of the criticisms have merit, the simplified theory of production does nevertheless indicate some basic forces and tendencies operating in the economy. The theorems should be understood as conditions that the economy tends toward, rather than conditions that are always and instantaneously achieved. It is rare for them to be attained exactly, but it is just as rare for substantial violations of the theorems to endure.

Only the simplest aspects of the theory were described above. Without much difficulty it could be extended to cover firms that produce more than one product, as almost

all firms do. With more difficulty it could be applied to firms whose decisions affect the prices at which they sell and buy (monopoly, monopolistic competition, monopsony). The behaviour of other firms that recognize the possibility that their competitors may retaliate (oligopoly) is still a theory of production subject to controversy and research.

Law of Variable Proportions

Meaning:

Law of variable proportions occupies an important place in economic theory. This law examines the production function with one factor variable, keeping the quantities of other factors fixed. In other words, it refers to the input-output relation when output is increased by varying the quantity of one input.

When the quantity of one factor is varied, keeping the quantity of other factors constant, the proportion between the variable factor and the fixed factor is altered; the ratio of employment of the variable factor to that of the fixed factor goes on increasing as the quantity of the variable factor is increased.

Since under this law we study the effects on output of variation in factor proportions, this is also known as the law of variable proportions. Thus law of variable proportions is the new name for the famous ”Law of Diminishing Returns” of classical economics. This law has played a vital role in the history of economic thought and occupies an equally important place in modern economic theory. This law has been supported by the empirical evidence about the real world.

The law of variable proportions or diminishing returns has been stated by various economists in the following manner:

As equal increments of one input are added; the inputs of other productive services being held constant, beyond a certain point the resulting increments of product will decrease, i.e., the marginal products will diminish,” (G. Stigler)

“As the proportion of one factor in a combination of factors is increased, after a point, first the marginal and then the average product of that factor will diminish.” (F. Benham)

“An increase in some inputs relative to other fixed inputs will, in a given state of technology, cause output to increase; but after a point the extra output resulting from the same addition of extra inputs will become less.” (Paul A. Samuelson)

Marshall discussed the law of diminishing returns in relation to agriculture. He defines the law as follows: “An increase in the capital and labour applied in the cultivation of

land causes in general a less than proportionate increase in the amount of product raised unless it happens to coincide with an improvement in the arts of agriculture."

It is obvious from the above definitions of the law of variable proportions (or the law of diminishing returns) that it refers to the behaviour of output as the quantity of one factor is increased, keeping the quantity of other factors fixed and further it states that the marginal product and average product will eventually decline.

Assumptions of the Law:

The law of variable proportions or diminishing returns, as stated above, holds good under the following conditions:

1. First, the state of technology is assumed to be given and unchanged. If there is improvement in the technology, then marginal and average products may rise instead of diminishing.

2. Secondly, there must be some inputs whose quantity is kept fixed. This is one of the ways by which we can alter the factor proportions and know its effect on output. This law does not apply in case all factors are proportionately varied. Behaviour of output as a result of the variation in all inputs is discussed under "returns to scale".

3. Thirdly the law is based upon the possibility of varying the proportions in which the various factors can be combined to produce a product. The law does not apply to those cases where the factors must be used in fixed proportions to yield a product.

When the various factors are required to be used in rigidly fixed proportions, then the increase in one factor would not lead to any increase in output, that is, the marginal product of the factor will be zero and not diminishing. It may, however, be pointed out that products requiring fixed proportions of factors are quiet uncommon. Thus, the law of variable proportion applies to most of the cases of production in the real world.

The law of variable proportions is illustrated in Table 2.1.and Fig. 2.3. We shall first explain it by considering Table 2.1. Assume that there is a given fixed amount of land, with which more units of the variable factor labour, is used to produce agricultural output.

Table 2.1

Returns to Labour

Units of Labour	*Total Product (Quintals)*	*Marginal Product (Quintals)*	*Average Product (Quintals)*
L	Q	$\frac{\Delta Q}{\Delta L}$	$\frac{Q}{L}$
1	80	80	80
2	170	90	85
3	270	100	90
4	368	98	92
5	430	62	86
6	480	50	80
7	504	24	72
8	504	0	63
9	495	–9	55
10	480	–15	48

With a given fixed quantity of land, as a farmer raises employment of labour from one unit to 7 units, the total product increases from 80 quintals to 504 quintals of wheat. Beyond the employment of 8 units of labour, total product diminishes. It is worth noting that up to the use of 3 units of labour, total product increases at an increasing rate.

This fact is clearly revealed from column 3 which shows successive marginal products of labour as extra units of labour are used. Marginal product of labour, it may be recalled, is the increment in total output due to the use of an extra unit of labour.

It will be seen from Col. 3 of Table 2.1, that the marginal product of labour initially rises and beyond the use of three units of labour, it starts diminishing. Thus when 3 units of labour are employed, marginal product of labour is 100 and with the use of 4th and 5th units of labour marginal product of labour falls to 98 and 62 respectively.

Beyond the use of eight units of labour, total product diminishes and therefore marginal product of labour becomes negative. As regards average product of labour, it rises upto the use of fourth unit of labour and beyond that it is falling throughout.

Three Stages of the Law of Variable Proportions:

The behaviour of output when the varying quantity of one factor is combined with a fixed quantity of the other can be divided into three distinct stages. In order to understand these three stages it is better to graphically illustrate the production function with one factor variable.

This has been done in Fig. 2.3. In this figure, on the X-axis the quantity of the variable factor is measured and on the X-axis the total product, average product and marginal product are measured. How the total product, average product and marginal product a variable factor change as a result of the increase in its quantity, that is, by increasing the quantity of one factor to a fixed quantity of the others will be seen from Fig. 2.3.

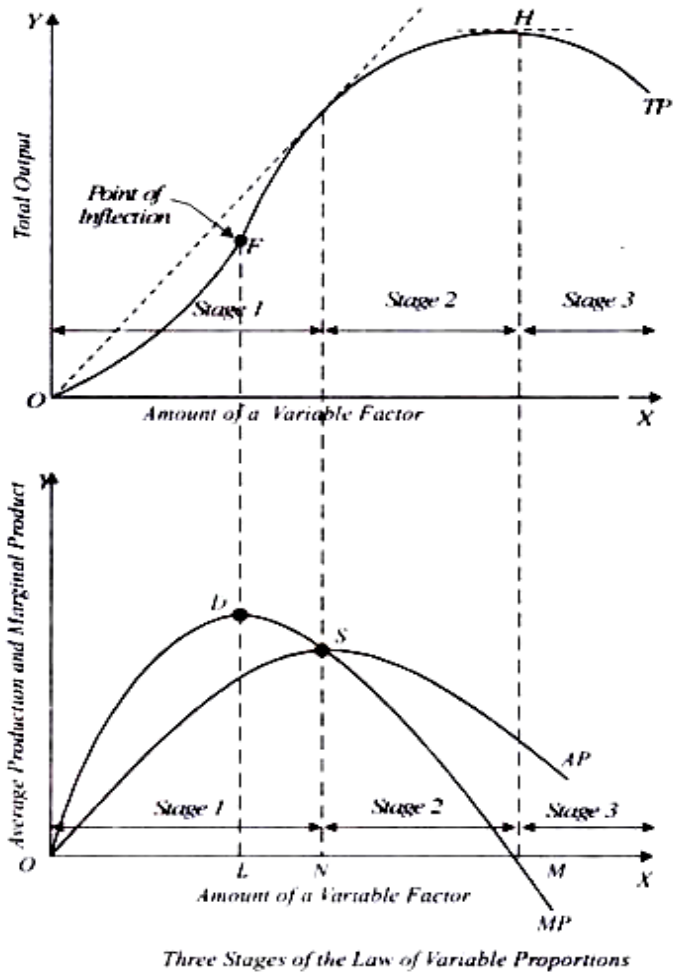


Fig.2.3

In the top panel of this figure, the total product curve TP of variable factor goes on increasing to a point and alter that it starts declining. In the bottom panel average and marginal product curves of labour also rise and then decline; marginal product curve starts declining earlier than the average product curve.

The behaviour of these total, average and marginal products of the variable factor as a result of the increase in its amount is generally divided into three stages which are explained below:

Stage 1:

In this stage, total product curve TP increases at an increasing rate up to a point. In Fig. 2.3. from the origin to the point F, slope of the total product curve TP is increasing, that is, up to the point F, the total product increases at an increasing rate (the total product curve TP is concave upward upto the point F), which means that the marginal product MP of the variable factor is rising.

From the point F onwards during the stage 1, the total product curve goes on rising but its slope is declining which means that from point F onwards the total product increases at a diminishing rate (total product curve TP is concave down-ward), i.e., marginal product falls but is positive.

The point F where the total product stops increasing at an increasing rate and starts increasing at the diminishing rate is called the point of inflection. Vertically corresponding to this point of inflection marginal product is maximum, after which it starts diminishing.

Thus, marginal product of the variable factor starts diminishing beyond OL amount of the variable factor. That is, law of diminishing returns starts operating in stage 1 from point D on the MP curve or from OL amount of the variable factor used.

This first stage ends where the average product curve AP reaches its highest point, that is, point S on AP curve or CW amount of the variable factor used. During stage 1, when marginal product of the variable factor is falling it still exceeds its average product and so continues to cause the average product curve to rise.

Thus, during stage 1, whereas marginal product curve of a variable factor rises in a part and then falls, the average product curve rises throughout. In the first stage, the quantity of the fixed factor is too much relative to the quantity of the variable factor so that if some of the fixed factor is withdrawn, the total product will increase. Thus, in the first stage marginal product of the fixed factor is negative.

Stage 2:

In stage 2, the total product continues to increase at a diminishing rate until it reaches its maximum point H where the second stage ends. In this stage both the marginal product and the average product of the variable factor are diminishing but remain positive.

At the end of the second stage, that is, at point M marginal product of the variable factor is zero (corresponding to the highest point H of the total product curve TP). Stage 2 is very crucial and important because as will be explained below the firm will seek to produce in its range.

Stage 3: Stage of Negative Returns:

In stage 3 with the increase in the variable factor the total product declines and therefore the total product curve TP slopes downward. As a result, marginal product of the variable factor is negative and the marginal product curve MP goes below the X-axis. In this stage the variable factor is too much relative to the fixed factor. This stage is called the stage

of negative returns, since the marginal product of the variable factor is negative during this stage.

It may be noted that stage 1 and stage 3 are completely symmetrical. In stage 1 the fixed factor is too much relative to the variable factor. Therefore, in stage 1, marginal product of the fixed factor is negative. On the other hand, in stage 3 the variable factor is too much relative to the fixed factor. Therefore, in stage 3, the marginal product of the variable factor is negative.

The Stage of Operation:

Now, an important question is in which stage a rational producer will seek to produce. A rational producer will never choose to produce in stage 3 where marginal product of the variable factor is negative. Marginal product of the variable factor being negative in stage 3, a producer can always increase his output by reducing the amount of the variable factor.

It is thus clear that a rational producer will never be producing in stage 3. Even if the variable factor is free, the rational producer will stop at the end of the second stage where the marginal product of the variable factor is zero.

At the end point M of the second stage where the marginal product of the variable factor is zero, the producer will be maximising the total product and will thus be making maximum use of the variable factor. A rational producer will also not choose to produce in stage 1 where the marginal product of the fixed factor is negative.

A producer producing in stage 1 means that he will not be making the best use of the fixed factor and further that he will not be utilising fully the opportunities of increasing production by increasing quantity of the variable factor whose average product continues to rise throughout the stage 1. Thus, a rational entrepreneur will not stop in stage 1 but will expand further.

Even if the fixed factor is free (i.e., costs nothing), the rational entrepreneur will stop only at the end of stage 1 (i.e., at point N) where the average product of the variable factor is maximum. At the end point N of stage 1, the producer they will be making maximum use of the fixed factor.

It is thus clear from above that the rational producer will never be found producing in stage 1 and stage 3. Stage 1 and 3 may, therefore, be called stages of economic absurdity or economic non-sense. The stages 1 and 3 represent non-economic regions in production function.

A rational producer will always seek to produce in stage 2 where both the marginal product and average product of the variable factor are diminishing. At which particular point in this stage, the producer will decide to produce depends upon the prices of factors. The stage 2 represents the range of rational production decisions.

We have seen above how output varies as the factor proportions are altered at any given moment. We have also noticed that this input-output relation can be divided into three stages. Now, the question arises as to what causes increasing marginal returns to the variable factor in the beginning, diminishing marginal returns later and negative marginal returns to the variable factor ultimately.

Causes of Initial Increasing Marginal Returns to a Factor:

In the beginning, the quantity of the fixed factor is abundant relative to the quantity of the variable factor. Therefore, when more and more units of a variable factor are added to the constant quantity of the fixed factor, the fixed factor is more intensively and effectively utilised.

This causes the production to increase at a rapid rate. When, in the beginning the variable factor is relatively smaller in quantity, some amount of the fixed factor may remain unutilised and therefore when the variable factor is increased fuller utilisation of the fixed factor becomes possible with the result that increasing returns are obtained.

The question arises as to why the fixed factor is not initially taken in an appropriate quantity which suits the available quantity of the variable factor. Answer to this question is provided by the fact that generally those factors are taken as fixed which are indivisible. Indivisibility of a factor means that due to technological requirements a minimum amount of that factor must be employed whatever the level of output.

Thus, as more units of variable factor are employed to work with an indivisible fixed factor, output greatly increases in the beginning due to fuller and more effective utilisation of the latter. Thus, we see that it is the indivisibility of some factors which causes increasing returns to the variable factor in the beginning.

The second reason why we get increasing returns to the variable factor in the initial stage is that as more units of the variable factor are employed the efficiency of the variable factor itself increases. This is because when there is a sufficient quantity of the variable factor, it becomes possible to introduce specialisation or division of labour which results in higher productivity. The greater the quantity of the variable factor, the greater the scope of specialisation and hence the greater will be the level of its productivity or efficiency.

Causes of Diminishing marginal Returns to a Factor:

The stage of diminishing marginal returns in the production function with one factor variable is the most important. The question arises as to why we get diminishing marginal returns after a certain amount of the variable factor has been added to a fixed quantity of the other factor.

As explained above, increasing returns to a variable factor occur initially primarily because of the more effective and fuller use of the fixed factor becomes possible as more units of the variable factor are employed to work with it.

Once the point is reached at which the amount of the variable factor is sufficient to ensure the efficient utilisation of the fixed factor, then further increases in the variable factor will cause marginal and average products of a variable factor to decline because the fixed factor then becomes inadequate relative to the quantity of the variable factor.

In other words, the contributions to the production made by the variable factor after a point become less and less because the additional units of the variable factor have less and less of the fixed factor to work with. The production is the result of the co-operation of various factors aiding each other. Now, how much aid one factor provides to the others depends upon how much there is of it.

Eventually, the fixed factor is abundant relative to the number of the variable factor and the former provides much aid to the later. Eventually, the fixed factor becomes more and more scarce in relation to the variable factor so that as the units of the variable factor are increased they receive less and less aid from the fixed factor. As a result, the marginal and average products of the variable factor decline ultimately.

The phenomenon of diminishing marginal returns, like that of increasing marginal returns, rests upon the indivisibility of the fixed factor. As explained above, the important reason for increasing returns to a factor in the beginning is the fact that the fixed factor is indivisible which has to be employed whether the output to be produced is small or large.

When the indivisible fixed factor is not being fully used, successive increases in a variable factor add more to output since fuller and more efficient use is made of the indivisible fixed factor. But there is generally a limit to the range of employment of the variable factor over which its marginal and average products will increase.

There will usually be a level of employment of the Variable factor at which indivisible fixed factor is being as fully and efficiently used as possible. It will happen when

the variable factor has increased to such an amount that the fixed indivisible factor is being used in the "best or optimum proportion" with the variable factor.

Once the optimum proportion is disturbed by further increases in the variable factor, returns to a variable factor (i.e., marginal product and average product) will diminish primarily because the indivisible factor is being used too intensively, or in other words, the fixed factor is being used in non-optimal proportion with the variable factor.

Just as the marginal product of the variable factor increases in the first stage when better and fuller use of the fixed indivisible factor is being made, so the marginal product of the variable factor diminishes when the fixed indivisible factor is being worked too hard.

If the fixed factor was perfectly divisible, neither the increasing nor the diminishing returns to a variable factor would have occurred. If the factors were perfectly divisible, then there would not have been the necessity of taking a large quantity of the fixed factor in the beginning to combine with the varying quantities of the other factor.

In the presence of perfect divisibility, the optimum proportion between the factors could have always been achieved. Perfect divisibility of the factors implies that a small firm with a small machine and one worker would be as efficient as a large firm with a large machine and many workers.

The productivity of the factors would be the same in the two cases. Thus, we see that if the factors were perfectly divisible, then the question of varying factor proportions would not have arisen and hence the phenomena of increasing and diminishing marginal returns to a variable factor would not have occurred. Prof. Bober rightly remarks: "Let divisibility enter through the door, law of variable proportions rushes out through the window."

Mrs.Joan Robinson goes deeper into the causes of diminishing returns. She holds that the diminishing marginal returns occur because the factors of production are imperfect substitutes for one another. As seen above, diminishing returns occur during the second stage since the fixed factor is now inadequate relatively to the variable factor. Now, a factor which is scarce in supply is taken as fixed.

When there is a scarce factor, quantity of that factor cannot be increased in accordance with the varying quantities of the other factors, which, after the optimum proportion of factors is achieved, results in diminishing returns.

If now some factors were available which perfect substitute of the scarce fixed factor was, then the scarce fixed factor during the second stage would have been made up by the

increase in supply of its perfect substitute with the result that output could be expanded without diminishing returns.

Thus, even if one of the variable factors which we add to the fixed factor were perfect substitute of the fixed factor, then when, in the second stage, the fixed factor becomes relatively deficient, its deficiency would have been made up the increase in the variable factor which is its perfect substitute.

Thus, Joan Robinson says, "What the Law of Diminishing Returns really states is that there is a limit to the extent to which one factor of production can be substituted for another, or, in other words, that the elasticity of substitution between factor is not infinite.

If this were not true, it would be possible, when one factor of production is fixed in amount and the rest are in perfectly elastic supply, to produce part of the output with the aid of the fixed factor, and then, when the optimum proportion between this and other factors was attained, to substitute some other factor for it and to increase output at constant cost." We, therefore, see that diminishing returns operate because the elasticity of substitution between factors is not infinite.

Explanation of Negative Marginal Returns to a Factor:

As the amount of a variable factor continues to be increased to a fixed quantity of the other factor, a stage is reached when the total product declines and the marginal product of the variable factor becomes negative.

This phenomenon of negative marginal returns to the variable factor in stage 3 is due to the fact that the number of the variable factor becomes too excessive relative to the fixed factor so that they obstruct each other with the result that the total output falls instead of rising.

Besides, too large a number of the variable factor also impairs the efficiency of the fixed factor. The proverb "too many cooks spoil the broth" aptly applies to this situation. In such a situation, a reduction in the units of the variable factor will increase the total output.

Laws of Returns to scale: The Isoquant Approach

The Laws of Returns to Scale:

The laws of returns to scale can also be explained in terms of the isoquant approach. The laws of returns to scale refer to the effects of a change in the scale of factors (inputs) upon output in the long-run when the combinations of factors are changed in some proportion. If by increasing two factors, say labour and capital, in the same proportion, output increases in exactly the same proportion, there are constant returns to scale. If in order to

secure equal increases in output, both factors are increased in larger proportionate units, there are decreasing returns to scale. If in order to get equal increases in output, both factors are increased in smaller proportionate units, there are increasing returns to scale.

The returns to scale can be shown diagrammatically on an expansion path "by the distance between successive 'multiple-level-of-output' isoquants, that is, isoquants that show levels of output which are multiples of some base level of output, e.g., 100, 200, 300, etc."

Increasing Returns to Scale:

Figure 2.3 shows the case of increasing returns to scale where to get equal increases in output, lesser proportionate increases in both factors, labour and capital, are required.

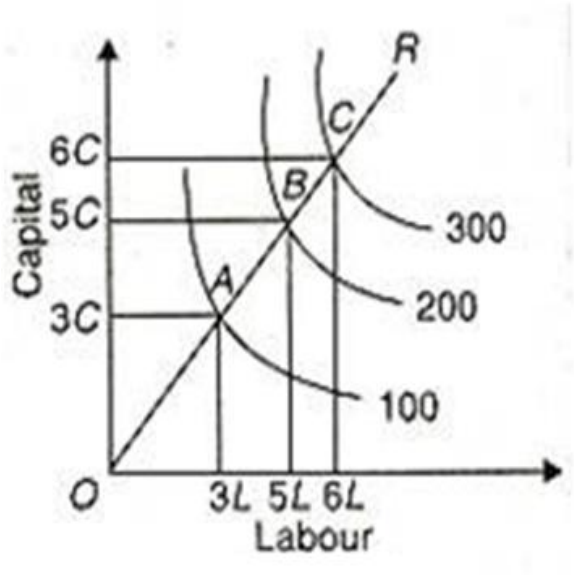


Fig.2.3

It follows that in the figure:

100 units of output require 3C +3L

200 units of output require 5C + 5L

300 units of output require 6C + 6L

So that along the expansion path OR, OA > AB > BC. In this case, the production function is homogeneous of degree greater than one.

The increasing returns to scale are attributed to the existence of indivisibilities in machines, management, labour, finance, etc. Some items of equipment or some activities have a minimum size and cannot be divided into smaller units. When a business unit expands, the returns to scale increase because the indivisible factors are employed to their full capacity.

Increasing returns to scale also result from specialisation and division of labour. When the scale of the firm expands there is wide scope for specialisation and division of labour. Work can be divided into small tasks and workers can be concentrated to narrower range of processes. For this, specialized equipment can be installed. Thus with specialization, efficiency increases and increasing returns to scale follow.

Further, as the firm expands, it enjoys internal economies of production. It may be able to install better machines, sell its products more easily, borrow money cheaply, procure the services of more efficient manager and workers, etc. All these economies help in increasing the returns to scale more than proportionately.

Not only this, a firm also enjoys increasing returns to scale due to external economies. When the industry itself expands to meet the increased 'long-run demand for its product, external economies appear which are shared by all the firms in the industry. When a large number of firms are concentrated at one place, skilled labour, credit and transport facilities are easily available. Subsidiary industries crop up to help the main industry. Trade journals, research and training centres appear which help in increasing the productive efficiency of the firms. Thus these external economies are also the cause of increasing returns to scale.

Decreasing Returns to Scale:

Figure 2.4 shows the case of decreasing returns where to get equal increases in output, larger proportionate increases in both labour and capital are required.

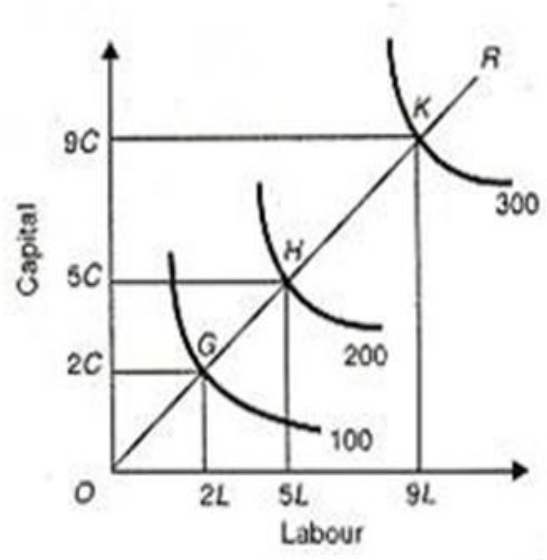


Fig.2.4

It follows that:

100 units of output require 2C + 2L

200 units of output require 5C + 5L

300 units of output require 9C + 9L

So that along the expansion path OR, OG < GH < HK.

In this case, the production function is homogeneous of degree less than one.

Returns to scale may start diminishing due to the following factors. Indivisible factors may become inefficient and less productive. The firm experiences internal diseconomies. Business may become heavy and produce problems of supervision and coordination. Large management creates difficulties of control and rigidities. To these internal diseconomies are added external diseconomies of scale. These arise from higher factor prices or from diminishing productivities of the factors.

As the industry continues to expand the demand for skilled labour, land, capital, etc. rises. There being perfect competition, intensive bidding raises wages, rent and interest. Prices of raw materials also go up. Transport and marketing difficulties emerge. All these factors tend to raise costs and the expansion of the firms leads to diminishing returns to scale so that doubling the scale would not lead to doubling the output.

Constant Returns to Scale:

Figure 2.5 shows the case of constant returns to scale. Where the distance between the isoquants 100, 200 and 300 along the expansion path OR is the same, i.e., OD = DE = EF. It means that if units of both factors, labour and capital, are doubled, the output is doubled. To treble output, units of both factors are trebled.

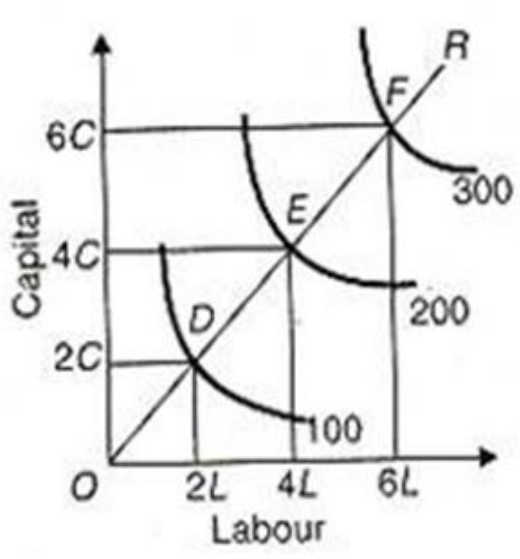


Fig.2.5

It follows that:

100 units of output require 1 (2C + 2L) = 2C + 2L

200 units of output require 2(2C + 2L) = 4C + 4L

300 units of output require 3(2C + 2L) = 6C + 6L

The returns to scale are constant when internal economies enjoyed by a firm are neutralised by internal diseconomies so that output increases in the same proportion. Another reason is the balancing of external economies and external diseconomies. Constant returns to scale also result when factors of production are perfectly divisible, substitutable, homogeneous and their supplies are perfectly elastic at given prices.

That is why, in the case of constant returns to scale, the production function is homogeneous of degree one. The various production functions were explained in terms of the traditional analysis. The technique involved here is similar to the indifference curve technique used in consumption theory.

Isoquants:

An isoquant (isoproduct) is a curve on which the various combinations of labour and capital show the same output. According to Cohen and Cyert, "An isoproduct curve is a curve along which the maximum achievable rate of production is constant." It is also known as a production indifference curve or a constant product curve. Just as indifference curve shows the various combinations of any two commodities that give the consumer the same amount of satisfaction (iso-utility), similarly an isoquant indicates the various combinations of two factors of production which give the producer the same level of output per unit of time. Table 2.1 shows a hypothetical isoquant schedule of a firm producing 100 units of a good.

Table 2.2: Isoquant Schedule:

Combination	Units of Capital	units of Labour	Total Output (in units)
A	9	5	100
B	6	10	100
C	4	15	100
D	3	20	100

This Table 2.2 is illustrated on Figure 2.5 where labour units are measured along the X-axis and capital units on the Y-axis. The first, second, third and the fourth combinations are shown as A, S, C and D respectively. Connect all these points and we have a curve IQ.

This is an isoquant. The firm can produce 100 units of output at point A on this curve by having a combination of 9 units of capital and 5 units of labour. Similarly, point B shows a combination of 6 units of capital and 10 units of labour; point C shows a combination of 4 units of capital and' 15 units of labour; and point D, a combination of 3 units of capital and 20 units of labour to yield the same output of 100 units.

An isoquant map shows a number of isoquants representing different amounts of output. In Figure 2.5, curves IQ, IQ_1 and IQ_2 show an isoquant map. Starting from the curve IQ which yields 100 units of product, the curve IQ_1, shows 200 units and the IQ_2 curve 300 units of the product which can be produced with altogether different combinations of the two factors.

Isoquants vs. Indifference Curves:

An isoquant is analogous to an indifference curve in more than one way. In it, two factors (capital and labour) replace two commodities of consumption. An isoquant shows equal level of product while an indifference curve shows equal level of satisfaction at all points. The properties of isoquants, as we shall study below, are exactly similar to those of indifference curves. However, there are certain differences between isoquants and indifference curves.

Firstly, an indifference curve represents satisfaction which cannot be measured in physical units. In the case of an isoquant the product can be measured in physical units.

Secondly, on an indifference map one can only say that a higher indifference curve gives more satisfaction than a lower one, but it cannot be said how much more or less satisfaction is being derived from one indifference curve as compared to the other, whereas one can easily tell by how much output is greater on a higher isoquant in comparison with a lower isoquant.

In Figure 2.6 output on the curve $1Q_I$ is double, and on the IQ_2 triple than on the curve IQ. Lastly, since satisfaction on indifference curves cannot be measured in physical units, they are given arbitrary numbers 1, 2, 3, 4, etc. The isoquants have an added advantage over the former because they can be labelled in physical units, as 100, 200, 300, etc. in Figure 24.1, to indicate the output level to which each curve corresponds.

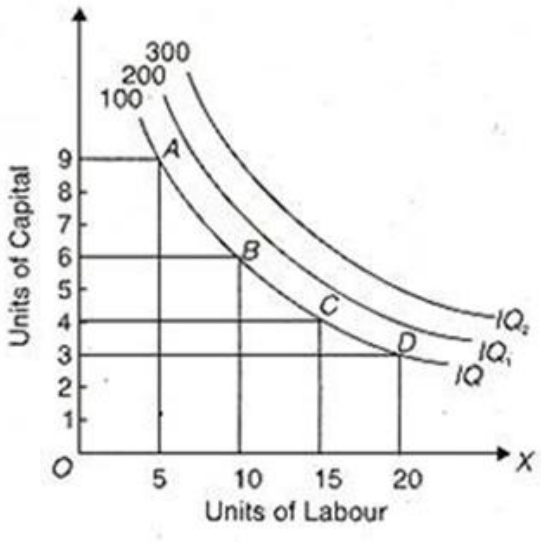


Fig.2.6

Properties of Isoquants:

Isoquants possess certain properties which are similar to those of indifference curves.

(1) Isoquants are negatively inclined:

If they do not have a negative slope, certain logical absurdities follow. If the isoquant slopes upward to the right, it implies that both capital and labour increase but they produce the same output. In Figure 2.7 (A), combination B on the IQ curve having a larger amount of both capital and labour ($OC_1 + OL_1 > OC + OL$) will yield more output than before. Therefore, point A and B on the IQ curve cannot be of equal product.

Suppose the isoquant is vertical as shown in Figure 2.7 (B), which implies a given amount of labour is combined with different units of capital. Since OL of labour and OC_1 of capital will produce a larger amount than produced by OL of labour and OC of capital, the isoquant IQ cannot be a constant product curve.

Take Figure 2.7 (C) where the isoquant is horizontal which means combining more of labour with the same quantity of capital. Here OC of capital and OL_1 of labour will produce a larger or smaller amount than produced by the combination OC of capital and OL of labour. Therefore, a horizontal isoquant cannot be an equal product curve.

Thus it is clear that an isoquant must slope downward to the right as shown in Figure 2.7 (D) where points A and B on the IQ curve are of equal quantity. As the amount of capital decreases from OC to OC_1 and that of labour increases from OL to OL_1 so that output remains constant.

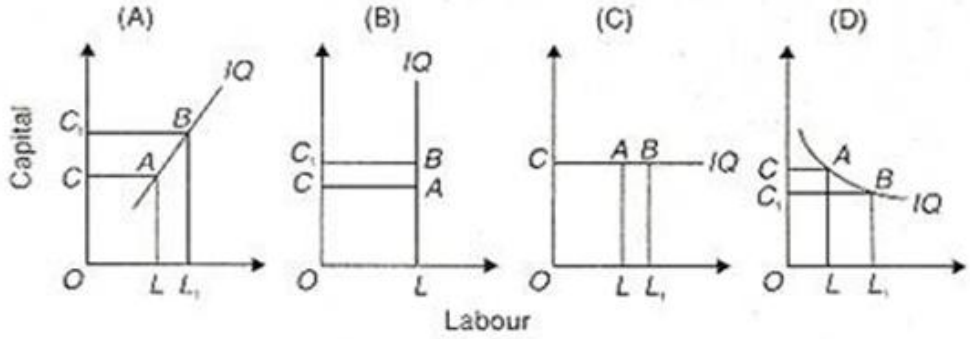


Fig.2.7

(2) An Isoquant lying above and to the right of another represents a higher output level. In Figure 2.8 combination B on IQ_1 curve shows larger output than point A on the curve IQ. The combination of OC of capital and OL of labour yields 100 units of product while OC_1 of capital and OL_1 of labour produce 200 units. Therefore, the isoquant IQ_1 which lies above and to the right of the isoquant IQ, represents a larger output level.

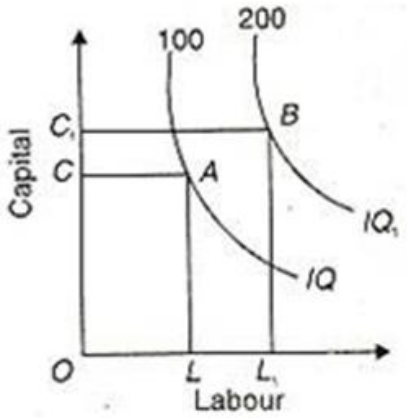


Fig.2.8

(3) No two isoquants can intersect each other. On the isoquant IQ, combination A = B. And on the isoquant IQ_1 combination R=S. But combination S is preferred to combination B, being on the higher portion of isoquant IQ_1. On the other hand, combination A is preferred to R, the former being on the higher portion of the isoquant IQ. To put it algebraically, it means that S> B and R< A. But this is logically absurd because S combination is as productive as R and A combination produces as much as B. Therefore, the same combination cannot both be less and more productive at the same time. Hence two isoquants cannot intersect each other.

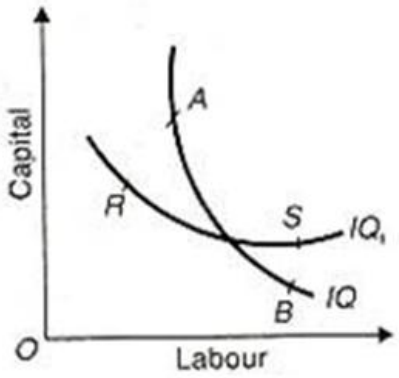


Fig.2.9

(4) Isoquants need not be parallel because the rate of substitution between two factors is not necessarily the same in all the isoquant schedules.

(5) In between two isoquants there can be a number **of isoquants showing various levels of output which the combinations of the two factors** can yield. In fact, in between the units of output 100, 200, 300, etc. represented on isoquants there can be innumerable isoquants showing 120, 150, 175,235, or any other higher or lower unit.

(6) Units of output shown on isoquants are arbitrary. The various units of output such as 100, 200, 300, etc., shown in an isoquant map are arbitrary. Any units of output such as 5, 10, 15, 20 or 1000, 2000, 3000, or any other units can he taken.

(7) No isoquant can touch either axis. If an isoquant touches X-axis, it would mean that the product is being produced with the help of labour alone without using capital at all. This is a logical absurdity for OL units of labour alone are incapable of producing anything. Similarly, OC units of capital alone cannot produce anything without the use of labour. Therefore IQ and lQ_1 cannot be isoquants, as shown in Figure 2.10.

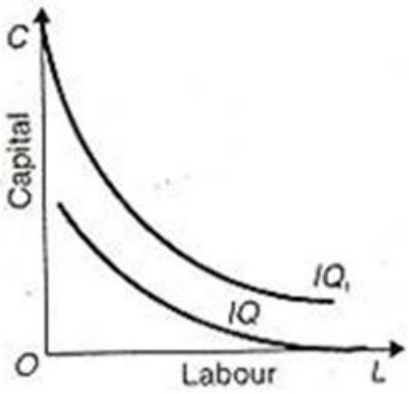


Fig.2.10

(8) Each isoquant is convex to the origin:

As more units of labour are employed to produce 100 units of the product, lesser and lesser units of capital are used. This is because the marginal rate of substitution between two factors diminishes. In Figure 2.11, in order to produce 100 units of the product, as the producer moves along the isoquant from combination A to B and to C and D, he gives up smaller and smaller units of capital for additional units of labour. To maintain the same output of 100 units, BR less of capital and relatively RC more of labour is used.

If he were producing this output with the combination D, he would be employing CT less of capital and relatively TD more of labour. Thus the isoquants are convex to the origin due to diminishing marginal rate of substitution. This fact becomes clear from successively smaller triangles below the IQ curve Δ ASB > ΔBRC > Δ CTD.

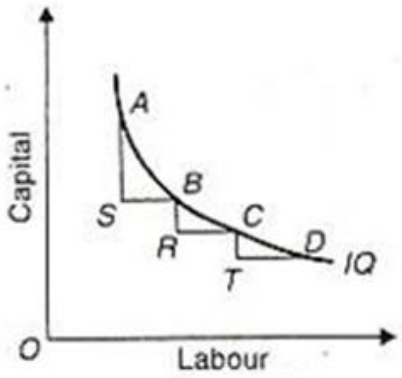


Fig.2.11

(9) Each isoquant is oval-shaped:

It is elliptical which means that at some point it begins to recede from each axis. This shape is a consequence Labour of fact that if a producer uses more of capital or more of labour or more Fig. 2.12 of both than is necessary, the total product will eventually decline.

The firm will produce only in those segments of the isoquants which are convex to the origin and lie between the ridge lines.

This is the economic region of production. In Figure 2.12, oval-shaped isoquants are shown. Curves OA and OB are the ridge lines and in between them economically feasible units of capital and labour can be employed to produce 100, 200, 300 and 400 units of the product. For example, OT units of labour and ST units of the capital can produce 100 units of the product, but the same output can be obtained by using the same quantity of labour OT and less quantity of capital VT.

Thus only an unwise entrepreneur will produce in the dotted region of the isoquant 100. The dotted segments of an isoquant are the waste- bearing segments. They form the uneconomic regions of production. In the upper dotted portion, more capital and in the lower dotted portion more labour than necessary is employed. Hence GH, JK, LM, and NP segments of the elliptical curves are the iso- quants.

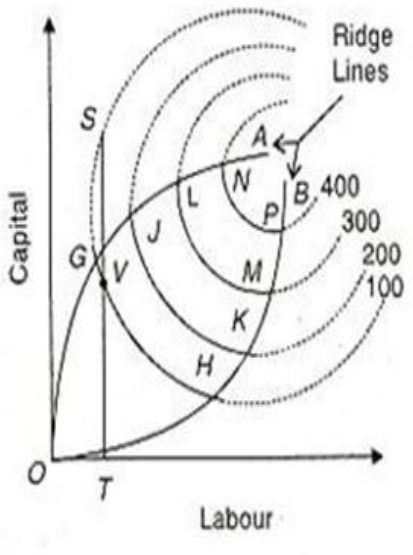


Fig.2.12

Isocost Curves:

Having studied the nature of isoquants which represent the output possibilities of a firm from a given combination of two inputs, we pass on to the prices of the inputs as represented on the isoquant map by the isocost curves. These curves are also known as outlay lines, price lines, input-price lines, factor-cost lines, constant-outlay lines, etc. Each isocost curve represents the different combinations of two inputs that a firm can buy for a given sum of money at the given price of each input.

Figure, 2.13 (A) shows three isocost curves AB, CD and EF, each represents a total outlay of 50, 75 and 100 respectively. The firm can hire OC of capital or OD of labour with Rs. 75. OC is 2/3 of OD which means that the price of a unit of labour is 1½ times less than that of a unit of capital. The line CD represents the price ratio of capital and labour. Prices of factors remaining the same, if the total outlay is raised, the isocost curve will shift upward to the right as EF parallel to CD, and if the total outlay is reduced it will shift downwards to the left as AB. The isocosts are straight lines because factor prices remain the same whatever the outlay of the firm on the two factors. The isocost curves represent the locus of all

combinations of the two input factors which result in the same total cost. If the unit cost of labour (L) is w and the unit cost of capital (C) is r, then the total cost: TC = wL + rC. The slope of the isocost line is the ratio of prices of labour and capital i.e., w/r.

The point where the isocost line is tangent to an isoquant represents the least cost combination of the two factors for producing a given output. If all points of tangency like LMN are joined by a line, it is known as an output- factor curve or least-outlay curve or the expansion path of a firm. Salvatore defines expansion path as "the locus of points of producer's equilibrium resulting from changes in total outlays while keeping factor prices constant." It shows how the proportions of the two factors used might be changed as the firm expands.

For example, in Figure 2.13 (A) the proportions of capital and labour used to produce 200 (IQ_1) units of the product are different from the proportions of these factors used to produce 300 (IQ_2) units or 100 (OQ) units at the lowest cost.

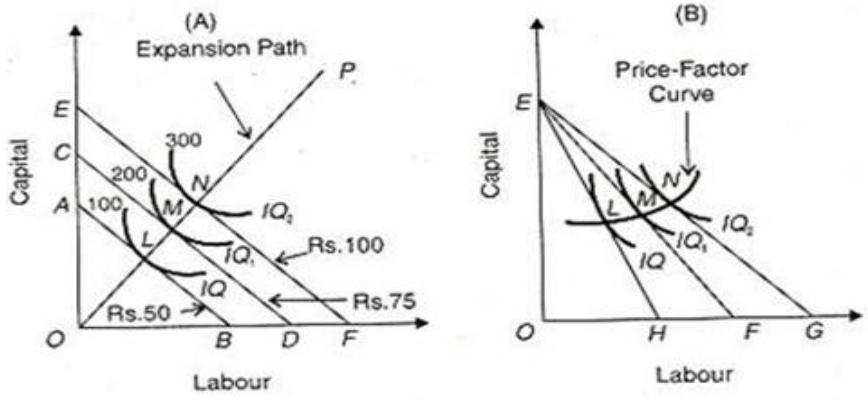


Fig.2.13

Like the price-income line in the indifference curve analysis, a relative cheapening of one of the factors to that of another will extend the isocost line to the right. If one of the factors becomes relatively dearer, the isocost line will contract inward to the left. Given the price of capital, if the price of labour falls, the isocost line EF in Panel (B) will extend to the right as EG and if the price of labour rises, the isocost line EF will contract inward to the left as EH. if the equilibrium points L, M, and N are joined by a line, it is called the price-factor curve.

The Principle of Marginal Rate of Technical Substitution:

The principle of marginal rate of technical substitution (MRTS or MRS) is based on the production function where two factors can be substituted in variable proportions in such a way as to produce a constant level of output.

The marginal rate of technical substitution between two factors C (capital) and L (labour), $MRTS_{LC}$ is the rate at which L can be substituted for C in the production of good X without changing the quantity of output. As we move along an isoquant downward to the right, each point on it represents the substitution of labour for capital.

MRTS is the loss of certain units of capital which will just be compensated for by additional units of labour at that point. In other words, the marginal rate of technical substitution of labour for capital is the slope or gradient of the isoquant at a point. Accordingly, slope = $MRTS_{LC} = - \Delta C/A L$. This can be understood with the aid of the isoquant schedule, in Table 2.2

Table 2.3: Isoquant Schedule:

Combination	Labour	Capital	$MRTS_{LC}$	Output
1	5	9	—	100
2	10	6	3:5	100
3	15	4	2:5	100
4	20	3	L;5	100

The above table shows that in the second combination to keep output constant at 100 units, the reduction of 3 units of capital requires the addition of 5 units of labour, $MRTS_{LC}$ = 3:5. In the third combination, the loss of 2 units of capital is compensated for by 5 more units of labour, and so on.

In Figure 2.14 at point B, the marginal rate of technical substitution is AS/SB, at point G, it is BT/TG and at H, it is GR/ RH.

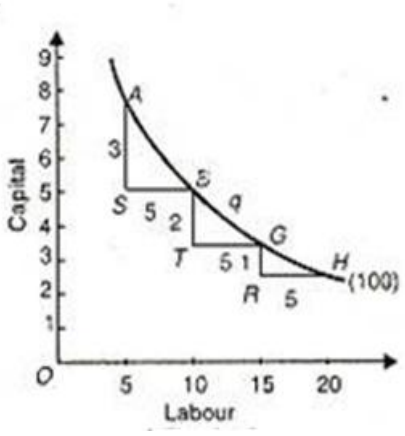


Fig.2.14

The isoquant AH reveals that as the units of labour are successively increased into the factor-combination to produce 100 units of good X, the reduction in the units of capital becomes smaller and smaller. It means that the marginal rate of technical substitution is

diminishing. This concept of the diminishing marginal rate of technical substitution (DMRTS) is parallel to the principle of diminishing marginal rate of substitution in the indifference curve technique.

This tendency of diminishing marginal substitutability of factors is apparent from Table 2.3 and Figure 2.14. The $MRTS_{LC}$ continues to decline from 3:5 to 1:5 whereas in the Figure 2.14 the vertical lines below the triangles on the isoquant become smaller and smaller as we move downward so that GR < BT < AS. Thus, the marginal rate of technical substitution diminishes as labour is substituted for capital. It means that the isoquant must be convex to the origin at every point.

The Law of Variable Proportions:

The behaviour of the law of variable proportions or of the short-run production function when one factor is constant and the other variable can also be explained in terms of the isoquant analysis. Suppose capital is a fixed factor and labour is a variable factor. In Figure 2.15, OA and OB are the ridge lines and it is in between them that economically feasible units of labour and capital can be employed to produce 100, 200, 300, 400 and 500 units of output.

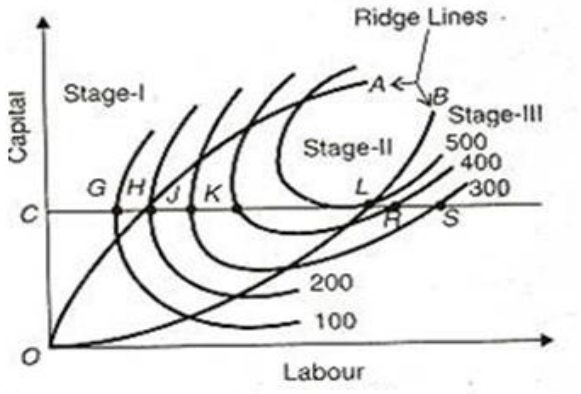


Fig.2.15

It implies that in these portions of the isoquants, the marginal product of labour and capital is positive. On the other hand, where these ridge lines cut the isoquants, the marginal product of the inputs is zero. For instance, at point H the marginal product of capital is zero, and at point L the marginal product of labour is zero. The portion of the isoquant that lies outside the ridge lines, the marginal product of that factor is negative. For instance, the marginal product of capital is negative at G and that of labour at R.

The law of variable proportions says that, given the technique of production, the application of mote and more units of a variable factor, say labour, to a fixed factor, say capital, will, until a certain point is reached, yield more than proportional increases in output, and thereafter less than proportional increases in output.

Since the law refers to increases in output, it relates to the marginal product. To explain the law, capital is taken as a fixed factor and labour as a variable factor. The isoquants show different levels of output in the figure. OC is the fixed quantity of capital which therefore forms a horizontal line CD. As we move from C to D towards the right on this line, the different points show the effects of the combinations of successively increasing quantities of labour with fixed quantity of capital OC.

To begin with, as we move from C to G to H, it shows the first stage of increasing marginal returns of the law of variable proportions. When CG labour is employed with OC capital, output is 100. To produce 200 units of output, labour is increased by GH while the amount of capital is fixed at OC.

The output has doubled but the amount of labour employed has not increased proportionately. It may be observed that GH < CG, which means that smaller additions to the labour force have led to equal increment in output. Thus C to H is the first stage of the law of variable proportions in which the marginal product increases because output per unit of labour increases as more output is produced.

The second stage of the law of variable proportions is the portion of the isoquants which lies in between the two ridge lines O A and OB. It is the stage of diminishing marginal returns between points H and L. As more labour is employed, output increases less than proportionately to the increase in the labour employed. To raise output to 300 units from 200 units, HJ labour is employed. Further, JK quantity of labour is required to raise output from 300 to 400 and KL of labour to raise output from 400 to 500.

So, to increase output by 100 units successively, more and more units of the variable factor (labour) are required to be applied along with the fixed factor (capital) , that is KL>JK>HJ. It implies that the marginal product of labour continues to decline with the employment of larger quantities to it. Thus as we more from point H to K, the effect of increasing the units of labour is that output per unit of labour diminishes as more output is produced. This is known as the stage of diminishing returns.

If labour is employed further, we are outside the lower ridge line OB and enter the third stage of the law of variable proportions. In this region which lies beyond the ridge line

OB there is too much of the variable factor (labour) in relation to the fixed factor (capital). Labour is thus being overworked and its marginal product is negative. In other words when the quantity of labour is increased by LR and RS, the output declines from 500 to 400 and to 300. This is the stage of negative marginal returns.

We arrive at the conclusion that a firm will find it profitable to produce only in the second stage of the law of variable proportions for it will be uneconomical to produce in the regions to the left or right of the ridge lines which form the first stage and the third stage of the law respectively.

Relation between Returns to Scale and Returns to a Factor (Law of Returns to Scale and Law of Diminishing Returns)

Returns to a factor and returns to scale are two important laws of production. Both laws explain the relation between inputs and output. Both laws have three stages of increasing, decreasing and constant returns. Even then, there are fundamental differences between the two laws.

Returns to a factor relate to the short period production function when one factor is varied keeping the other factor fixed in order to have more output, the marginal returns of the variable factor diminish. On the other hand, returns to scale relate to the long period production function when a firm changes its scale of production by changing one or more of its factors.

We discuss the relation between the returns to a factor (law of diminishing returns) and returns to scale (law of returns to scale) on the assumptions that:

(1) There are only two factors of production, labour and capital.

(2) Labour is the variable factor and capital is the fixed factor.

(3) Both factors are variable in returns to scale.

(4) The production function is homogeneous.

Given these assumptions, we first explain the relation between constant return to scale and returns to a variable factor in terms of Figure 2.16 where OS is the expansion path which shows constant returns to scale because the difference between the two isoquants 100 and 200 on the expansion path is equal i.e., OM = MN. To produce 100 units, the firm uses OC + OL quantities of capital and labour and to double the output to 200 units, double the quantities of labour and capital are required so that OC_1 + OL_2 lead to this output level at point N. Thus there are constant returns to scale because OM = MN.

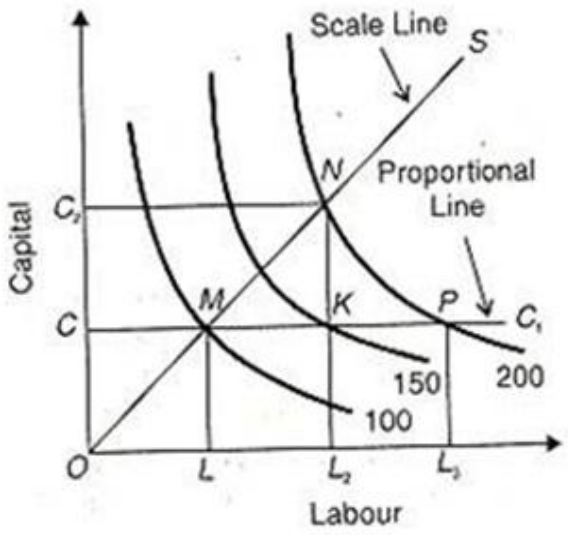


Fig.2.16

To prove that returns to the variable factor, labour, diminish, we take OC of capital as the fixed factor, represented by the CC, line. Keeping C as constant, if the amount of labour is doubled by LL_2, we reach point K which lies on a lower isoquant 150 than the isoquant 200. By keeping C constant, Ц if the output is to be doubled from 100 to 200 units, then L_3 units o of labour will be required. But $L_3 > L_2$. Thus by doubling the units of labour with constant C_2, the output less than doubles. It is 150 units at point K instead of 200 units at point P. This shows that the marginal returns of the variable factor, labour, have diminished.

As pointed out by Stonier and Hague, "So, if production function were always homogeneous of the first degree and if returns to scale were always constant, marginal physical productivity (returns) would always fall."

The relation between diminishing returns to scale and return to a variable factor is explained with the help of Figure 2.17 where OS is the expansion path which depicts diminishing returns to scale because the segment MN>OM. It means that in order to double the output from 100 to 200, more than double the amounts of both factors are required.

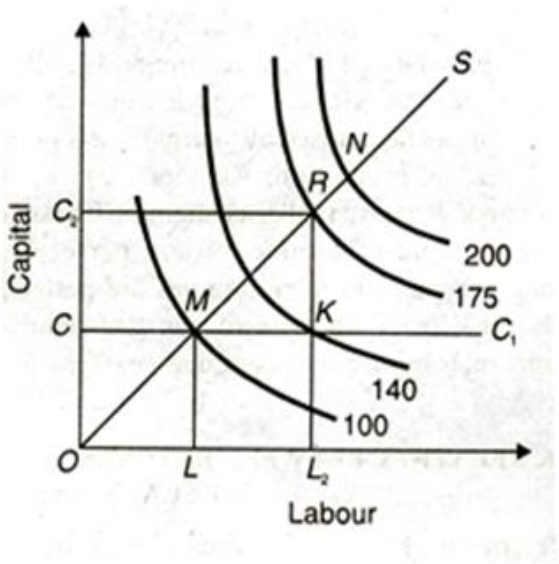


Fig.2.17

Alternatively, if both factors are doubled to OC_2+ OL_2 they lead to the lower output level isoquant 175 at point R than the isoquant 200 which shows diminishing returns to scale. If C is kept constant and the amount of variable factor, labour, is doubled by LL_2 we reach point K which lies on a still lower level of output represented by the isoquant 140. This proves that the marginal returns (or physical productivity) of the variable factor, labour, have diminished.

Now we take the relation between increasing returns to scale and returns to a variable factor. This is explained in terms of Figure 2.18 (A) and (B). In Panel (A), the expansion path OS depicts increasing returns to scale because the segment OM > MN. It means that in order to double the output from 100 to 200, less than double the amounts of both factors will be required. If C is kept constant and the amount of variable factor, labour, is doubled by LL_2 the level of output is reached at point K which shows diminishing marginal returns as represented by the lower isoquant 160 than the isoquant 200 when returns to scale are increasing.

In case the returns to scale are increasing strongly, that is, they are highly positive they will offset the diminishing marginal returns of the variable factor, labour. Such a situation leads to increasing marginal returns. This is explained in Panel (B) of Figure 2.18 where on the expansion path OS, the segment OM > MN, thereby showing increasing returns to scale. When the amount of the variable factor, labour, is doubled by LL_2 while keeping C as constant, we reach the output level K represented by the isoquant 250 which is at a higher level than the isoquant 200. This shows that the marginal returns of the variable factor, labour, have increased even when there are increasing returns to scale.

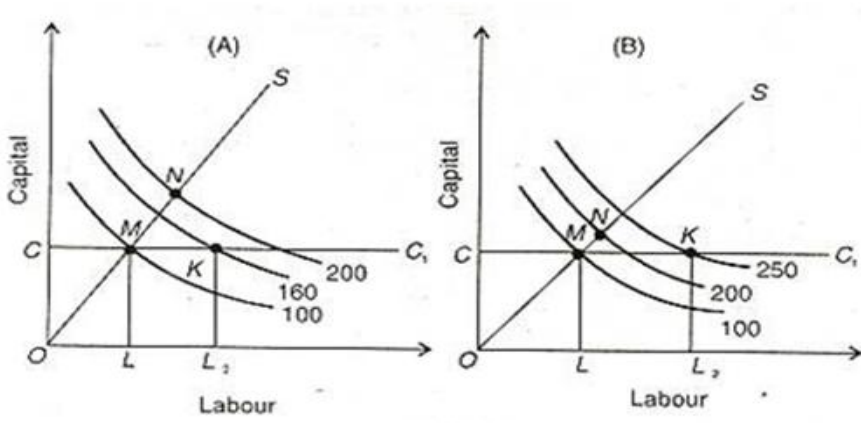


Fig.2.18

Conclusion:

It can be concluded from the above analysis that under a homogeneous production function when a fixed factor is combined with a variable factor, the marginal returns of the variable factor diminish when there are constant, diminishing and increasing returns to scale. However, if there are strong increasing returns to scale, the marginal returns of the variable factor increase instead of diminishing.

The Law of Diminishing Marginal Rate of Substitution

The marginal rate of substitution is the rate of exchange between some units of goods X and У which are equally preferred. The marginal rate of substitution of X for Y $(MRS)_{xy}$ is the amount of Y that will be given up for obtaining each additional unit of X. This rate is explained below in Table.2.4

Table 2.4

Marginal Rate of Substitution

(1) *Combination*	(2) *X*	(3) *Y*	(4) *MRS of X for Y*
L	1	9	—
M	2	6	3:1
N	3	4	2:1
P	4	3	1:1

To have the second combination and yet to be at the same level of satisfaction, the consumer is prepared to forgo 3 units of Y for obtaining an extra unit of X. The marginal rate of substitution of X for У is 3:1. The rate of substitution will then be the number of units of У for which one unit of X is a substitute. As the consumer proceeds to have additional units of X, he is willing to give away less and less units of У so that the marginal rate of substitution falls from 3:1 to 1:1 in the fourth combination (Col. 4).

In Fig. 2.19 above at point M on the indifference curve I, the consumer is willing to give up 3 units of У to get an additional unit of X. Hence, MRS_{xy} =3. As he moves along the curve from M to N, MRS_{xy}, =2. When the consumer moves downwards along the indifference curve, he acquires more of X and less of Y. The amount of У he is prepared to give up to get additional units of X becomes smaller and smaller.

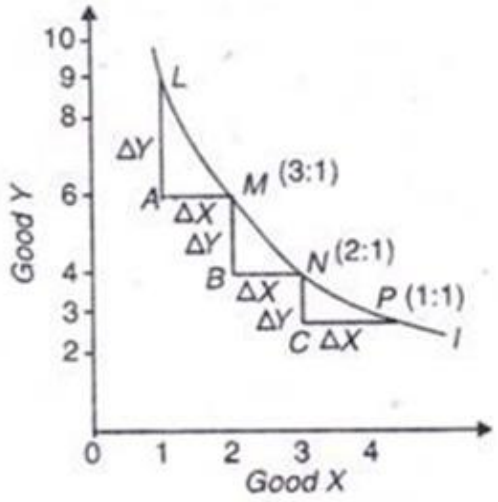


Fig.2.19

The marginal rate of substitution of X for У (MRS_{xy}) is, in fact, the slope of the curve at a point on the indifference curve, such as points M, N or P in Fig.2.19. Thus MRS_{xy} = ΔY/ΔX

It means that the MRS_{xy} is the ratio of change in good Y to a given change in X. In the figure, there are three triangles on the I_1 curve whose vertical sides LA, MB and NC represent AУ which diminish and the horizontal sides AM, BN and CP signify AX which remains the same.

At point M, MRS_{xy} = LA/AM at N it is MB/BN. This also shows that as the consumer moves downwards along the curve, he possesses additional units of X, and gives up lesser and lesser units of Y, i.e., the MRS_{xy} diminishes. It is due to this law of diminishing MRS that an indifference curve is convex to the origin.

Exceptions of DMRS Law:

However, this law is not applicable in the case of perfect substitutes and complementary goods. These are the exceptions of the DMRS law whereby an indifference curve is not convex to the origin but is a straight line and L- shaped.

1. Straight Line Indifference Curve:

If MRS of X for Y or Y for X is diminishing, the indifference curve must be convex to the origin. If it is constant, the indifference curve will be a straight line sloping downwards to the right at a 45° angle to either axis, as in Fig.4.20. This is the case of perfect substitute goods like Lux and Godrej soap, Tata and Brooke Bond Tea, etc.

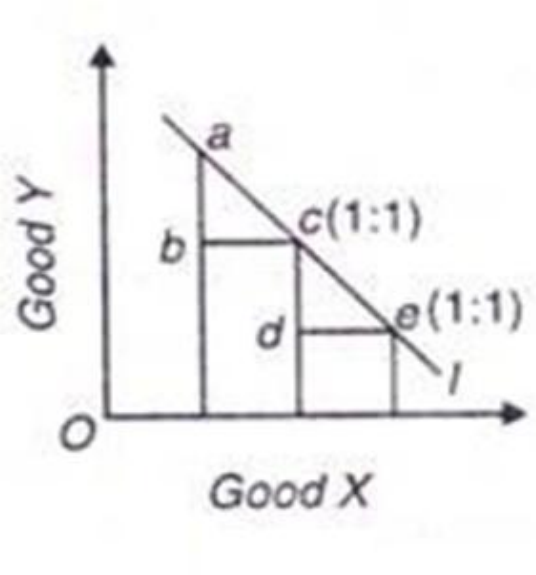


Fig.2.20

When a consumer substitutes Lux for Godrej or vice versa, his satisfaction remains the same. Thus MRS_y for perfect substitutes is constant, i.e., MRS_{xy}

This is clear from equal triangles, Δabc = Δcde, below the I curve.

L -Shaped Indifference Curve:

When two goods are used simultaneously in a constant ratio such as left shoe and right shoe, the indifference curve is L-Shaped or of 90° angle. Such a curve is for perfect complementary goods and their MRS is always zero MRS_{xy} =0. Figure.2.21 shows preferences of consumer for left and right shoes. Since shoes are perfect complementary, at point B of I curve an additional right shoe will not increase his satisfaction 90° until he gets another left shoe of the same size.

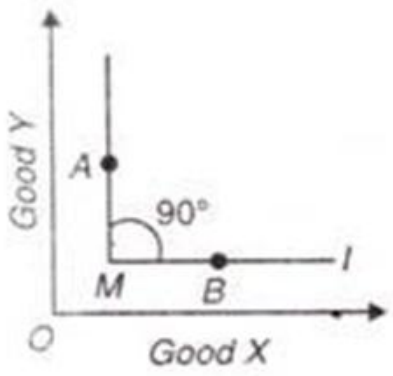


Fig.2.21

The same is the case at point I A where he gets an additional left shoe without another right shoe. So the consumer will get full satisfaction at point M of the I curve where he purchases a left and right shoe of the same size and his $MRS_{xy} = 0$. This is because he cannot substitute right and left shoes.

Cobb-Douglas Production Function:

Charles W. Cobb and Paul H. Douglas studied the relationship of inputs and outputs and formed an empirical production function, popularly known as Cobb-Douglas production function. Originally, C-D production function applied not to the production process of an individual firm but to the whole of the manufacturing production.

The Cobb-Douglas production function is expressed by

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

where Q is output and L and A' are inputs of labour and capital respectively. A, α and β are positive parameters where $\alpha > 0$, $\beta > 0$. The equation tells that output depends directly on L and K and that part of output which cannot be explained by L and K is explained by A which is the 'residual', often called technical change.

The marginal products of labour and capital are the functions of the parameters A, α and β and the ratios of labour and capital inputs. That is,

$MP_L = \partial Q/\partial L = \alpha AL^{\alpha-1}K^{\beta}$

$MP_K = \partial Q/\partial K = \beta AL^{\alpha}K^{\beta-1}$

The two parameters a and P taken together measure the degree of the homogeneity of the function.

In other words, this function characterises the returns to scale thus:

$\alpha + \beta > 1$: Increasing returns to scale

$\alpha + \beta = 1$: Constant returns to scale

$\alpha + \beta < 1$: Decreasing returns to scale.

Although the C-D production function is a multiplicative type and is non-linear in its general form, it can be transferred into linear function by taking it in its logarithmic form. That is why, this function is also known as log linear function, which is

$\text{Log } Q = \log A + a \log L + p \log K$

It is easier to compute C-D function when expressed in log linear form.

Properties of C-D Production Function:

(i) There are constant returns to scale.

(ii) Elasticity of substitution is equal to one.

(iii) A and p represent the labour and capital shares of output respectively.

(iv) A and p are also elasticities of output with respect to labour and capital respectively.

(v) If one of the inputs is zero, output will also be zero.

(vi) The expansion path generated by C-D function is linear and it passes through the origin.

(vii) The marginal product of labour is equal to the increase in output when the labour input is increased by one unit.

(viii) The average product of labour is equal to the ratio between output and labour input.

(ix) The ratio α / β measures factor intensity. The higher this ratio, the more labour intensive is the technique and the lower is this ratio and the more capital intensive is the technique of production.

Importance of C-D Production Function

(i) It suits to the nature of all industries.

(ii) It is convenient in international and inter-industry comparisons.

(iii) It is the most commonly used function in the field of econometrics.

(iv) It can be fitted to time series analysis and cross section analysis.

(v) The function can be generalised in the case of 'n' factors of production.

(vi) The unknown parameters a and p in the function can be easily computed.

(vii) It becomes linear function in logarithm.

(viii) It is more popular in empirical research.

Limitations of C-D Production Function

It has the following limitations:

(i) The function includes only two factors and neglects other inputs.

(ii) The function assumes constant returns to scale.

(iii) There is the problem of measurement of capital which takes only the quantity of capital available for production.

(iv) The function assumes perfect competition in the factor market which is unrealistic.

(v) It does not fit to all industries.

(vi) It is based on the substitutability of factors and neglects complementarity of factors.

(vii) The parameters cannot give proper and correct economic implication.

Constant Elasticity of Substitution Production Function:

The CES production function is otherwise known as Homohighplagic production function. Arrow, Chenery, Minhas and Solow have developed the Constant Elasticity of Substitution (CES) function. This function consists of three variables Q, K and L, and three parameters A, a and 0. It may be expressed in the form

$$Q = A\,[\alpha\, C^{-\Theta} + (1-\alpha)\, L^{-\Theta}]^{-1/\Theta}$$

where Q is the total output, K is capital, and L is labour. A is the efficiency parameter indicating the state of technology and organisational aspects of production. It shows that with technological and/or organisational changes, the efficiency parameter leads to a shift in the production function, a (alpha) is the distribution parameter or capital intensity factor coefficient concerned with the relative factor shares in the total output, and 0 (theta) is the substitution parameter which determines the elasticity of substitution. And $A > 0$; $0 < \alpha < 1$; $\Theta > -1$.

In the CES production function, the elasticity substitution is constant and not necessarily equal to unity.

Mukherji has generated the CES function by introducing more than two inputs.

Properties of CES Production Function:

1. The CES function is standardised of degree one. Thus like Cobb-Douglas production function, the ES function displays constant returns to scale.
2. In the CES production function, the average and marginal products in the variables C and L are standardised of degree zero like all linear standardised production functions.

3. From the above property the incline of an Isoquant i.e. Marginal rate of Technical Substitution MRTS of capital for labour can be represented convex to the origin.
4. As a consequence of the above, if L and C are substitutable (infinity) for each other an increase in C will require less of L for a given productivity. As a consequent, the MP of L will increase. Thus the MP of an input will increase when the other input is increased.

Merits

It has the following merits and they are listed below.

1. The CES function is a standardised of grade one, it is more general, it covers all type of returns
2. This function takes account of a number of parameters
3. This functions takes account of raw materials among its inputs
4. CES functions are very easy to approximate and are free from impractical postulations.

Limitations

1. This production function regards only two inputs. It can be comprehensive to more than two units. However, it becomes very complex and intricate arithmetically to use it for more than two inputs.
2. The distribution parameter or capital intensity factor co-efficient α is not dimensionless.
3. If data are fitted to the CES function, the value of the competence a parameter cannot be made independent of the other units representing in this production function.
4. If the CES function is used to explain the production of a firm, it cannot be used to explain the aggregate production function of all the firms in the industry. Thus it involves the problem of aggregation of production function of diverse firms in the industry.
5. It suffers from the drawback that elasticity of substitution amidst any part of inputs is the same which does not materialize to be realistic.
6. In approximation of the parameters of CES production function we may come across a large number of problems like choice of exogenous variables, judgment procedure and the problem of multi collinear.
7. There is little possibility of identifying the production function under technological change.

Conclusion

Despite these limits, CES functions is constructive in its application to prove Euler's theorem, to reveal constant returns to scale to show that average and marginal products are standardised degree zero and to determine the elasticity of substitution.

The Traditional Theory of Costs:

The traditional theory of costs analyses the behaviour of cost curves in the short run and the long run and arrives at the conclusion that both the short run and the long run curves are U-shaped but the long-run cost curves are flatter than the short-run cost curves.

(A) Firm's Short-Run Cost Curves:

The short run is a period in which the firm cannot change its plant, equipment and the scale of organisation. To meet the increased demand, it can raise output by hiring more labour and raw materials or asking the existing labour force to work overtime.

Short-Run Total Costs:

The scale of organisation being fixed, the short-run total costs are divided into total fixed costs and total variable costs.

TC = TFC + TVC

Total Costs or TC

Total costs are the total expenses incurred by a firm in producing a given quantity of a commodity. They include payments for rent, interest, wages, taxes and expenses on raw materials, electricity, water, advertising, etc.

Total Fixed Costs or TFC:

Are those costs of production that do not change with output. They are independent of the level of output. In fact, they have to be incurred even when the firm stops production temporarily. They include payments for renting land and buildings, interest or borrowed money, insurance charges, property tax, depreciation, maintenance expenditures, wages and salaries of the permanent staff, etc. They are also called overhead costs.

Total Variable Costs or TVC:

Are those costs of production that change directly with output. They a rise when output increases, and fall when output declines. They include expenses on raw materials, power, water, taxes, hiring of labour, advertising etc., They are also known as direct costs.

The relation between total costs, variable costs and fixed costs is presented in Table 1, where column (1) indicates different levels of output from 0 to 10 units. Column (2) indicates

that total fixed costs remain at Rs. 300 at all levels of output. Column (3) shows total variable costs which are zero when output is nothing and they continue to increase with the rise in output.

In the beginning they rise quickly, and then they slow down as the firm enjoys economies of large scale production with further increases in output and later on due to diseconomies of production, the variable costs start rising rapidly. Column (4) relates to total costs which are the sum of columns (2), and (3) i.e., TC – TFC + TVC. Total costs vary with total variable costs when the firm starts production.

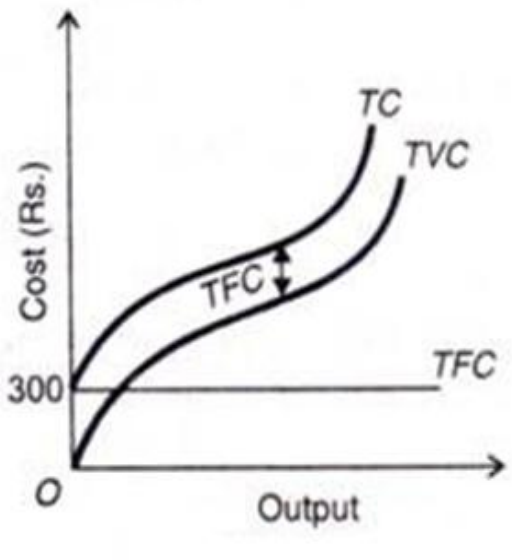


Fig.2.22

The curves relating to these three total costs are shown diagrammatically in Figure 2.22. The TC curve is a continuous curve which shows that with increasing output total costs also increase. This curve cuts the vertical axis at a point above the origin and rises continuously from left to right. This is because even when no output is produced, the firm has to incur fixed costs.

Table 2.5

COST FUNCTION IN THE SHORT-RUN

TO (1)	*TFC* (2)	*TVC* (3)	*TC* (4) (2+3)	*AFC* (5) (2÷1)	*AVC* (6) (3÷1)	*ATC* (7) (5+6) or(4÷1)	*MC* (8) (from 4)
	Rs	Rs	Rs	Rs	Rs	Rs	Rs
0	300	0	300	300	0	300	—
1	300	300	600	300	300	600	300
2	300	400	700	150	200	350	100
3	300	450	750	100	150	250	50
4	300	500	800	75	125	200	50
5	300	600	900	60	120	180	100
6	300	720	1020	50	120	170	120
7	300	890	1190	42.9	127.1	170	170
8	300	1100	1400	37.5	137.5	175	210
9	300	1350	1650	33.3	150	183.3	470
10	300	2000	2300	30	200	230	650

The TFC curve is shown as parallel to the output axis because total fixed costs are the same (Rs.300) whatever the level of output. The TVC curve has an inverted-S shape and starts from the origin O because when output is zero, the TVCs are also zero. They increase as output increases.

So long as the firm is using less variable factors in proportion to the fixed factors, the total variable costs rise at a diminishing rate. But after a point, with the use of more variable factors in proportion to the fixed factors, they rise steeply because of the application of the law of variable proportions. Since the TFC curve is a horizontal straight line, the TC curve follows the TVC curve at an equal vertical distance.

Short-Run Average Costs:

In the short run analysis of the firm, average costs are more important than total costs. The units of output that a firm produces do not cost the same amount to the firm. But they must be sold at the same price. Therefore, the firm must know the per unit cost or the average cost. The short-run average costs of a firm are the average fixed costs, the average variable costs, and the average total costs.

Average Fixed Costs equal total fixed costs at each level of output divided by the number of units produced:

AFC = TFC /Q

The average fixed costs diminish continuously as output increases. This is natural because when constant total fixed costs are divided by a continuously increasing unit of output, the result is continuously diminishing average fixed costs. Thus the AFC curve is a

downward sloping curve which approaches the quantity axis without touching it, as shown in Figure 3. It is a rectangular hyperbola.

Short-Run Average Variable Costs (or SAVC) equal total variable costs at each level of output divided by the number of units produced:

SAVC = TVC/Q

The average variable costs first decline with the rise in output as larger quantities of variable factors is applied to fixed plant and equipment. But eventually they begin to rise due to the law of diminishing returns. Thus the SAVC curve is U-shaped, as shown in Figure 2.23.

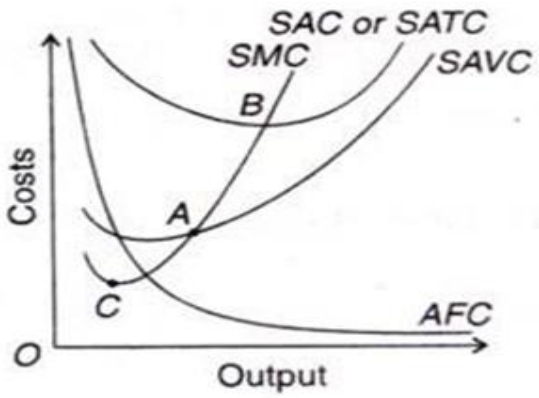


Fig.2.23

Short-Run Average Total Costs (or SATC or SAC) are the average costs of producing any given output.

They are arrived at by dividing the total costs at each level of output by the number of units produced:

SAC or SATC = TC/Q TFC/Q + TVC/Q = AFC+ AVC

Average total costs reflect the influence of both the average fixed costs and average variable costs. At first average total costs are high at low levels of output because both average fixed costs and average variable costs are large. But as output increases, the average total costs fall sharply because of the steady decline of both average fixed costs and average variable costs till they reach the minimum point.

This results from the internal economies, from better utilisation of existing plant, labour, etc. The minimum point B in the figure represents optimal capacity. As production is increased after this point, the average total costs rise quickly because the fall in average fixed costs is negligible in relation to the rising average variable costs.

The rising portion of the SAC curve results from producing above capacity and the appearance of internal diseconomies of management, labour, etc. Thus the SAC curve is U-shaped, as shown in Figure 2.23.

Why is SAC curve U-shaped?

The U-shape of the SAC curve can also be explained in terms of the law of variable proportions. This law tells that when the quantity of one variable factor is changed while keeping the quantities of other factors fixed, the total output increases but after some time it starts declining.

Machines, equipment and scale of production are the fixed factors of a firm that do not change in the short run. 'On the other hand, factors like labour and raw materials are variable. When increasing quantities of variable factors are applied on the fixed factors, the law of variable proportions operates.

When, say the quantities of a variable factor like labour are increased in equal quantities, production rises till fixed factors like machines, equipment, etc. are used to their maximum capacity. In this stage, the average costs of the firm continue to fall as output increases because it operates under increasing returns.

Due to the operation of the law of increasing returns when the variable factors are increased further, the firm is able to work the machines to their optimum capacity. It produces the optimum output and its average costs of production will be the minimum which is revealed by the minimum point of the SAC curve, point B in Figure 2.23.

It the firm tries to raise output after this point by increasing the quantities of the variable factors, the fixed factors like machines would be worked beyond their capacity. This would lead to diminishing returns. The average costs will start rising rapidly. Hence, due to the working of the law of variable proportions the short-run AC curve is U-shaped.

Short Run Marginal Cost:

A fundamental concept for the determination of the exact level of output of a firm is the marginal cost.

Marginal cost is the addition to total cost by producing an additional unit of output:

$SMC = \Delta TC/\Delta Q$

Algebraically, it is the total cost of n + 1 units minus the total cost of n units of output $MC_n = TC_{n+1} - TC_n$. Since total fixed costs do not change with output, therefore, marginal fixed cost is zero. So marginal cost can be calculated either from total variable costs or total costs. The result would be the same in both the cases. As total variable costs or total costs first fall and

then rise, marginal cost also behaves in the same way. The SMC curve is also U-shaped, as shown in Figure 2.23.

Conclusion:

Thus the short-run cost curves of a firm are the SAVC curve, the AFC curve, the SAC curve and the SMC curve. Out of these four curves, the AFC curve is insignificant for the determination of the firm s exact output and is, therefore, generally neglected.

(B) Firm's Long-Run Cost Curves:

In the long run, there are no fixed factors of production and hence no fixed costs. The firm can change its size or scale of plant and employ more or less inputs. Thus in the long run all factors are variable and hence all costs are variable.

The long run average total cost or LAC curve of the firm shows the minimum average cost of producing various levels of output from all-possible short-run average cost curves (SAC). Thus the LAC curve is derived from the SAC curves. The LAC curve can be viewed as a series of alternative short-run situations into any one of which the firm can move.

Each SAC curve represents a plant of a particular size which is suitable for a particular range of output. The firm will, therefore, make use of the various plants up to that level where the short-run average costs fall with increase in output. It will not produce beyond the minimum short-run average cost of producing various outputs from all the plants used together.

Let there be three plants represented by their short-run average cost curves SAC_1 SAC_2 and SAC_3 in Figure 4. Each curve represents the scale of the firm. SAC_1depicts a lower scale while the movement from SAC_2 to SA C_1 shows the firm to be of a larger size. Given this scale of the firm, it will produce up to the least cost per unit of output. For producing ON output, the firm can use SAC_1or SAC_2 plant.

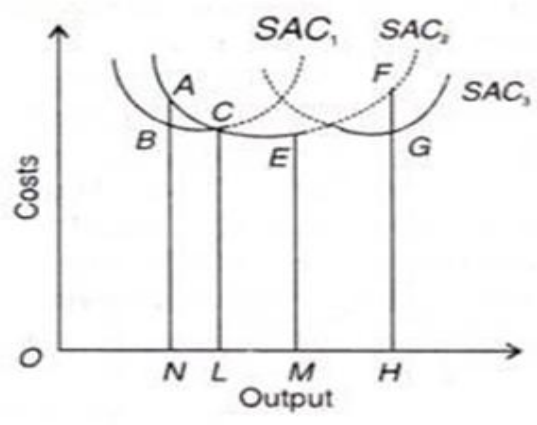


Fig.2.24

The firm will, however, use the scale of plant represented by SAC_3since the average cost of producing ON output is NB which is less than NA, the cost of producing this output on the SAC_2 plant. If the firm is to produce OL output, it can produce at either of the two plants. But it would be advantageous for the firm to use the plant SA C_2 for the OL level of output.

But it would be more profitable for the firm to produce the larger output OM at the lowest average cost ME from this plant. However, for output OH, the firm would use the SAC_1 plant where the average cost HG is lower than HF of the SAC_2 plant. Thus in the long-run in order to produce any level of output the firm will use that plant which has the minimum unit cost.

If the firm expands its scale by the three stages represented by $SAC_1SAC_2$and SAC_3 curves, the thick wave-like portions of these curves form the long-run average cost curve. The dotted portions of these SAC curves are of no consideration during the long run because the firm would change the scale of plant rather than operate on them.

But the long-run average cost curve LAC is usually shown as a smooth curve fitted to the SAC curves so that it is tangent to each of them at some point, as shown in Figure 2.25, where $SAC_{1,}SAC_{2,}$ SAC_3, SAC_4 and $SAC_5$are the short-run cost curves. It is tangent to all the SAC curves but only to one at its minimum point.

The LAC is tangent to the lowest point E of the curve SAC_3 in Figure2.25 at OQ optimum output. The plant SAC_3 which produces this OQ optimum output at the minimum cost QE is the optimum plant, and the firm producing this optimum output at the minimum cost with this optimum plant is the optimum firm. If the firm produces less than the optimum output OQ, it is not working its plant to full capacity and if it produces beyond it is overworking its plants. In both the cases, the plants SAC_2 and SAC_4 have higher average costs of production than the plant SAC_3

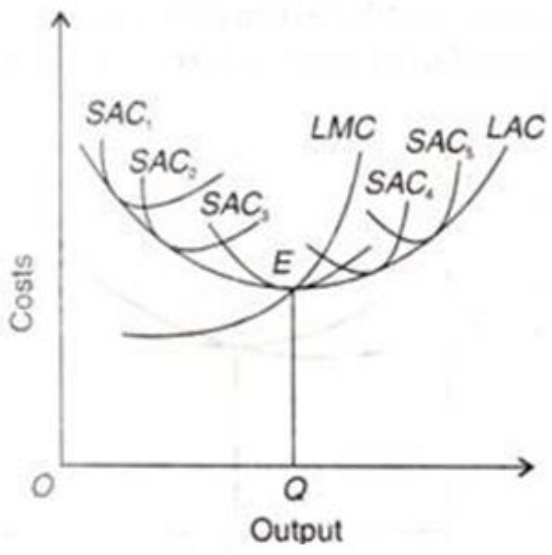


Fig.2.25

The LAC curve is known as the "envelope" curve because it envelopes all the SAC curves. According to Prof. Chamberlin, "It is composed of plant curves; it is the plant curve. But it is better to call it a "planning" curve because the firm plans to expand its scale of production over the long run."

The long-run marginal cost (LMC) curve of the firm intersects SAC_1 and LAC curves at the minimum point E.

LAC Curve Flatter than SAC Curve:

Though the long-run average cost (LAC) curve is U-shaped, yet it is flatter than the short-run average cost (SAC) curve. It means that the LAC curve first falls slowly and then rises gradually after a minimum point is reached.

1. Initially, the LAC gradually slopes downwards due to the availability of certain economies of scale like the economical use of indivisible factors, increased specialisation and the use of technologically more efficient machines or factors. The returns to scale increase because of the indivisibility of factors of production. When a business unit expands, the returns to scale increase because the indivisible factors are employed to their maximum capacity. Further, as the firm expands, it enjoys internal economies of production. It may be able to install better machines, sell its products more easily, borrow money cheaply, procure the services of more efficient manager and workers, etc. All these economies help in increasing the returns to scale more than proportionately.
2. After the minimum point of the long-run average cost is reached, the LAC curve may flatten out over a certain range of output with the expansion of the scale of

production. In such a situation, the economies and diseconomies balance each other and the LAC curve has a disc base.

3. With further expansion of scale, the diseconomies like the difficulties of coordination, management, labour and transport arise which more than counterbalance the economies so that the LAC curve begins to rise. This happens when the indivisible factors become inefficient and less productive due to the over expansion of the scale of production. Moreover, when supervision and coordination become difficult, the per unit cost increases. To these internal diseconomies are added external diseconomies of scale. These arise from higher factor prices or from diminishing productivities of factors. As the industry continues to expand, the demand for skilled labour, land, capital, etc. rises. Transport and marketing difficulties also emerge. Prices of raw materials go up. All these factors lead to diminishing returns to scale and tend to raise costs.

Conclusion:

The LAC curves first falls and then rises more slowly than the SAC curve because in the long run all costs become variable and few are fixed. The plant and equipment can be altered and adjusted to the output. The existing factors can be worked fully and more efficiently so that both the average fixed costs and average variable costs are lower in the long run than in the short run. That is why, the LAC curve is flatter than the SAC curve.

Similarly, the LMC curve is flatter than the SMC curve because all costs are variable and there are few fixed costs. In the short-run, the marginal cost is related to both the fixed and variable costs. As a result, the SMC curve falls and rises more swiftly than the LMC curve. The LMC curve bears the usual relation to the LAC curve. It first falls and is below the LAC curve. Then rises and cuts the LAC curve at its lowest point E and is above the latter throughout its length, as shown in Figure 2.26.

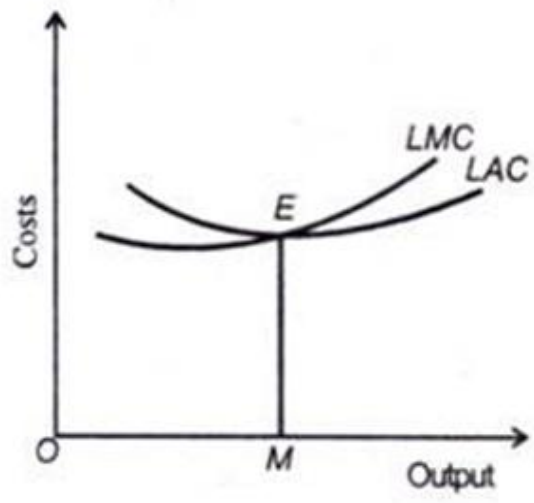


Fig.2.26

The modem theory

The modem theory of costs differs from the traditional theory of costs with regard to the shapes of the cost curves. In the traditional theory, the cost curves are U-shaped. But in the modem theory which is based on empirical evidences, the short-run SAVC curve and the SMC curve coincide with each other and are a horizontal straight line over a wide range of output. So far as the LAC and LMC curves are concerned, they are L-shaped rather than U-shaped. We discuss below the nature of short- run and long-run cost curves according to the modem theory.

(1) Short-Run Cost Curves:

As in the traditional theory, the short-run cost curves in the modem theory of costs are the AFC, SAVC, SAC and SMC curves. As usual, they are derived from the total costs which are divided into total fixed costs and total variable costs.

But in the modem theory, the SAVC and SMC curves have a saucer-type shape or bowl-shape rather than a U-shape. As the AFC curve is a rectangular hyperbola, the SAC curve has a U-shape even in the modem version. Economists have investigated on the basis of empirical studies this behaviour pattern of the short-run cost curves.

According to them, a modern firm chooses such a plant which it can operate easily with the available variable direct factors. Such a plant possesses some reserve capacity and much flexibility. The firm installs this type of plant in order to produce the maximum rate of output over a wide range to meet any increase in demand for its product.

The saucer-shaped SAVC and SMC curves are shown in Figure 2.27. To begin with, both the curves first fall upto point A and the SMC curve lies below the SAVC curve. “The

falling part of the SAVC shows the reduction in costs due to the better utilisation of the fixed factor and the consequent increase in skills and productivity of the variable factor (labour).

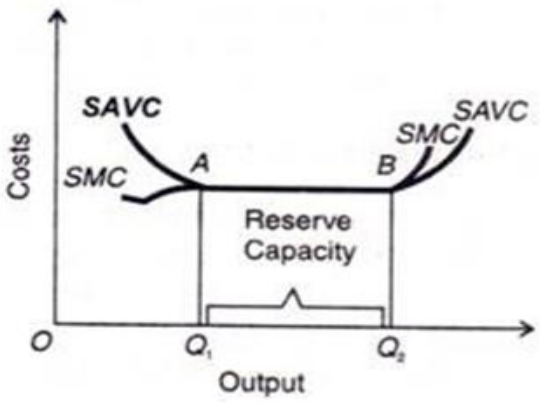


Fig.2.27

With better skills, the wastes in raw materials are also being reduced and a better utilisation of the whole plant is reached." So far as the flat stretchz of the saucer-shaped SAVC curve over $Q_{:1}Q_2$ range of output is concerned, the empirical evidence reveals that the operation of a plant within this wide range exhibits constant returns to scale.

The reason for the saucer-shaped SAVC curve is that the fixed factor is divisible. The SAV costs are constant over a large range, up to the point at which all of the fixed factor is used. Moreover, the firm's SAV costs tend to be constant over a wide range of output because there is no need to depart from the optimal combination of labour and capital in those plants that are kept in operation.

Thus there is a large range of output over which the SAVC curve will be flat. Over that range, SMC and SAVC are equal and are constant per unit of output. The firm will, therefore, continue to produce within Q_1Q_2 reserve capacity of the plant, as shown in Fig. 2.27.

After point B, both the SAVC and SMC curves start rising. When the firm departs from its normal or the load factor of the plant in order to obtain higher rates of output beyond Q_2, it leads to higher SAVC and SMC. The increase in costs may be due to the overtime operations of the old and less efficient plant leading to frequent breakdowns, wastage of raw materials, reduction in labour productivity and increase in labour cost due to overtime operations. In the rising portion of the SAVC curve beyond point B, the SMC curve lies above it.

The short-run average total cost curve (SATC or SAC) is obtained by adding vertically the average fixed cost curve (AFC) and the SAVC curve at each level of output.

The SAC curve, as shown in Figure2.28, continues to fall up to the OQ level of output at which the reserve capacity of the plant is fully exhausted.

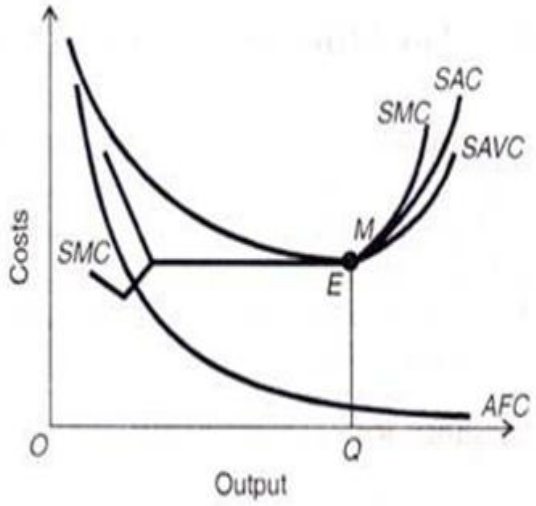


Fig.2.28

Beyond that output level, the SAC curve rises as output increases. The smooth and continuous fall in the SAC curve upto the OQ level of output is due to the fact that the AFC curve is a rectangular hyperbola and the SAVC curve first falls and then becomes horizontal within the range of reserve capacity. Beyond the OQ output level, it starts rising steeply. But the minimum point M of the SAC curve where the SMC curve intersects it, is to the right of point E of the SAVC curve. This is because the SAVC curve starts rising steeply from point E while the AFC curve is falling at a very low rate.

(2) Long-Run Cost Curves:

Empirical evidence about the long-run average cost curve reveals that the LAC curve is L-shaped rather than U-shaped. In the beginning, the LAC curve rapidly falls but after a point "the curve remains flat, or may slope gently downwards, at its right-hand end." Economists have assigned the following reasons for the L-shape of the LAC curve.

1. Production and Managerial Costs:

In the long run, all costs being variable, production costs and managerial costs of a firm are taken into account when considering the effect of expansion of output on average costs. As output increases, production costs fall continuously while managerial costs may rise at very large scales of output. But the fall in production costs outweighs the increase in managerial costs so that the LAC curve falls with increases in output. We analyse the behaviour of production and managerial costs in explaining the L-shape of the LAC curve.

Production Costs:

As a firm increases its scale of production, its production costs fall steeply in the beginning and then gradually. This is due to the technical economies of large scale production enjoyed by the firm. Initially, these economies are substantial. But after a certain level of output when all or most of these economies have been achieved, the firm reaches the minimum optimal scale or mini mum efficient scale (MES).

Given the technology of the industry, the firm can continue to enjoy some technical economies at outputs larger than the MES for the following reasons:

(a) from further decentralisation and improvement in skills and productivity of labour; (b) from lower repair costs after the firm reaches a certain size; and

(c) by itself producing some of the materials and equipment cheaply which the firm ne eds instead of buying them from other firms.

Managerial Costs:

In modern firms, for each plant there is a corresponding managerial set-up for its smooth operation. There are various levels of management, each having a separate management technique applicable to a certain range of output. Thus, given a managerial set-up for a plant, its managerial costs first fall with the expansion of output and it is only at a very large scale output, they rise very slowly.

To sum up, production costs fall smoothly and managerial costs rise slowly at very large scales of output. But the fall in production costs more than offsets the rise in managerial costs so that the LAC curve falls smoothly or becomes flat at very large scales of output, thereby giving rise to the L-shape of the LAC curve.

In order to draw such an LAC curve, we take three short-run average cost curves SAC_1 SA C_2, and SAC_3representing three plants with the same technology in Figure 9. Each SAC curve includes production costs, managerial costs, other fixed costs and a margin for normal profits. Each scale of plant (SAC) is subject to a typical load factor capacity so that points A, B and C represent the minimal optimal scale of output of each plant.

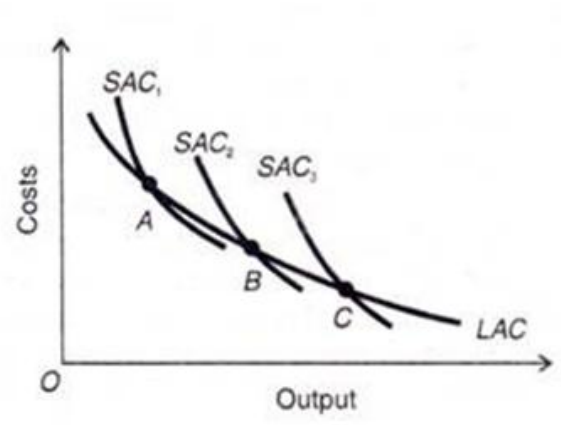


Fig.2.29

By joining all such points as A, B and C of a large number of SACs, we trace out a smooth and continuous LAC curve, as shown in Figure 2.29. This curve does not turn up at very large scales of output. It does not envelope the SAC curves but intersects them at the optimal level of output of each plant.

2. Technical Progress:

Another reason for the existence of the L-shaped LAC curve in the modern theory of costs is technical progress. The traditional theory of costs assumes no technical progress while explaining the U-shaped LAC curve. The empirical results on long-run costs conform the widespread existence of economies of scale due to technical progress in firms.

The period, between which technical progress has taken place, the long-run average costs show a falling trend. The evidence of diseconomies is much less certain. So an upturn of the LAC at the top end of the size scale has not been observed. The L-shape of the LAC curve due to technical progress is explained in Figure 2.30.

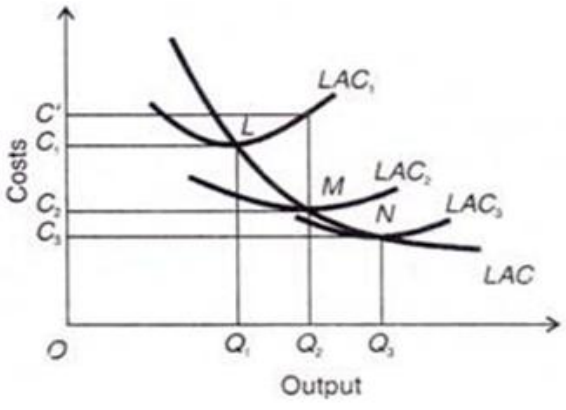


Fig.2.30

Suppose the firm is producing OQ_1 output on LAC_1curve at a per unit cost of OC_1 If there is an increase in demand for the firm's product to OQ_2,with no change in technology,

the firm will produce OQ_2 output along the LAC_1 curve at a per unit cost of OC_2. If, however, there is technical progress in the firm, it will install a new plant having LAC_2 as the long-run average cost curve. On this plant, it produces OQ_2 output at a lower cost OC_2 per unit.

Similarly, if the firm decides to increase its output to OQ_3 to meet further rise in demand technical progress may have advanced to such a level that it installs the plant with the LAC_3 curve. Now it produces OQ_3output at a still lower cost OC_3 per unit. If the minimum points, L, M and N of these U- shaped long-run average cost curves LAC_1, LAC_2 and LAC_3 are joined by a line, it forms an L-shaped gently sloping downward curve LAC.

3. Learning:

Another reason for the L-shaped long- run average cost curve is the learning process. Learning is the product of experience. If experience, in this context, can be measured by the amount of a commodity produced, then higher the production is, the lower is per unit cost.

The consequences of learning are similar to increasing returns. First, the knowledge gained from working on a large scale cannot be forgotten. Second, learning increases the rate of productivity. Third, experience is measured by the aggregate output produced since the firm first started to produce the product.

Learning-by-doing has been observed when firms start producing new products. After they have produced the first unit, they are able to reduce the time required for production and thus reduce their per unit costs. For example, if a firm manufactures airframes, the fall observed in long-run average costs is a function of experience in producing one particular kind of airframe, in general.

One can, therefore not airframes, draw a "learning curve" which relates cost per airframe to the aggregate number of airframes manufactured so far, since the firm started manufacturing them. Figure2.31 shows a learning curve LAC which relates the cost of producing a given output to the total output over the entire time period.

Growing experience with making the product leads to falling costs as more and more of it is produced. When the firm has exploited all learning possibilities, costs reach a minimum level, M in the figure 2.31. Thus, the LAC curve is L-shaped due to learning by doing.

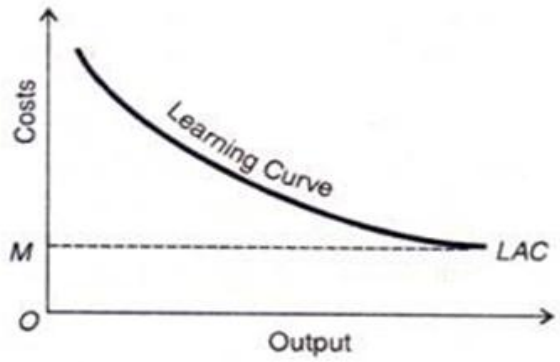


Fig.2.31

Relation between LAC and LMC Curves:

In the modern theory of costs, if the LAC curve falls smoothly and continuously even at very large scales of output, the LMC curve will lie below the LAC curve throughout its length, as shown in Figure 2.32.

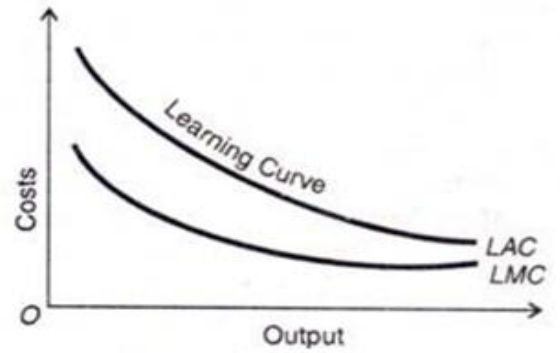


Fig.2.32

If the LAC curve is downward sloping up to the point of a minimum optimal scale of plant or a minimum efficient scale (MES) of plant beyond which no further scale economies exist, the LAC curve becomes horizontal. In this case, the LMC curve lies below the LAC curve until the MES point M is reached, and beyond this point the LMC curve coincides with the LA C curve, as shown in Figure 2.33.

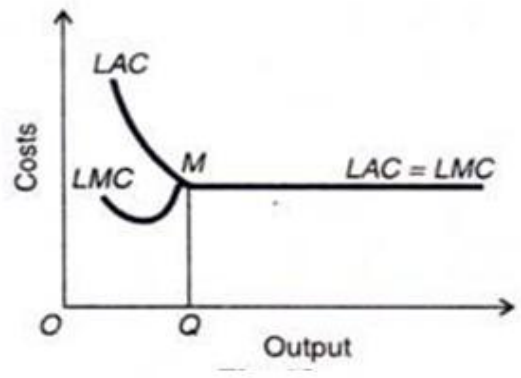


Fig.2.33

Conclusion:

The majority of empirical cost studies suggest that the U-shaped cost curves postulated by the traditional theory are not observed in the real world. Two major results emerge predominantly from most studies. First, the SAVC and SMC curves are constant over a wide-range of output.

Second, the LAC curve falls sharply over low levels of output, and subsequently remains practically constant as the scale of output increases. This means that the LAC curve is L-shaped rather than U-shaped. Only in very few cases diseconomies of scale were observed, and these at very high levels of output.

UNIT — III - PRICE AND OUTPUT DETERMINATION

Equilibrium of the firm and industry under perfect competition - Price and output determination under monopoly - Price and output determination under monopolistic competition - meaning of concepts Monopsony, Bilateral monopoly, oligopoly and duopoly

Equilibrium of the firm and industry under perfect competition:

Meaning:

A perfectly competitive market is one in which the number of buyers and sellers is very large, all engaged in buying and selling a homogeneous product without any artificial restrictions and possessing perfect knowledge of the market at a time.

Definitions

1. According to Mrs. Joan Robinson "Perfect Competition prevails, when the demand for the output of each producer is perfectly elastic".
2. According to Leftwhich, "Perfect Competition is a market situation in which there are many firms selling identical products with no firm large enough relative to the entire market to be able to influence market price".

Features

A market is said to be operating under Perfect Competition, if it satisfies the following conditions:

1. The first important feature of perfect competition is the existence of a large number of buyers and sellers in the market.
2. The second condition of perfectly competitive market is that all sellers are selling homogeneous or identical products. In other words, there is no product differentiation.
3. The next feature of perfect competition is that the firms are free to enter or leave the industry
4. Another condition of perfect competition is the existence of perfect knowledge on the part of the buyers and sellers regarding the market conditions.
5. The existence of perfect competition depends upon the perfect mobility of factors of production.
6. In a perfectly competitive market, there are no transport costs.
7. The last feature of perfect competition is that there is only one price for the commodity.

Distinction between pure and perfect competitions

Pure Competition

The Pure Competition is one which satisfies only the first three features of perfect competition, namely, existence of large number of buyers and sellers, existence of a homogeneous product and free entry and exit for firms

Perfect Competition

In addition to the features of Pure Competition, the Perfect Competition satisfies the other conditions namely,

1. Existence of perfect knowledge on the part of the buyers and sellers regarding the market conditions
2. Perfect mobility of factors of production
3. There are no transport costs.
4. There is only one price for the product

Valuation under perfect completion or price determination under perfect competition

The equilibrium price and output are determined by the interaction of total demand and total supply. Equilibrium price is that price at which quantity demanded of the product is equal to the quantity supplied. The demand curve of the market will slope downwards from left to right. This is, because, producers increase the supply at a higher price and vice versa

Table 3.1

Total Demand and Total Supply

Price of rice (Rs. Per Kg)	Total Demand (in Kg.)	Total Supply (in Kg.)
6	2	20
5	6	16
4	8	12
3	10	10
2	14	8
1	20	4

In Table 3.1 when the price of rice is Rs.6 per Kg., the supply is 20 kg., but the demand for rice is only 2 kg. Hence, 18 kg. of rice remain unsold. This will push down the price. When price falls from Rs.6 to Rs.5, the demand rises from 2 kg to 6 kg, while the supply falls from 20 kg to 16 kg, still the supply is in excess of the demand. Thus, the excess supply causes a

further downward trend on price. This process continues still the price is reached at Rs.3 per kg. at which both demand and supply are equal (10kg) this is the equilibrium price.

Let us begin from a low price. When the price is Re.1 per kg., the demand (20 kg) exceeds the supply (4 kg). Thus, there is a shortage of supply (excess demand) of 16 kg. in relation to higher demand. This will rise the price from Re.1 to Rs.2 per kg. This process continues till the price is reached at Rs.3 per kg. at which both demand and supply are equal (10 kg.). This is the equilibrium price.

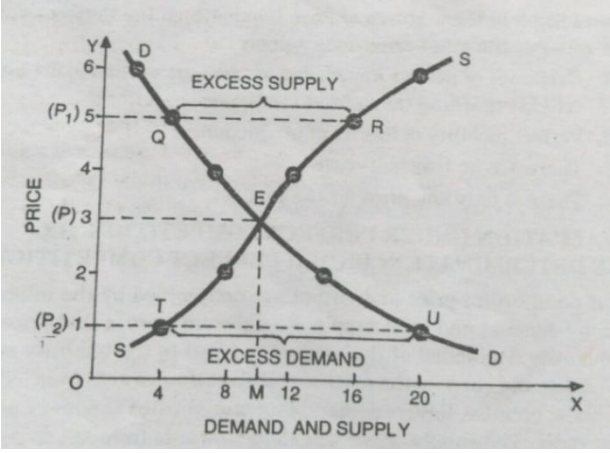


Fig.3.1

Explanation

In figure 3.1, X axis represents demand and supply and Y axis represents price. DD is the demand curve and SS is the supply curve. Both these curves intersect each other at point 'E', the equilibrium point. At this point, equilibrium price is OP (Rs.3) and the equilibrium output is OM (10 kg), where quantity demanded is equal to quantity supplied. When price increases from OP (the equilibrium price) to OP_1, the supply is P_1R while the demand is only P_1Q. Thus, excess supply is QR. That is, more is offered for sale than what the people demanded. Hence, in order to dispose the excess supply, the competing sellers will be induced to bring down the price from OP_1 to OP, the equilibrium price, where the demand equal to supply.

Normal profits

For equilibrium in a perfectly competitive market, two conditions must be fulfilled. Firstly, each individual firm must be in equilibrium. This will happen when every firm in the industry is earning maximum profit by equating marginal revenue with marginal cost. Secondly, the industry as a whole must be in equilibrium. This will happen when there is no

tendency for firms either to enter or leave the industry. This requires that all the entrepreneurs in the industry must earn enough money to make the move worth-while. In technical economic language, we describe the situation by saving that every entrepreneur in the industry is earning "Normal Profits". Therefore, "Normal Profits" for an entrepreneur in an industry are those profits which are just sufficient to induce him to stay in the industry.

The fact that we have introduced Normal Profit into our analysis means that we must make a slight alteration in the average cost curve. In order to decide whether a firm will be willing to stay in the industry in the Long-Run or not, it will be more useful to include normal profits in average cost curve. In future, the average cost curve which we shall draw will include normal profits. We shall assume that each entrepreneur must earn normal profit, if he is to stay in the industry in the Long-Run.

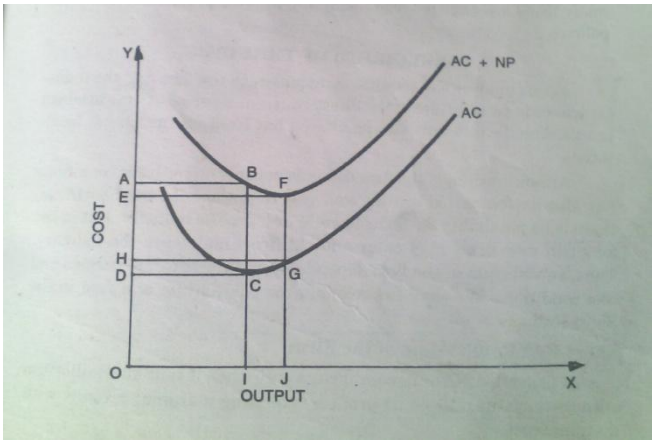


Fig.3.2

In Figure 3.2 .X axis represents output and Y axis represents cost. AC is the Average Cost curve excluding Normal Profits and AC + NP is the Average Cost curve including Normal Profits. It will be seemed that as output rises from OI to OJ, the vertical distance between the two curves diminishes. It means Normal Profit per unit of output declines. When output is OI, Normal Profit per unit of output is BC. When output is OJ, it is FG. However, the areas of rectangles ABCD and EFGH showing total normal profit will be equal.

Price and output determination under monopoly

We have seen that under Perfect Competition both Average Revenue curve (Demand Curve) and Marginal Revenue curve are the same and also perfectly elastic and is represented by a horizontal straight line parallel to X axis. This is so, because, under Perfect Competition, a seller can sell as much as he likes at the prevailing price. But under Monopoly, this is not so. Under Monopoly, the Average Revenue curve will be a down-ward sloping curve.

Further, the Marginal Revenue curve will also a down-ward sloping curve but it will be steeper (less) than the 'AR' curve. In other words, Marginal revenue curve lies below the Average Revenue curve. This is so, because a firm under Monopoly can sell larger quantities only when it reduces the price.

The principle of profit maximisation under Perfect Competition is applicable for Monopoly also. Therefore, the Monopolist will also maximise his profit by equating the Marginal Cost with the Marginal Revenue. The Monopolist will be in equilibrium when he gets maximum profit.

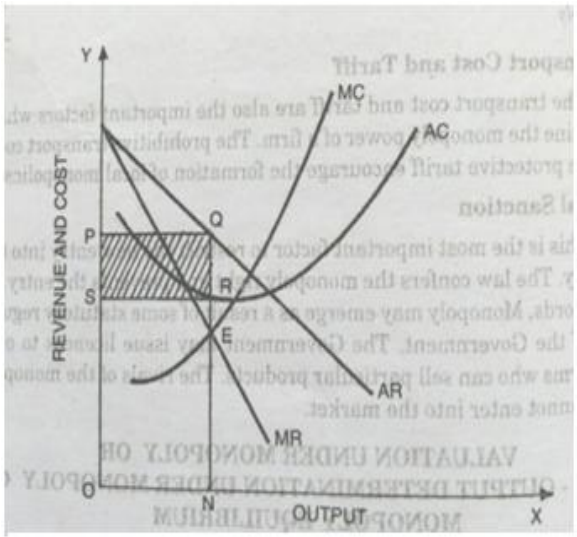


Fig.3.3

In Figure 3.3 X axis represents output and Y axis represents revenue and cost. AR is Average Revenue curve, MR is Marginal Revenue cure, AC is Average Cost curve and MC is Marginal Cost curve. The Monopolist, comes to equilibrium at point 'E' where MR=MC and produced 'ON' units of the commodity. At this level of output, the Average Revenue (price) is 'OP'(NQ) and Average Cost is 'NR'. Therefore, profit per unit = NQ – NR i.e., RQ. The total output is ON. Thus, the total profit is profit per unit of output multiplied by total output i.e., RQ x SR (SR=ON). The total profit, therefore, is "PQRS" (the shaded area).

Evils of Monopoly

1. Since Monopolist is a price maker, prices will be high under Monopoly. Hence, consumers are exploited by the Monopolist.
2. Further, the factors of production are also paid less under Monopoly than under Perfect Competition.
3. Monopoly leads to inequality of income and wealth.

4. Under Perfect Competition, each firm will adopt new methods of production. But, under Monopoly this is not so.
5. Under Monopoly, consumers will loss their freedom to choice to some extent
6. Monopoly tends to aggravate the unemployment problem due to underutilisation of resources

Control of Monopoly

1. Monopoly may be controlled by legislative measures. Anti-monopoly legislations should be enacted to check the growth of Monopoly
2. The Government may promote measures of competition by giving licences, permits etc., to new firms against monopoly firms.
3. Establishment of consumer's movement will effectively restrain Monopoly.
4. According to Prof. Pigou and Prof. E.A.G. Robinson, publicity against monopoly prices, profits and exploitations will also control Monopoly
5. The Government can also regulate a Monopoly through taxation.
6. Establishment of co-operative organisations will control Monopoly in course of time
7. Nationalisation of harmful monopolies will put an end to the Monopoly.

Price and output determination under monopoly

Meaning

It is a market situation in which there are many sellers but there is single buyer of a commodity. It is also known as "Buyer's Monopoly"

Example

For example, sugarcane cultivators are many, but a single sugar mill is the only one buyer in a particular region.

Price Determination Under Monopsony

The conditions of equilibrium of a profit maximising Monopsonist are the same as that of Monopoly. The Monopsonist will maximise his profit when marginal costs equals marginal revenue. It is illustrated in Figure 3.4

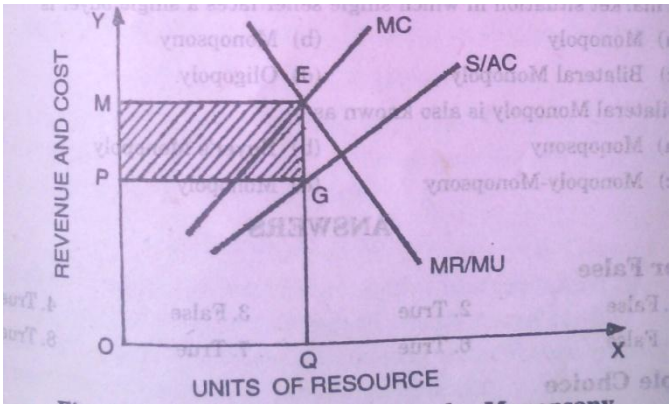


Fig.3.4

In Figure 3.4 MC is the marginal cost curve, 'S' is the market supply curve or average cost curve and MRP is the marginal revenue curve or marginal utility curve. The point 'E' is the equilibrium point where MR=AC. At this point, Monopsonist buys OQ quantity at OP (QG) price which is the supply price for that output. But revenue per unit of output is QE. Therefore, surplus per unit is GE. Hence, the Monopsonist makes a surplus equal to PGEM.

Bilateral Monopoly

Meaning

It is a market situation in which single seller (Monopolist) faces a single buyer (Monopsonist). In other words, the existence of one seller and one buyer in a market is known as 'Bilateral Monopoly'. Thus, under Bilateral Monopoly, the seller is Monopolist and the buyer is Monopolist. That is, under Bilateral Monopoly, Monopolist faces the Monopsonist. It is also known as the "Monopoly-Monopsony" situation. This Bilateral Monopoly usually exists in the public sector, e.g., Railways, Port Authority etc.

Valuation Under Bilateral Monopoly

The equilibrium price and output cannot be determined by demand and supply. It can be determined by the non-economic factors like bargaining power of the buyer and seller. It is illustrated in Figure 3.5.

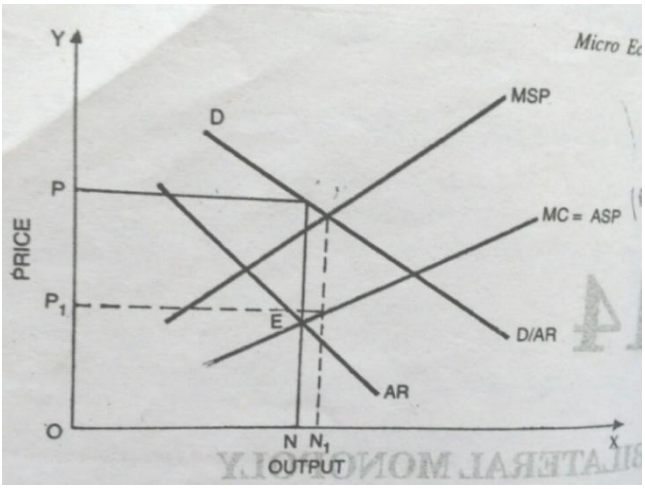


Fig.3.5

In Figure 3.5 DD is the demand curve for the buyer and also the marginal utility curve for the buyer. Further it is the average revenue or demand curve for the seller. MR is the marginal revenue curve of the seller. MC is the marginal cost curve of the seller which is the average supply curve (ASP). The seller is in equilibrium at point 'E' the equilibrium point where MR=MC. At this point, the output is ON and price is OP. On the other hand, the buyer is in equilibrium by equating his marginal supply curve (MSP) with marginal utility curve or the price (DD) he is prepared to pay. He will be maximising satisfaction by purchasing ON_1 amount at OP_1price. The price OP_1 is much lower than the price OP. The seller (Monopolist) wants to sell ON quantities at OP_1 price. Hence, the price and output cannot be determined by the demand and supply. It can be determined by the non-economic factors like bargaining power of the buyer and seller.

Oligopoly

Meaning

"Oligo" means "a few" and "Poly" means "selling". Thus, Oligopoly refers to a market situation in which there are "more than two" or "a few" sellers selling either homogeneous products or products which are having close substitutes but not perfect substitutes.

Definitions

1. Prof. Stigler defines Oligopoly as that, "situation in which a firm bases its market policy in part on the expected behaviour of a few close rivals".
2. According to Prof. Leftwhich, "An Oligopolistic industry is one in which the number of sellers is small enough for the activities of a single seller to affect other firms and for activities of other firms to affect him".

Characteristic features of Oligopoly

1. The first feature of the Oligopoly is that there are more than two or a few sellers in the market.
2. The next feature is that the sellers sell the homogeneous products or products which are having close substitutes.
3. The another important feature of an Oligopoly is the interdependence between the firms. That means, the price and output decisions of one firm will affect the others.
4. The most striking feature of Oligopoly is indeterminate demand. That means, no firm in Oligopoly can forecast the nature and position of its demand. Because, it depends upon the price and output decisions of other firms.
5. Another character of an Oligopoly is the enormous selling cost (advertisement cost) incurred by the competing firms.
6. Another peculiarity of Oligopoly is the conflicting attitude of the firms.
7. Another important character of an Oligopoly is price rigidity. That means, the prices in Oligopoly are constant.
8. Another condition of an Oligopoly is that, sometimes a firm takes up leadership (price leader) in fixing the price of the product. The other firms in the industry simply follow the price leader and accept the price fixed by it.

Kinds of Oligopoly

1. Pure or Perfect Oligopoly and Differentiated or Imperfect Oligopoly

An Oligopoly is said to be Pure or Perfect when it sells homogeneous products. On the other hand, Oligopoly is Imperfect or Differentiated, when it sells different products.

2. Open Oligopoly and Closed Oligopoly

An Open Oligopoly refers to that market situation in which the new firms are free to enter into the market and compete with the existing firms. But a Closed Oligopoly refers to the market situation in which the new forms cannot enter into the market.

3. Collusive Oligopoly and Non-Collusive Oligopoly or Competitive Oligopoly

An Oligopoly in which firms have agreement or collusion regarding price and output is called as Collusive Oligopoly, whereas Oligopoly in which firms have no such agreement or collusion is called Non-Collusive or Competitive Oligopoly.

4. Partial Oligopoly and Full Oligopoly

A Partial Oligopoly is one which is dominated by the price leader. On the contrary, Full Oligopoly is one which has no such price leader.

5. Syndicated Oligopoly and Organised Oligopoly

Syndicated Oligopoly refers to the market situation in which the firms sell their products through a centralised syndicate, while an Organised Oligopoly is one in which the forms organised themselves in a central organisation for fixing prices, output and quotas.

Kinked Demand Curve or Sweezy's Model or Pricing under Oligopoly

Prof. Paul M. Sweezy has used the "Kinked Demand Curve" to explain price rigidity under Oligopoly. Hence, his Model is called as "Kinked Demand Model". This Model represents a condition in which the firm has no incentive either to increase or to decrease the price, but keeps the price rigid. This is because, the firm fears that if the firm increases the price, it will loss its customers and if the price is decreased, the rival firms will decrease the price and hence profits of all firms would be reduced. Hence, the firms sticks on to the present price, until a drastic change in its demand and cost conditions.

Under these conditions, the demand curve of the firm under Oligopoly will be kinked. This means, that the demand curve will have a "kink" or "bend" at the present price as shown in Figure 3.6

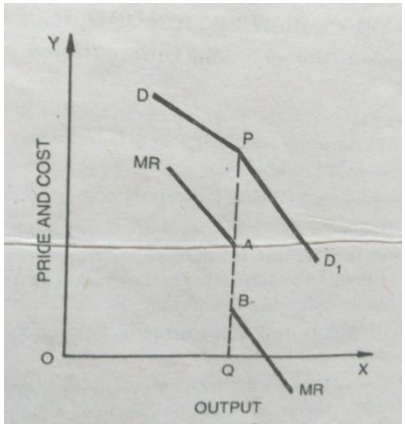


Fig.3.6

In Figure 3.6 DD_1 is the demand curve and MRMR is the marginal revenue curve. The demand curve has a kink at the point 'P'. QP is the price at which the firm is selling OQ units of the product. The anticipated demand of the form is DP. That is, above the point, the demand curve DP is elastic. This shows that if the firm increases the price above QP while all the firms maintain QP price, it will loss its customers. Hence, the demand for the product of the firm would fall and consequently the total revenue and profit of the firm would be reduced. So, the firm keeps the price rigid at QP. The corresponding marginal revenue curve declines as shown in Figure 3.6.

Below the price QP, the anticipated demand of the firm will be PD_1 which is inelastic. It is inelastic, because if the firm reduces the price, that price will be followed by other firms and hence the profits of all firms would be reduced. Hence, the firm sticks on the present price QP. At this level, the MR curve is negative as shown in Figure…Thus, when the demand curve is DP, MR is positive and if the demand curve is PD_1, MR is negative. So, the firm keeps the price rigid at QP. The firm has no scope of better profit in either way.

The peculiar feature in Figure 3.6 is that, there is a gap or discontinuity (AB) in the MR curve below the kink point i.e., between MR and MR_1. The gap depends upon the elasticity of demand above and below the kink. The gap will be greater, if the elasticity is greater above the kink and inelasticity is greater below the kink. The price will not change, unless there is drastic change in demand and cost conditions of the firm concerned.

The Kinked Demand Model explains the price rigidity, but does not explain how the price is determined under Oligopoly. Further, this Model cannot be applied to Oligopoly with product differentiation and price leadership or Collusive Oligopoly.

Price Leadership

Meaning

Under Oligopoly, if the industry is dominated by one large firm, then that firm is called as the "Leader" of the industry. The Leader will be the "Price Leader:. This situation is called "Price Leadership". Under price Leadership, the price leader fixes the price of the product for the entire industry. The other firms in the industry accept the price fixed by the price leader and adjust their output to this price. The price leader is having the lowest cost of production.

Price-Output Determination under Price Leadership

Economists have developed various models to determine the price output under price leadership. Let us now explain a simple model to determine price-output under price leadership on the basis of the following assumptions:

1. There are only two firms say, A and B and firm A is the price leadership having lower cost of production than the firm B.
2. The firms produce homogenous product, so that consumers are indifferent between the firms
3. Both the firms have equal share in the market. That is, they are facing same demand curve which will be half of the total demand.

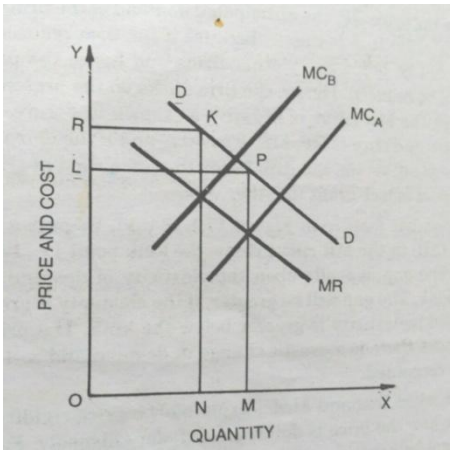


Fig.3.7

In Figure 3.7, X axis represents quantity and Y axis represents price and cost. DD is the combined demand curve for the goods of firms A and B. The two firms share the market demand 50 per cent each. MR is the marginal revenue curve of each firm, MC_A is the marginal cost curve of the firm A and MC_B is the marginal cost curve of the firm B. Since we have assumed that the firm A is the price leader and is having lower cost of production than the firm B, MC_A is drawn below than MC_B

Firm A will be maximising its profit by selling OM amount of output at OL (MP) price, because at the output ON, its marginal cost (MC_A) is equal to its marginal revenue.

Firm B will be maximising its profit by selling ON amount of output at OR (NK) price, because at the output ON, its marginal cost (MC_B) is equal to its marginal revenue.

It can be seen that the profit maximising price OR of the firm B is higher than the profit maximising price OL of the firm A. Since it is assumed that the products of the two firms are homogenous, the two firms will have the same price. Unless the firm B reduce the price from OR to OL, it will loss its customers and automatically it will leave the industry. Therefore, the firm B accepts the price fixed by the firm A. This shows that, in this situation, the firm A is the price leader and the firm B has to follow it. Thus both the firms A and B will charge the same price OL and sell the same output OM.

UNIT- IV : ALTERNATIVE THEORIES OF FIRM AND DISTRIBUTION

*Bain's limit Pricing theory of recent development - Prof. Baurnol's Theory of Sales - Revenue maximization - Neo - classical Theory of Marginal Productivity theory Euler's Theorem - Factors Pricing under different market conditknis - **Modem** Theories of Rent, Wages, Interest and Profit (including Shakle's Theory of profit) .*
Micro Theories of distribution.
a) Richardian b) Marxian c) Kalecki and d) Kaldor

Bain's Limit Pricing Theory of Recent Development

Introduction:

Bain formulated his 'limit-price' theory in an article 1949 and he published his major work Barriers to competition in 1956. In his article, he aimed at explaining why firms of long period of time were keeping their price at a level of demand where the elasticity was below unity, that is, they did not change the price which would maximize their revenue. He arrived at the conclusion that the traditional theory was unable to explain this empirical fact due to the commission from the pricing decision of important factor, namely, the threat of potential entry. Traditional theory was concerned only with actual entry, which resulted in the long-run equilibrium of the firm and the industry (where P=LAC). However, the price, Bain argued did not fall to the level of LAC in the long-run because of the existence of barriers to entry while at the same time price was not set at the level compatible with profit maximisation because of the threat of potential entry. In fact, he maintained that price was set at the level above the LAC (= pure competition price) and below the monopoly price (where MC=MR short-run profits are maximised). We can explain this behaviour by assuming that there are barriers to entry, and that the existing firms did not set the monopoly price but the 'limit price' that is, the highest price which the established firms believe they can charge without inducing entry.

Bain's theory is based on the following assumption

1. There is a determinate long-run demand curve for industry output, which is unaffected by price adjustment of sellers or by entry. Hence, the market marginal revenue curve is determinate. The long-run industry-demand curve shows the expected sales at different prices maintained over long periods.
2. There is effective collusion among the established oligopolists.
3. The established firms compute a limit price, below which entry will not occur. The level at which the limit price will be set depends
    a. On estimation of cost of the potential entrant

b. On the market elasticity of demand
c. On the shape and level of the LAC
d. On the size of the market, and
e. On the number of firms in the industry

4. Above the limit price, entry is attracted and there is considerable uncertainty concerning the sales of the established firms (post entry) and
5. The established firms seek the maximisation of their own long run profit

Bain has developed two models: Model (1) Where there is no collusion with the new entrant and Model (2) where there is collusion with the new entrant.

In his book, "Barriers to New Competition" (1956) Bain has changed the basis of his study. In his 1949 articles, he tried to explain why the actual price set by the oligopolist was below the monopoly price. His answer was that the retention of price below the monopoly was due to threat of potential entry, the firms charged the limit price, lower than the monopoly price, because the entry prevention policy secured maximum long run profits.

Bain has tried to explain why price is set above the competitive price, that price which is equal to the long run AC. His conclusion is that the limit price is above the competitive price due to barriers to entry. He went on to develop a theory of pricing to prevent entry without explicitly discussing its rationality, that is without saying whether this policy aims at profit maximization, or whether entry prevention is a goal in itself for achieving long run survival, or keeping market share constant.

Bain's ideals on entry prevention may be studied in terms of (1) Bain's concept of competition and entry (2) Barriers to entry and (3) Bain's theory of pricing: equilibrium of the industry.

The concept of Competition

Bain makes distinction between two types of competition: actual and competition between established firms and potential competition arising from potential entry of firms outside the Actual competition is probably of high importance as a regulator of business activity especially in oligopolistic markets. However, the threat of potential entry is also an important determinant of the pricing policies of firms.

In the traditional theory of pure competition and monopolistic competition, entry is discussed explicit (we may note that in the traditional theory of monopoly entry is explicitly assumed to be barred in both the short and long run). In both models, entry is treated as actual

entry, whose effects on the long run. The latter implies that entry comes from completely new firms.

The former implies no consideration of the effects on current price output decisions of the threat arising from potential entrants.

Both models assume that the entry is free. That is, there are no legal restrictions. Not only that, the models assumes that the entry is easy in the long run. In other words, there are no insurmountable barriers to entry. The consequence of the assumption of free entry in pure competition is that in the long run equilibrium of each firm and on the industry is attained at

P = LAC = LMC

and costs are at a minimum.

The consequence of the assumption of free entry in a monopolistic market in the long run is that equilibrium is reached with the tangency solution.

P = LAC and P > LMC

But costs are not at a minimum. Because, excess capacity is a necessary consequence of the falling demand curve.

In oligopolistic markets, however, conventional theory has in general been silent. The oligopoly models of Cournot, Bertand, Edgeworth and Chamberlin are closed models. They do not allow for entry. The number of firms in the model is assumed constant. The models of collusive solutions are also closed models. Cartels are inherently unstable and they break down once entry is allowed while in the traditional models of price leadership the entrant is usually assumed to be a small firm which follows the leader.

Fellner was the first economist to mention that the threat of potential entry is one of the reasons why the collusive oligopolists do not charge the monopoly price. Hatt and Hitch, PW.S Andrews and R.S. Edwards have underlined the importance of potential entry as a determinant of pricing decisions. We shall, however, examine Bain's contribution to the analysis of entry barriers.

The Concept of entry in Bain's Theory

Bain considers entry as the establishment of a new firm which builds or introduces new productive capacity that was not used for production in the establishment of the new firm.

Bain excludes from his entry concept (a)the take over of an existing firm by some other firm which constitutes change in ownership, (b) the expansion of capacity by an

existing firm (c)cross entry, that is, entry by a firm already established in another industry, which adds the product of this industry to its line of products.

Barriers to Entry

Bain introduce the concept of 'the condition of entry' which he defined as a margin by which established firms can raise their above the competitive price level persistently without attracting entry. Symbolically we have

$$E \quad \frac{PL - Pc}{Pc}$$

Where E = condition of entry

Pl = Limit Price

Pc = Competitive price, that is, price under pure competition in the long run (pc = LAC) Solving for PL we find

PL = Pc (1+E)

Thus the condition of entry, E, in fact the premium accruing to the established firms in industry from charging a price. PL, higher than the pure competitive price Pc, without acting entry.

According to Bain the time period which is implied in the above definition of the condition of entry is long enough to encompass a typical range of varying conditions of and, demand factor prices and the like. This period normally, might of as 5 to 10 year (Barriers to New Competition, P.7)

In Bain's theory entry is ling run phenomenon. The lag of entry is an important determinant of the barriers to entry. For example the establishment of a firm in women's garments may require four months and in cement one to two years, the longer the lag, the less the threat of entry and hence the greater the gap between the limit price PL and the competitive price Pc.

The lag of entry, that is, the tie required for a new firm to be established depends on many factors which constitute the barriers to entry. Bain distinguishes four main barriers to entry: (1) product differentiation barrier, (2) absolute cost advantage of established firms, (3) economics scale and (4) large initial capital requirements.

We may add to this list the legal barriers to entry which are imposed by law. The firms protected by such legal barriers may charge any price without the fear of attracting new entrants.

On the basis of empirical findings, Bain noted that there were very strong (important) barriers to entry in industries such as automobiles, cigarettes, liquor, tractors and typewriters. However, in interpreting Bain's finding, we should keep in mind that Bain considered the barriers faced by completely new firms. This is a serious limitation of his generalization. If e take into account that the most typical form of entry is that of an already established firm which diversifies its production, or integrates vertically, or takes over other firms, or merges with other existing firms, then Bain's results may not be valid in the modern business world. Cross entry and within entry practically eliminated the importance of barriers to entry.

Industry Equilibrium

In oligopolistic markets, each of the large firms, whether they act in collusion or singly, will study the barriers to entry and will set a price low enough to prevent entry.

Costs are the same for all established firms, they can arrive at the limit price sooner or later by independent action, via trial and error. Collusion is not necessary if costs and efficiency are identical although with collusive action the limit price will be arrived at quicker and with less danger of 'spoiling the market' by provoking a price war. However, if efficiencies differ among established firms, the limit price will be set by the most efficient, least cost firm.

The threat of potential entry, defined by the strength of barriers to entry, is crucial in pricing decisions. Potential competition is dealt with by charging the limit price. Bain introduced the concept of 'condition of entry' which may be defined as a percentage by which established firms can raise their price above the competitive level (Pc) without attracting entry.

$$E \quad \frac{PL - Pc}{Pc}$$

Where E = Condition of entry

Pc = competitive Price (=PAC)

PL = limiting price

Rearranging the above formula we obtain,

$$PL = Pc\ (1+E)$$

Rearranging the above formula we obtain,

$$PL = Pc\ (1+E)$$

In this form, we can see that the price limiting entry is determined by the competitive rice (=The LAC of the most efficient firm) and the premium E which is a measure of the barriers to entry.

If entry is easy, barriers are non-existent (or negligible) thus E=O, and hence,

PL =Pc +=LAC

That is, the firms will be able to charge a price PL which will be higher than Pc and they will be earning abnormal profits (since PL >LAC)

The competitive level of price is equal to the LAC of the most-efficient firm. The long run cost curve is L shaped long run cost curves are approximately constant, as the firm size exceeds the minimal optimal scale. The LAC (=Pc) includes a normal profit, that is a normal return to capital. Thus, E is a premium, a margin above the long run competitive price.

In the real world, there are differences in the products produced by firms and in the efficiency of the firms. Firms have different demand and cost curves. The term 'industry' includes products for which a stable relative price quality pattern emerges, such a group of products can be 'treated' as one for demand purposes. In practice, it has been noticed that there are products which can be classified within a 'broad commodity group' which may for all practical purposes be characterized as a distinct 'industry'. Of course, the price quality pattern may not remain strictly stable, but the degree of stability is adequate to justify the retention of the industry concept and to produce an analysis on a partial equilibrium approach.

Thus in equilibrium, the price is not unique. There is a cluster of prices reflecting the differences in product quality and costs. Under these conditions, the 'condition of entry' (E) may be measured specifically as the maximum gap between price and the minimal cost of the most efficient firm(s) at which entry is prevented. In Bain's words, "the condition of entry (B) is measured by the long run gap between minimal cost and price which the most favoured firms can reach without attracting entry". The less efficient firms will be making concurrent price elevations.

There may be many potential entrants. But the one that is crucial for the pricing decision of the established firms is the potential firm with the lowest costs as compared with the other potential entrants. Thus, the relevant entrant is the 'most favoured' entrant in the sense that he has the lowest costs and sufficient to grant him a profitable existence.

Baumol's Theory of Sales – Revenue Maximization

Introduction

The sales Maximization Hypothesis has been put forward by Prof. Baumol. In his view, Maximization of sales rather than the maximization of profits is the ultimate objective of the firm. He says that the firm tries to promote sales and not merely as a means to further its other objectives such as operational efficiency and profits. Baumol thinks that sales maximization is the most valid assumption governing the behaviour of a firm by sales the means the revenue earned by selling the product. Therefore, it is also otherwise called "Revenue Maximization Hypothesis" Prof. Baumol very strongly believes that sales maximization has become the sales maximization of the entrepreneur (or) the manager of a firm and therefore they direct their energies in promoting and maximizing sales misted of profits.

Rationale for Revenue Maximization Hypothesis

It is an alternative to the principle of profit maximization. Baumol has given many reasons to justify revenue maximization hypothesis. In the modern firm, there is the separation of ownership from management. It gives discretion to managers to pursue goals which maximise their own utility and deviate form profit maximisation, which is desirable goal of owner. Given this discretion, Baumol suggest sales maximisation seems the most plausible goal of managers. From his personal experience as a consultant to large firms, Baumol found that manager's are preoccupied with maximization of sales rather than profits. He has given the following reasons for this attitude of the management.

Firstly, there is evidence that salaries and their (slack) earnings of top management are correlated more closely with sales than with profits.

Secondly, the banks and other financial institutions keep a close eye on the sales of firms and are more willing to finance firms with large and growing sales.

Thirdly, personnel problems are handled more satisfactorily when sales are growing.

Fourthly, large sales growing over time prestige to the managers, while large profits, go into the pockets of shareholders.

Fifthly, managers prefer a steady performance with satisfactory profits to spectacular profit maximisation projects.

Sixthly, large growing sales strengthen the power to adopt competitive skills.

The executives always give top priority for sales maximisation. In the words of Baumol, almost every time I have come across a case of conflict between profits and sales,

the businessmen with whom I worked left little doubt as to where their hearts lay a programme which explicitly proposes any cut in sales volume, whatever the profit consideration, I likely to meet a cold reception.

According to Baumol, the top managers become to a certain risk avoiders. Although Baumol recognises the interdependence of firms as the main features of oligopolistic markets, he argues that in day-to-day decision-making, management often explicitly or implicitly on the premise that its decision will produce no changes in the behaviour of those with whom they are competing it is only when the firm takes more radical decisions, such as the launching of major advertising campaign or the introduction of a radically new line of products that management usually does consider the probable competitive response. But often, even in a crucial decisions, and almost always in routine policy making, only the most cursory mention is paid to competitive reactions.

The desire of top management for a 'quiet life' has led large enterprises to some tacit collusion; firms depend on each other to behave in an 'orderly' way. However, we may note that businessmen are not completely indifferent to action of petitions. Being sales maximisers, they are very alert to any changes in the share of their market. Top management will ignore competition only to the firm's market and do not with the desired rate of growth of the sales of the firm.

Other evidence cited by Baumol essentially suggests that short run revenue 'maximisation may be constant with ling run profit maximisation. However, Baumol states that revenue maximisation in many oligopoly firms.

The minimum acceptable level of profits which serves as a constraint on a firm's attempts to maximise in revenue is not regarded as purely arbitrary. Baumol feels that this constraint is determined by the capital market. In his own words, "typical oligopolic firm, while large in the market for its own product, is relatively small in the capital market. This means that, in obtaining capital by the issue of stock, it must be prepared to meet competitive pricing conditions – they yield on its stock price rises can remunerate stockholders adequately. If this is so, each company's minimum rate of profits is set competitively in terms of the current market value of its securities.

Baumol presented two basic models. The first is a static single period model, the second is a multi-period dynamic model of growth of sales revenue maximisation. Each model has two versions, one without and one with advertising activities.

The Basic Assumption of Baumol's static models

1. The time horizon of a form is a single period
2. During this period the firm attempts to maximise its sales revenue not physical volume of output subject to profit constraint.
3. The minimum profits constraints is exogenously determined by the demands and expectations of the shareholders, the banks and other financial institutions
4. 'Conventional' cost and revenue functions are assumed. That is, Baumol accepts cost curves are U shaped and the demand curve of the firm is downward sloping.

The new element in the single produce model, is the introduction of advertising as a major instrument (policy variable) of the firm – Baumol argues that in the real world non price competition is the typical form of competition on oligopolistic markets.

One of the important findings of Baumol's single product model is that sales, revenue increases with advertising expenditure (that is OR/DA>O where a = advertising expenditure). This implies that advertising will always short the demand curve of the firm will sell a larger quantity and earn a large revenue. The price is assumed to remain constant.

A firm, in an oligopolistic market will prefer to increase its sales by advertising rather than by a cut in price. And the sales maximiser will normally have a higher advertising expenditure than a profit maximiser.

Multi-product firm (without advertising)

If we assume that the firm has a given amount of resources and given costs and wants to allocate them among various commodities it produces so as to maximise sales revenue it will reach the same equilibrium solution as the profit maximisers. Formally the condition for the equilibrium of the multi product firm with given resources and costs is which reads: The firm is in equilibrium when the ratio of the marginal revenue from any two commodities (i and j) is equal to the ratio of their marginal costs.

$$\frac{dR/dXi}{dR/dXj} = \frac{\delta c/dXi}{\delta c/dXj}$$

Note: If the resources (and costs) are not given and the firm can allocate increased quantities of factors to be produced of any product, the profit maximiser will have a different equilibrium product mix than the sales maximiser.

The multi-product model, with advertising also aims at sales revenue maximisation subject to minimum profit constraint

Baumol's Dynamic model

The most serious weakness of the static model is the short time horizon of the firm and the treatment of the profit constraint as an endogenously determined magnitude, in the dynamic model, the time horizon is extended and the profit constraint is endogenously determined.

The assumptions of the dynamic model: The firm attempts to maximise the rate of growth of sales over lifetime. Profit is the main means of financing growth of sales and as such is an instrumental variable whose value if endogenously determined. Demand and costs have the traditional shape; demand is downward falling and costs are U shaped. In the dynamic model, profit is not a constraint (as in the static model) but an instrumental variable. It is a means whereby the top management will achieve its goal of a maximum rate of growth of sales.

If we assume that the sales revenue (R) grows at a rate of growth (g)per cent, then the firm attempts to maximise the present value of the stream of sales revenue (S)over its life time by choosing appropriate values for the current (initial) level of sales revenue(R) and its growth rate (g). It is obvious that S is positively related to both R and g. the present value of the stream of revenue will be higher R and g values. Thus the firm should choose as alarge as possible values of R and g.

Some comments on Baumol's Model

The sales-maximisation hypothesis cannot be tested against competing behavioural hypothesis unless the demand and cost functions of individual firms are measured. However, such data are commonly unknown to the firms.

It has been argued that in the long run, the sales maximisation, the profit maximisation hypothesis yield identical solutions: because profits attain their normal level in the long run and the minimum profit constraint will coincide with the maximum attainable (normal) level of profit. But this argument cannot be accepted without any empirical evidence to support it. The sales maximisation theory does not show how equilibrium in an industry, in which all firms are sales maximisers, will be attained. The relationship between the firm and the industry is not established by Baumol.

Baumol's model is based on the implicit assumption that firm has market power, that is, it can have control on its price and expansion policies. The firm can take decisions without being affected by competitors. This Baumol rules out interdependence ex-hypothesis and hence his theory cannot explain the basic problem of uncertainty in non-collusive oligopoly

markets. The theory cannot explain observed market situation in which price is kept for considerable time periods in the range of inelastic demand. The theory ignores not only actual competition, but also the threat of potential competition.

MM Peston in his article, "On the Sales Maximisation Hypothesis" (Economics, 1959) put forth the argument that sales maximisation is not in compatible with the goal of long run profit maximisation. He has argued that a firm may be willing to keep sales at a higher level, even though they are unprofitable in the short run, with hope that eventually (in the long run) the product will become profitable once established in the market. Such behaviour is common for new products for which the firm expects no profits or even losses at the initial stage of their introduction. However, firms except to earn profits once their product becomes known in the market and captures a share at least equal to the minimum optimum scale. This behaviour does not by itself provide a proof that the firm is a sales maximiser or a profit maximiser. Moreover, in Baumol's model, sales and profit are not competing goal up to the level of output at which profit is maximised. Thus Peston's argument does not seem to invalidate Baumol's theory

Baumol's claim that because in his model output will be larger than the output of a profit maximiser, the sales maximisation hypothesis implies a lower degree of misallocation, increase in the welfare of the society. This claim is not necessarily true. (The whole argument rest on the shape of the demand and cost curve as well as on the way by which measures society's optimal output)

Euler's Product Exhaustion Theorem (ML Jhingan)

Introduction

If all factors were paid rewards equal to their marginal products, would the total product be just exactly exhausted? In other words, if each factor is rewarded equal to its marginal product, the total product should be disposed of without any surplus of deficit. The problem of providing that the total product will be just exhausted if all factors are paid rewards equal to their marginal products has been called "Adding-up Problem"

Euler's theorem can be solved as under. Let C and L be the quantities of two factors of production, capital and labour respectively and P the total product of these factors. Then P= *f* (C,L). In other words, if P is a linear homogenous function (*f*) of C and L, the following equation will hold:

$$P= \frac{\delta f}{\delta C} C+ \frac{\delta f}{\delta L} L \qquad \ldots\ldots(1)$$

If the quantities of all inputs C and L are increased k-fold, the output P will also increase k-fold. Then the production function becomes.

$$kP=f(kC, kL)$$

By taking the total derivate of kP with respect to k, we have

$$\frac{dk}{} \quad P = \frac{\delta f}{\delta kC} \cdot \frac{\delta kC}{\delta k} + \frac{\delta f}{\delta kL} \cdot \frac{dkL}{dk}$$

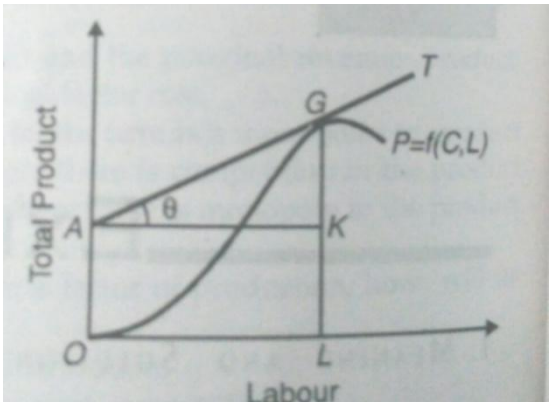


Fig.4.1

Where $\delta f / \delta C$ is the marginal product of capital and $\delta f / \delta L$ is the marginal product of labour. And $\delta f / \delta C$. C is the share of capital in the product P, and $\delta f / \delta L$. L is the share of labour in the total product. The above equation states that the marginal product of capital ($\delta f / \delta C$) multiplied by units of capital employed (C)plus the marginal product of labour ($\delta f / \delta L$) multiplied by the number of labourers (L)exactly equals the total product, P. Thus total factor payments exhaust the total value of the product.

Diagrammatic represents of Euler's Theorem. To restate, Euler's Theorem is

$$P= \frac{\delta f}{\delta C} C+ \frac{\delta f}{\delta L} L$$

It is illustrated in Fig. where labour is taken on the horizontal axis and the total product on the vertical axis. The curve OP is the total product curve or the production function: P = *f* (C, L). The tangent T on the OP curve at point G represents constant returns to scale. The slope at point G is equal to

$$\frac{\delta f}{\delta L} = \tan \theta = \frac{GK}{AK} \qquad (P= f(C,L)$$

$$= \frac{GK}{OL} \qquad (OL=AK)$$

Now $\frac{\delta f}{\delta L} L = \frac{GK}{OL}. OL = GK \qquad (L=OL)$

Which is the share of labour in total product GL.

From equation (1), we have

Now $\frac{\delta f}{\delta C} C = P \frac{\delta f}{\delta L} L$

$= P - GK$ (GK from equation 2)

$= GL - GK$ (total product P is equal to GL)

$= KL$

Which is the share of capital in total product GL

Thus $P = \frac{\delta f}{\delta C} C + \frac{\delta f}{\delta L} L$

$GL = GK + KL$

Hence the total product (GL) is fully exhausted (or distributed) between the two factors, capital (KL) and labour (GK)

Assumptions:

Euler's theorem (or the adding up problem) is based on the following assumptions

First, it assumes a linear homogeneous production function of the first degree which implies constant returns to scale.

Second, it assumes that the factors are complementary, i.e., if a variable factor increases, it increases the marginal productivity of the fixed factor.

Third, it assumes that factors of production are perfectly divisible

Fourth, the relative shares of the factors are constant and independent of the level of the product.

Fifth, there is a stationary, riskless economy where there are no profits

Sixth, there is perfect competition

Last, it is applicable only in the long run.

Explanation

Given these assumptions, Wicksteed proved with the help pf Euler's theorem that when each factor was paid according to its marginal product, the total product would be

exactly exhausted. This is based on the assumption of a linear homogenous function. Wicksteed did not differentiate competition and constant returns to scale; the product exhaustion theorem was universally valid.

Wicksteed's solution was treated by Edgeworth with mockery and Pareto objected to the assumption of constant returns to scale. Wicksell, Walras and Barone also criticised him. They pointed out that the production function does not yield a horizontal long-run average cost curve (LRAC) but a U-shaped LRAC curve. The U-shaped LRAC curve curve first shows decreasing returns to scale, then constant and in the end increasing returns to scale, "Where Wicksteed went wrong" writes Hicks, "was his assumption that he could argue from the shape of the curve at one particular point to the general shape of the curve".

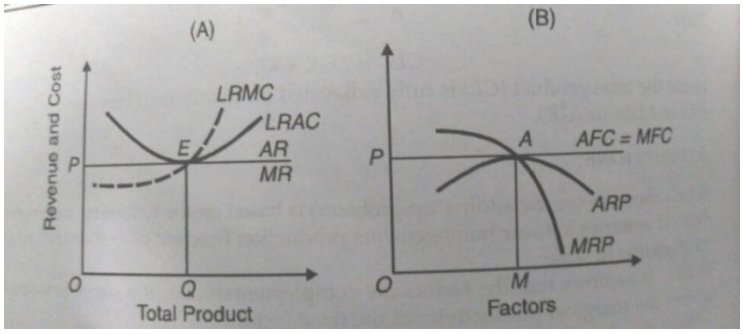


Fig.4.2

Wicksell proved that the product exhaustion problem held under perfectly competitive conditions in the long-run when profits were zero. He regarded it as a condition of equilibrium at the minimum point of firm's long-run average cost curve (LRAC) where the linear homogenous production function was satisfied. Suppose an entrepreneur is left with more than the marginal product of the resource he owns after paying all other resources their marginal products. Then all owners of resources are induced to become hiring agents and in the process th difference between the total product and the rewards to factors is eliminated. Conversely, if the residual left with the entrepreneur is less then his marginal product, after paying the other resources their marginal products, he will cease to be a producer and lend his services for its marginal product. Thus a firm under competitive conditions will produce at a level where the total product is exactly distributed according to the marginal product of the factor.

This solution of the product exhaustion theorem is based on a profitless long-run, perfectly competitive equilibrium position of a firm which operates at the minimum point, E of its LRAC curve, as shown in panel (A) of fig. At this point the firm is in full equilibrium, the marginal revenue productivity (MRP) of the factors being equal to the combined marginal

cost of the factors (MFC). This is shown in panel (B) of the fig. where MRP =MFC at point A. It is at point A that the total product OQ is exactly distributed to OM factors and nothing is left over.

As studied above, the product exhaustion problem is solved with a linear homogenous production function: $P = \frac{\delta f}{\delta C} C + \frac{\delta f}{\delta L} L$

If, however, there are diminishing returns to scale, less than the total product will be paid to the factors

$$P > \frac{\delta f}{\delta C} C + \frac{\delta f}{\delta L} L$$

In such a situation, there will be super-normal profits in the industry. They will attract new firms into the industry. As a result, output will increase, price will fall and profits will be eliminated in the long-run. In this way, the distributive shares of the factors as determined by their marginal productivities will completely exhaust the total product.

Criticism

In reality, constant returns to scale are incompatible with competitive equilibrium, for if long-run cost curve of the firm is horizontal and coincides with the price line the size of the firm is indeterminate; if it is below the price line the firm will become a monopoly concern, and if it is above the price line, the firm will cease to exist.

While in the case of increasing returns to scale, more than the total product will be distributed, because doubling the factors will more than double the total product. But increasing returns are incompatible with perfect competition, since the economies of production lead to the lowering of cost of production and in the long-run there is a tendency towards the establishment of a monopoly.

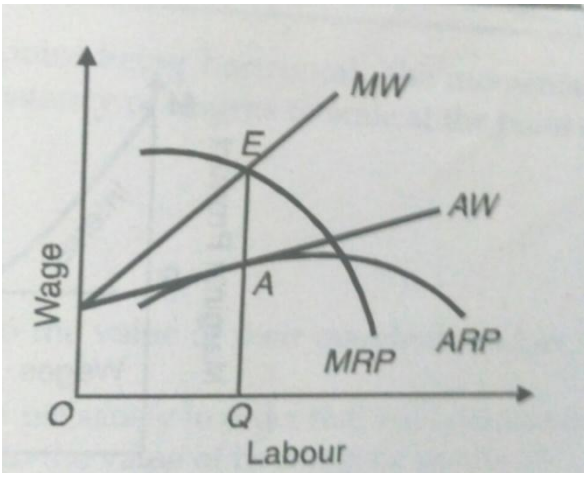


Fig.4.3

The whole analysis is based on the assumption that factors are fully divisible. Since the entrepreneur cannot be varied, we have not taken him as a separate factor. In fact, entrepreneurship disappears in the stationary economy. When there is full equilibrium at the minimum point of the LRAC curve, there is no uncertainty and profits disappear altogether. So the assumption of an entrepreneurless economy is justified for the solution of the adding-up problem. But once uncertainty appears, the entrepreneur becomes a residual claimant and the exhaustion of the production problem disappears.

Under imperfect or monopolistic competition the total product add up to more than the share paid to each factor, that is, P is greater than C and L. Taking an imperfect labour market, the average and marginal wage curve (AW and MW) slope upward and the average and marginal revenue product curves (ARP and MRP) are inverted U-shaped, as shown in Fig. Equilibrium is established at point E where the MRP curve cuts the MW curve form above. The firm employs OQ units of labour by paying QA wage which is less than the marginal revenue product of labour QE. Thus workers are paid less than their marginal productivity when there is imperfect competition. This argument applies not only to labour but to all shares even under constant returns to scale in the industry.

The product exhaustion theorem, however, holds true under monopolistic competition when the firm is in equilibrium. At equilibrium, the marginal cost curve cuts the marginal revenue curve and the average revenue curve is tangent to the average cost curve. It follows that the total outlay for factors and the total revenue product will be equal. If now a small change in factors is made, keeping their prices constant, the increase in the total revenue product is approximately proportional to the increase in the outlay for factors. Thus if each factor included in the cost curve is paid according to its marginal revenue product at equilibrium, the total product of the firm will be exactly exhausted among them. But if there is monopoly, payment in accordance with marginal product will not exhaust the total product.

Modern Theories of Rent

Meaning

In ordinary language rent is the price or reward or remuneration given for the use of land or a house or a machine to the owner. But, in Economics, Rent or "Economic Rent" refers to that part of the payment made by a tenant to his landlord for the use of his land, i.e., free gift of nature only.

Definitions

1. According to Marshall, "the income derived from the ownership of land and other free gifts of nature is commonly called rent"
2. David Ricardo defined rent as "that portion of the produce of the earth which is paid to the landlord for the use of the original and indestructible powers of the soil".

Kinds of Rent

Economic Rent and Contract Rent

The payment made by an agriculturist (tenant) to his landlord need not be equal to economic rent. A part of this payment may consist of interest on capital invested in the land by the landlord in the form of buildings, fences, drainage; well etc. the part of the payment which is made for the use of land only is "Economic Rent". But the total amount consists of payment for the use of land, interest on capital invested in the land by the landlord in the form of buildings, fences, drainage, well etc., paid by a tenant to the landlord is "Contract Rent". It is called Contract Rent, because it is determined by a contract between the tenant and the landlord.

Situation Rent

Rent not only depends upon differences in the fertility of the land but also the differences in the situation of land. Some lands enjoy situational advantage. For example, some lands are situated very near to the market which may help to save a lot of transport costs. Even if all lands ate equally fertile, lands possessing situational advantage command some superiority over other lands. Thus, rent arises on account of situation is called "Situation Rent"

Scarcity Rent

Rent not only depends upon differences in fertility and situation of land but also the scarcity of land in relation to demand. If population increases, the demand for land increases. Therefore, the no-rent land is also subjected to intensive cultivation and it yields a surplus over cost. This surplus is the rent. This rent arises because of scarcity pf land and hence it is called "Scarcity Rent".

The Ricardian Theory of Rent

Introduction

The Classical theory of Rent in called "Ricardian Theory of Rent" after David Ricardo (1773-1823), an eminent English Classical Economist of the 19th Century. This theory is the basis for all discussions on the problem of rent.

How does Rent Arise?

Ricardo explains the emergence of rent with the help of an illustration. Suppose some people go to a newly discovered island and settle down there, there are three grades of land namely A,B, and C in that island. 'A' being the most fertile, 'B' less fertile and 'C' the least fertile. They will cultivate all the most fertile lands (A grade) available. Since the land in abundant and idle, there may be no need not pay rent as long as such best lands are freely available. Given a certain amount of labour and capital, the yield per acre on 'A' grade land is 40 bags of corn. Suppose another batch of people goes and settles down in the same island after some time. Therefore, population increases. Hence the demand for agricultural produce will increase. The most fertile land (A grade) alone cannot produce all the food grains that are needed on account of the operation of the Law of Diminishing Returns. So the less fertile land (B grade) will have to be brought under cultivation in order to meet the needs of the growing population. For the same amount of labour and capital employed in 'A' grade land, the yield per acre 'B' grade land is 30 bags of corn. This surplus of 10 bags per acre appears on 'A' grade land. This is "Economic Rent" of "A" grade land. Suppose another batch of people goes and settles down in the same island. Therefore, population increases still further. So the least fertile land (C grade) will have to be brought under cultivation. For the same amount of labour and capital, the yield per acre on 'C' grade land is 25 bags of corn. This, surplus of 'A' grade land is now raised to 15 bags, and it is the "Economic Rent" of the 'A' grade land. The surplus of 'B' grade land is 5 bags. This is the Economic rent of 'B' grade land.

In the above illustration, in 'C' grade land, cost of production is just equal to the price of its produce and therefore does not yield any rent. Hence, 'C' grade land is called "No-Rent Land" or "Marginal Land". Therefore, No-Rent Land or Marginal Land in the land in which cost of production is just equal to the price of its produce. The land which yields rent is called "Intra-Marginal Land".

The illustration discussed above can be shown through Table 4.1

Ricardian Theory of Rent

Table 4.1

Grades of land	Production (in bags)	Surplus (i.e., Rent in bags)
A	40	40-25=15
B	30	30-25=5

C	25	25-25=0

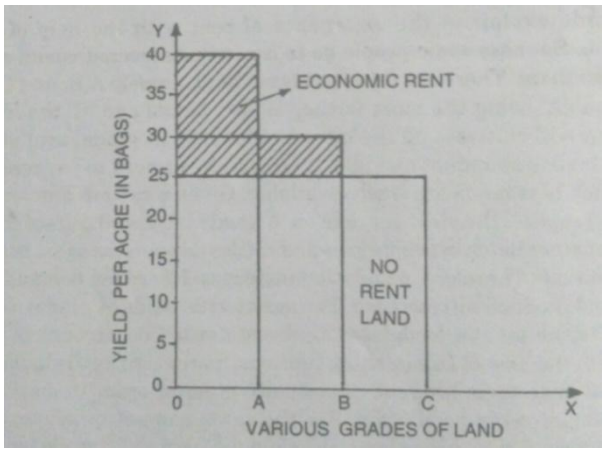


Fig.4.4

In fig… X axis represents various grades of land and Y axis represents yield per acre (in bags). OA, AB and BC are the A grade, B grade and C grade lands respectively. The application of equal amount of labour and capital on each of them give an yield represented by the rectangles standing just above the respective bases. The 'C' grade land is the "No-Rent Land", 'A' and 'B' grade lands are "Intra Marginal Lands". The Economic rent yield by 'A' and 'B' grade lands is equal to the shaded area of their perspective rectangles.

Assumptions

1. There is perfect competition in the economy
2. The supply of land is limited
3. There is difference in the fertility of land
4. There is marginal or no-rent land in the economy
5. Land has original and indestructible powers in the soil
6. The demand for agricultural products increases with population growth
7. The Law of Diminishing Returns applies
8. Rent arises in the long-run
9. Rent arises in the long-run
10. Land and capital are single factors
11. Rent is price determined
12. Land is cultivated in historical sequence, i.e., first the best land, then the less fertile and in this order.

Criticism

1. This theory assumes that cultivation of land is subject to the Law of Diminishing Returns. But the effect of the law could be checked by technological improvements
2. This theory also assumes Perfect Competition. But this assumption of Perfect Competition is unrealistic one because; in actual life Perfect Competition is very rare.
3. This theory is based on the assumption that rent does not enter into price. But in reality rent enters into price.
4. This theory assumes that there is no-rented land. But in reality, there may not be no-rented lands. Because, in practice, every land yields rent.
5. According to this theory, rent is due to the "original and indestructible powers of the soil". But the powers of the soil are not always original and also the fertile lands lost their fertility after being continuously cultivated. They are increased through scientific manuring.
6. Ricardo has ignored the alternative uses of land
7. According to Ricardo, most fertile lands are cultivated first. But most fertile lands are not always cultivated first.
8. It is not always possible to identify the powers of the soil
9. According to Ricardo, rent is due to the fertility of land. But critics pointed out that rent is due to fertility, but due to the scarcity of land.

Theories of wages

Meaning

Wages are a payment for the services of labour, whether mental or physical. Though in ordinary language an office executive, a minister or a teacher is said to receive a salary; a lawyer or a doctor a fee; and a skilled or unskilled worker a wage, yet in economics no such distinctions are made for different services and all if them are said to receive a wage. In other words, wages and less in terms of money wages and vice versa.

The subsistence theory of wages

The subsistence theory of wages was first formulated by the Physiocratic School of French Economists of 18th century. The theory was further developed and improved upon by German economists who called it as the Iron Law of Wages.

According to this theory, labour power is a commodity and its price is determined by its cost of production. Its cost of production is the minimum subsistence expenses required

for the support of the worker and his family so that continuous supply of labour is maintained. The exponents of this theory held that wages have a tendency to settle only at the subsistence level or cost of production. If at any time, wages exceeded this level, workers finding themselves in better economic position, would marry. As such, population would increase and with it the supply of labour. This tendency would continue to operate until wages were reduced to the minimum subsistence level. If, on the other hand, wages fell below the subsistence level, there would be reduction in population through starvation or diseases among workers, thereby bringing about a shortage in the supply of labour. This tendency would continue to operate until wages were again raised to the minimum subsistence level. Thus the subsistence theory of wages reaches the conclusion that the wage level can neither rise nor fall below the minimum subsistence level.

Criticism

This theory has been criticised on the following grounds

1. Wrongly based on the Malthusian Theory of Population: The subsistence theory of wages is based on the Malthusian Theory of Population which, in turn, is a highly controversial theory and never has its applications in Western economies
2. Historically not Correct: Critics also maintain that historically this theory has not been correct in its conclusions. Experience shows that a rise in wages is not necessarily accompanied by an increase in population, rather it is followed by a decline in the rate of growth of population
3. Unable to explain difference in wages: Labour market is characterised by heterogeneity of wage rates. If wage rates were to be equal to the subsistence level, then they would have been uniform for every one. This is also a proof that the theory is unrealistic.
4. One-sided theory: This theory approaches the problem of wage determination from the side of supply and completely ignores the demand side of labour. This theory is thus one sided.
5. Highly pessimistic: This theory is highly pessimistic for the working class. It presents a very dark picture of the future of the society
6. No attention to efficiency and productivity: According to this theory, wage rate for all workers tends to be equal to the minimum subsistence level. But it need be so because workers differ in efficiency and productivity.

Interest

Meaning

Interest is the reward paid by the borrower of capital to the lender of capital (capitalist) for the use of his capital. The followers of the Liquidity Preference Theory of Interest believe that interest is a price for surrendering liquidity preference.

Definition

According to Marshall, interest is "the price paid for the use of capital in any market" Keynes defines interest as "the premium which has to be offered to induce people to hold their wealth in some form other than hoarded money"

Keynes' Liquidity Preference theory of Interest (BM)

Introduction

Keynes propounded the Liquidity Preference Theory of Interest, in his famous book, "The General theory of Employment, Interest and Money". According to Keynes, interest is purely a monetary phenomenon because the rate of interest is calculated in terms of money. According to him, "interest is the reward for parting with liquidity" for a specified period

Keynes rejected the classical view of rate of interest as the reward of saving or waiting as such. The Classical Economists believed that saving was influenced by the rate of interest. But according to Keynes, saving depends upon the level of income. Of all assets, money is the most liquid asset and people prefer liquid form of asset. Therefore, if they are asked to surrender this liquidity, they must be rewarded. This reward is "Interest". According to Keynes, greater the desire for liquidity, higher shall be the rate of interest

Motives

According to Keynes, there are three motives for liquidity preference. They are:

1. The Transaction Motive
2. The Precautionary Motive
3. The Speculative Motive

1. The Transaction Motive

The transactions motive relates to the desire of the people to hold cash for the current transactions (i.e., day-to-day expenditure)

Individuals hold cash in order "to bridge the interval between the receipt of income and its expenditure". This is called the "Income Motive".

The businessmen and entrepreneurs hold cash in order to pay for wages and salaries, to pay for raw materials and transport and to meet all other current expenses incurred by business. This is called by Keynes as "Business Motive".

2. The Precautionary Motive

The Precautionary motive relates to the desire of the people to hold cash to meet unexpected expenditure such as sickness, accidents, fire, theft and unemployment.

Keynes described the combined sum of cash held for transactions and precautionary motives as "Active Balances" and labelled M_1. The demand for active balances can be referred to as L_1 and thus $L_1 = f(y)$.

3. The speculative Motive

The speculative motive relates to the desire of the people to hold cash in order to take advantage of market movements regarding the future changes in the price of bonds and securities in the capital market. If a man expects an increase in the price of bonds and securities in future, he may desire to hold money in liquid form. He does not hold more money, of he fear that the price of bonds and securities will fall in future. Any money held in excess of the amount required for transactions and precautionary motives has been described by Keynes as "Idle Balances"

Of these three motives, the speculative motive is more important in relation to the rate of interest.

Some times there will be a excessive purchase or excessive sale of securities in the stock exchange market. The tendency for excessive purchase of securities is known as the "Bullish sentiment" and the tendency for aggressive sale of securities is known as the "Bearish sentiment". The person who joins hands in these tendencies are called the members of the 'Bull' brigade and the 'Bear' brigade respectively.

The amount of money held under the speculative motive depends upon the rate of interest. Bond prices and the rate of interest are inversely related to each other. Low bond prices are indicative of high interest rates. Let us suppose that an individual wants to purchase a bond of Rs.1000, which carries 4% interest. Now on this bond, he earns an income of Rs.40. When the rate of interest goes up to 5 per cent, to earn the same income of Rs.40, an investment of Rs.800 will be sufficient. It means that bond price decrease with an increase in the rate of interest and vice versa. This shows that when the rate of interest is high, people will have a desire to hold less money for speculative purposes. When the rate of interest decreases, the bond holders sell their bonds and hold liquid cash. It means that lower the rate

of interest, greater will be the liquidity preference. Thus people will prefer to hold more or less cash depending on the future changes that are likely to take place in the prices of assets

Thus higher the rate of interest, the lower will be the speculative demand for money and vice versa. Geometrically, it is a smooth curve which slopes downwards from left to right, indicating that more money will be held at a lower rate if interest. Algebraically, the speculative demand for money is expressed as $M_2 = f(r)$

According to Keynes, Liquidity preference curve (LPC) becomes perfectly elastic at a certain rate of interest, such as OR_2. At a very low rate of interest people prefer to keep money in cash rather than lend it.

The argument is that at abnormally low interest rates virtually everyone would expect the interest rate to rise towards its normal level in the near future. In this situation, virtually, everyone would be expecting a fall in the price of bonds and therefore, capital losses are expected for bond holders.

The flat portion or the perfectly elastic portion of the liquidity preference curve is known as the liquidity trap.

Implication of the Liquidity Trap

The concept of liquidity trap has certain important implications.

Firstly, the monetary policy is out of commission. It cannot influence the rate of interest by following a cheap money policy. In this extreme case, the velocity of circulation falls as all increases in the money supply are added to 'idle' balances and the monetary policy is helpless as it cannot cut down the interest rate any further

Secondly, the rate of interest cannot fall to zero

Lastly, the policy of a general wage cut cannot be effective in the perfectly elastic portion of the liquidity preference curve.

Total Demand for Money

Since transactions and precautionary demand for money are income elastic, both can be lumped together. It is represented by M.

We know that $M1 = f(y)$

Likewise the speculative demand for money is expressed as M2

$M2 = f(r)$

The aggregate demand for money is $M = (M1 + M2)$

Therefore $M = f(y, r)$

Supply of money

Supply of money refers to the total quantity of money in the country for all purposes at any given time. The total supply of money consists of legal tender money and credit money or bank money. Money supply is a policy variable, and its supply is determined by the central bank of a country. At any particular point of time the supply of money is fixed. Given the supply of money, the rate of interest is determined by the demand for money.

Determination of the Rate of Interest

Like the price of any product, the rate of interest is determined at the level where the demand for money equals the supply of money. In the following figure, the vertical line QM represents the supply of money and L the total demand for money curve. Both the curve intersect at E_2 where the equilibrium rate of interest OR is established

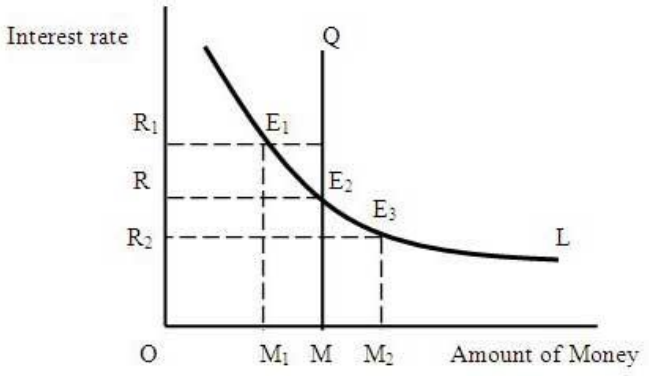


Fig.4.5

If there is any deviation from this equilibrium position an adjustment will take place through the rate of interest, and equilibrium E_2 will be re-established At the point E_1 the supply of money OM is greater than the demand for money OM_1. Consequently, the rate of interest will start declining from OR_1 till the equilibrium rate of interest OR is reached.

Similarly at OR_2 level of interest rate, the demand for money OM_2 is greater than the supply of money OM. As a result, the rate of interest OR_2 will start rising till it reaches the equilibrium rate OR It may be noted that, if the supply of money is increased by the monetary authorities, but the liquidity preference curve L remains the same the rate of interest will fall. If the demand for money increases and the liquidity preference curve sifts upward, given the supply of money, the rate of interest will rise

Criticism

1. Keynes considered interest purely as a monetary phenomenon, but rate of interest is not purely a monetary phenomenon. He ignores the real factors like productivity, waiting and time preference.

2. According to Keynes, "interest is the reward for parting with liquidity" for a specified period. But some critics pointed out that interest is not a reward for parting with liquidity but for the productivity of capital
3. In this theory, the term "liquidity" has not been clearly defined
4. Keynes makes the rate of interest independent of the demand for investment funds. It is not so independent. This demand for investment funds depends upon the marginal revenue productivity of capital
5. This theory does not explain the existence of different rates of prevailing in the market at the same time.
6. This theory only explains interest in the Short-run. It does not explain interest in the Long-run
7. According to Keynes, the rate of interest is determined by the speculative demand for money and the supply of money. Given the total supply of money, we cannot know how much money will be available to satisfy the speculative demand for money, unless we know how much the transactions demand for money is. And we cannot know the transaction demand for money, unless we know the level of income. Thus, the Keynesian theory, like the Classical Theory is indeterminate.

Profit

The Wage Theory of Profit

This theory was popularized and put forward by Prof. Taussig and Davenport the two most prominent economists. According to them – "Profits are best regarded as simply a form of wages. They accrue to the entrepreneur on account of his special ability."They have argued that there is very close similarity between a labourer and entrepreneur. Just as labourers receive wages for his services, similarly entrepreneurs receive profit for his service.

The entrepreneur performs mental labour like—teachers, doctors, lawyers etc. But the only difference between entrepreneur and other mental workers is that the entrepreneur receives profit for his special ability and hard work. This is a surplus amount which the entrepreneurs receive after meeting all expenses of production where as the wage forms a part of the cost of production.

Criticisms

a. Element of risk and uncertainty:

The entrepreneur's work is full of risk and uncertainty and profit is given to face this risk. But the workers receive wages simply for his labour. Risk and uncertainty part do not

incorporate anywhere in his activities. For labourer risk is of losing the job which is an extreme step.

b. Profit is flexible, it may vary:

Profits may rise or fall. It depends upon the business conditions and situations. But wage may remain stable and cannot fluctuate more in the short- period.

c. This theory is silent over the payment to shareholders:

The shareholders of any organisation or company do not perform any function but they receive the share of profits in the form of dividend for undertaking risk of money invested. This theory fails to explain this contention as to why they are paid.

d. Entrepreneurs windfall or chance profits:

The entrepreneur may receive windfall or chance profits but a worker cannot have opportunity to get wages of chance or windfalls.

The Marginal Productivity Theory of Profit

This theory was propounded by Prof. Marshall. According to him, "Profit is equal to the marginal productivity of the entrepreneur. He has said that the amount which the community is liable to produce with the help of entrepreneur over and above what it could produce with his help."

Recently Stigler and Stonier and Hague have said that "Profit is the reward of an entrepreneur which is determined by its marginal revenue productivity, the higher are the profits and lower the marginal revenue productivity, the lower are the profits of an entrepreneur."

Criticisms

a. This theory is based on unrealistic assumptions:

These unrealistic assumptions are homogeneity of entrepreneurs in an industry. As entrepreneurs' efficiency differ, therefore it is not possible that there will be one marginal revenue productivity curve for all entrepreneurs. So Profit cannot be same.

b. This theory fails to determine profit accurately:

Because efficiency of entrepreneurs differs, systems and methods of doing work differ, therefore. Profit cannot be calculated accurately.

c. The concept of marginal revenue productivity of entrepreneurship is a meaningless concept:

Because unlike other factors, there can be only one entrepreneur in a firm.

d. It is one sided theory:

This theory takes into account only the demand for entrepreneurs and do not take into account the supply or availability of entrepreneurs.

e. This is a static theory:

Where all entrepreneurs earn only normal profits, they have not considered that the world is dynamic also where some entrepreneurs can earn more than normal profits.

f. Not taken windfall gain and loss:

This theory has not taken into account the windfall or chance or gain or even monopoly profits.

The Dynamic Theory of Profit

Prof. J. B Clark propounded this theory in the year 1900. According to him – "Profit is the difference between the price and the cost of the production of the commodity". But Profit is the result of dynamic change. Further, Prof. Clark was of this opinion that in a stationary state having static economic conditions of demand and supply, there can be no real or pure profit as a surplus. In a stationary economy, the quantum of capital invested, methods of production, managerial organisation, technology, demand pattern etc. remain constant.

Under competitive conditions, price tends to equal average costs; hence, the surplus is zero. So, no pure profit but there may be some frictional profits emerging due to frictions in the system. But, this cannot be regarded as real Profits.

Profit is the result exclusively of six dynamic changes i.e.:

(1) Changes or increase in population,

(2) Changes in tastes and preferences,

(3) Multiplication of wants,

(4) Capital formation,

(5) Technological advancement and

(6) Changes in the form of business organisation.

On account of these changes the economy tends to be dynamic. Demand and supply conditions are altered. Some entrepreneurs may get advantageous business positions against others and may reap surplus over costs, as a real profit. In short, those who takes advantage of changing situation can earn real profits according to their efficiency.

Inefficient and careless producers who fail to move with dynamic changes may not get any real profit and may even incur losses. Thus, Clark's dynamic theory of Profit has an element of truth as it emphasis the dynamic aspect of Profit.

Criticisms

a. All changes are not foreseen:

Clark's theory fails to make any difference between a change that is foreseen and one that is unforeseen in advance. If the six generic changes as assumed by Prof. Clark are to be foreknown in advance then the effects of changes will not hold at all. In reality, all changes are not foreseen. Some are foreseen and some are not. So, to have a clear understanding of the problem, it is essential to separate its effects from those of change as such.

b. This theory gives artificial dichotomy:

In this connection Taussig has said that Clark's theory gives an artificial dichotomy of 'Profit' and 'Wages of management'.

c. All changes do not lead to Profit:

Clark's theory suggests that all dynamic changes lead to Profit. But critics are of this opinion that only unpredictable changes would give rise to profits. Predictable changes will not cause surplus to emerge on account of precise adjustments.

d. Here, the concept of frictional Profit is vague:

Clark's theory indicates that in a stationary state, there is only a frictional profit. But the concept of frictional profit is vague. But it is the normal profit which is earned in a stationary state.

e. Element of risk involved in business:

Clark's theory of Profit do not stress the element of risk involved in business due to dynamic changes. The best course is to combine elements of risk dynamic changes to understand the true nature of profit in a modern economy.

F.W. Hawley's the Risk Theory of Profit

This theory of Profit is associated with F. B. Hawley who has considered the risk-taking as the important function of an entrepreneur. The entrepreneur exposes his business to risk, and in turn he receives a reward in the form of Profit because the task of risk-taking is irksome.

It is definite that no entrepreneur will like to undertake risks if he gets only the normal return. Therefore, the reward for risk-taking must be higher than the actual value of the risk. Further, it has been said that the actual value of the risk.

Further, it has been said that more risky the business, the higher is the expected Profit rate. As Professor D. M. Holland has said that "riskier the industry or firm, the higher is its Profit rate." But he was warned that this tentative view must be tested in depth.

Criticisms

a. There cannot be functional relationship between Risk and Profit:

Those persons who dare to take high risks in certain businesses may not necessarily earn high profits.

b. Profit is not based on entrepreneur's ability:

In this connection Prof. Carve has said that "Profit is not based on entrepreneur's ability to undertake the risks of the business, but rather as his capability of risk avoidance."

c. It is an incomplete theory:

From business point of view, all enterprises are risky and an element of uncertainty is present there. But every entrepreneur aims at making large profits which is also uncertain. Therefore, Hawley's Risk Theory can also be called as an incomplete theory of Profit.

d. Amount of Profit not related to size of risk involved:

The amount of Profit is not in any way related to the size of the risk undertaken. If it were so related then every entrepreneur would involve himself into huge risks in order to earn larger profits.

e. Concentrates mostly on risk and not on anything else:

This theory mostly disregards many other factors attributable to Profit and just concentrate on risks and risks alone

Modern Theory or Perfect Competition or Demand and Supply Theory of Profit:

This modern theory of Profit defines the entrepreneur as a business enterprise itself and 'Profits' as his net income. In this theory profits have been regarded as the reward of an entrepreneur and are governed by the demand for and supply of entrepreneur.

Demand for Entrepreneurs:

The demand for entrepreneurs mostly depends upon the level of industrial development, the elements of uncertainty in the industry, the scale of production and the marginal revenue productivity of entrepreneurship. If the level of industrial progress is high, the scale of production is large and efficiency and productivity increase, the profits will be high. The marginal revenue productivity of entrepreneurship is the most important factor in influencing the demand for entrepreneurs.

Supply of Entrepreneurs:

Similarly, the supply of entrepreneurs depends upon various factors like the availability of capital, the existence of managerial and technical personal, the number of entrepreneurs and the condition of society etc. The larger the availability of capital, the larger

is the supply of entrepreneur's capital may be available in sufficient amount, but an entrepreneur has to depend largely on the managers and other technical personal for organising and running the business successfully.

If trained managerial and other personal are available in the market, the supply of entrepreneurs is bound to increase. Further, the economists are also of this opinion that the size of population is another factor that influences entrepreneurship. The larger the size of population, the higher will be the demand for various products which will attract more people to entrepreneurship and the supply of entrepreneurs will increase.

Criticisms

While criticising this theory Knight has said that Profit has been regarded as the reward for bearing non-insurable risks and uncertainties, then under perfect competition there can be no profit in the long-run. It is a static state where population, capital, technology, tastes, business organisation and income do not change.

If they change they can be predicted. Thus, there is no risk and uncertainty. The marginal revenue productivity curve of entrepreneurship would be zero. Therefore, Profit will also be zero. In a static state, profits exist because Profits are not competed away due to the presence of imperfect competition. So what entrepreneurs earn are monopoly profits rather than pure profits. It should be remembered that Manager-entrepreneurs earn wages of management and capitalist - entrepreneurs earn interest.

Prof. Schumpeter's Innovation Theory of Profit

Schumpeter deemed Profit as the reward to enterprise and innovation. In his opinion, the entrepreneur initiates innovation in the business and when he succeeds, he earns Profit as his reward. Now, the question is what is innovation? "Innovation means commercial application of new scientific inventions and discoveries."

An innovator is, therefore a businessman with vision, foresight, originality and is bold enough to bear high risks involved in undertaking new activities on a new basis. The innovator is not a scientist, but he successfully introduces new inventions on a commercial basis.

In giving opinion over this Samuelson has written as an example – "The scientific theory of radio wave was the brain-work of Maxwell. It was experienced upon by Hertz and its commercially profitable use was carried out by Marconi and Sarnoff, who are the innovators in radio manufacturing."

Innovation is of two types:

(i) Product innovations, and

(ii) Market innovations.

Product innovations affect the cost and quality of the product while market innovations include discovery and exploitation of new market, introducing new variety of products and product improvement, modes of advertising and sales propaganda etc. It has been said that any form of innovation leads to a Profit. It is called as innovational profit. This Profit is uncertain and unpredictable. It is temporary in nature.

Criticisms

a. Schumpeter has never considered Profit as the reward for risk-taking:

He is of this opinion that risk-taking is the function of the capitalist and not of the entrepreneur. It is the shareholders who undertake risks and thus earn profits.

b. There is no place of uncertainty in Schumpeter's innovation theory:

Profit is not the reward of uncertainty it is simply the wages of management.

c. This theory is incomplete:

Profit accrues to the entrepreneur for his organisational ability and nothing else. Therefore, this theory has been called as an incomplete explanation of the emergence of profits.

UNIT V : WELFARE ECONOMICS

Nature of Welfare Economics - Pigovian of Welfare Economics and Externalites Pareto's Optimum Social Welfare Criterion - Kaldor - Hicks New Welfare compensations Principle - Scitovsky's Double Criterion of new welfare - Bergson's Social welfare Function - Arrow's impossibility Theorem

A.C. Pigou's Economic of Welfare

Meaning of Welfare:

According to Pigou, welfare resides in a man's state of mind or consciousness which is made up of his satisfactions or utilities. The basis of welfare, therefore, is necessarily the extent to which an individual's desires are met.

Social welfare is regarded as the summation of all individual welfares in a society. Since general welfare is a very wide, complicated and impracticable notion, Pigou delimits the range of his study to economic welfare. As he himself observes, economic welfare is by no means an index of total welfare because many other elements in the latter, like the quality of work, one's environment, human relationships, status, housing, and public security are absent from economic welfare.

He, therefore, defines economic welfare as "that part of social (general) welfare that can be brought directly or indirectly into relation with the measuring rod of money." Thus economic welfare, in the Pigovian sense, implies the satisfaction of utility derived by an individual from the use of exchangeable goods and services.

Pigovian Welfare Conditions:

Pigou regard economic welfare and national income as essentially coordinate. It is on this basis that he lays down two conditions for maximisation of welfare. The first condition states that welfare is said to increase when national income increases. Given the same tastes and income distribution, an increase in the national income represents an increase in welfare. Pigou contends that in most cases the national income would increase even though the disutility of work also increases.

Second, for welfare maximisation the distribution of the national income is equally important. If national income remains constant, transfers 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 position of the latter is raised. This welfare condition is based on the dual Pigovian postulates of 'equal capacity for satisfaction and diminishing marginal utility of income.'

Pigou argues that different people derive the same satisfaction out of the same real income and that "people now rich are different in kind from the people now poor having in their fundamental nature greater capacities for enjoyment." With income subject to diminishing marginal utility, transfers of income from the rich to the poor will increase social welfare by satisfying the more intense wants of the latter at the expense of the less intense wants of the former. This it is economic equality that maximises welfare.

Dual Criterion:

To find out improvements in social welfare, Pigou adopts a dual criterion:

First, an increase in the national income 'brought about either by increasing some goods without dimishing others or by transferring factors to activities in which their social value is higher,' is regarded an improvement in welfare without reducing the share of the poor.

Second, any reorganization of the economy which increases the share of the poor without reducing the national income is also considered an improvement in social welfare.

Assumptions of Pigovian Conditions:

(1) Each individual tries to maximise his satisfaction from his expenditure on different goods and services.

(2) Satisfactions are comparable both interpersonally and interpersonally.

(3) The law of diminishing marginal utility of income applies. It means that the marginal utility of income falls, as income increases. As a result, the gain in utility of an additional amount of income to a poor man is greater than the loss of utility to a rich man from the same amount of income.

(4) There is equal capacity for satisfaction. It implies that different people derive the same satisfaction out of the same real income. Given these assumptions, it is possible to satisfy the Pigovian conditions of maximum social welfare on the basis of his dual criterion.

Criticism:

(1) The Notion of Maximisation is not clear:

Pigou lays emphasis on the maximisation of welfare, but he does not clarify the notion of maximisation. His 'maximum' is in fact the 'optimum', but it is a stable point. But this is not a correct view because the 'optimum' is not stable. It changes with the increase or decrease of national income.

(2) Pigou measures 'welfare' cardinally:

According to Pigou, welfare is measured in terms of utility or satisfaction. He regards social welfare as the summation of utilities of exchangeable goods and services to individuals. Economists do not agree with this view because quantitative measurement of utility is not possible. It is for this reason that modern economist's measure utility ordinally.

(3) National Income is not an Accurate Measure of Welfare:

Pigou's welfare conditions are related to the national income. But it is not easy to calculate national income. Again, social welfare does not increase by a mere increase in national income. It is possible that national income may increase due to inflationary rise in prices and the poor may become worse-off than before.

It is due to these reasons that modern economists measure welfare on the basis of 'choice' rather than by national income. For instance, when an individual chooses bundle A of some good rather than bundle 5, he undoubtedly derives greater satisfaction or utility from A. It is in this way that his welfare increases.

(4) Equal capacity for satisfaction:

According to Professor Robbins, Pigou's assumption of man's equal capacity for satisfaction does not make his notion of welfare a positive study. In his words, "This assumption rests on ethical principle rather than upon scientific demonstration; it is not a judgement of value."

(5) Pigou does not clarify the ethical relation of welfare:

Welfare economics is closely related to ethics but Pigou does not clarify it. Welfare economics is essentially a normative study in which value judgements and interpersonal comparisons are made. By not relating these concepts with his notion of welfare, Pigou s economics of welfare is not considered as an objective study of the causes of welfare.

Meaning of Externalities:

Divergences between private and social costs and benefits are known as externalities, external effects or external economics and diseconomies. Another term is spillovers or "neighbourhood effects". An external effect is assumed to exist whenever the production by a firm or the utility of an individual depends on some activity of another firm or individual through a means which is not bought and sold, such a means is not marketable, at least at present.

In other words, externalities may run from production to production and from production to consumption. They may also run from consumption to consumption and from

consumption to production. There are positive and negative externalities. The beneficial externalities are called positive externalities.

The costly externalities are called negative externalities. In other words, if social benefits exceed private benefits, it is a positive externality or external economy. If social costs exceed private costs, it is a negative externality or external diseconomy. Externalities are, in fact, market imperfections where the market offers no price for service or disservice.

These externalities lead to misallocation of resources and cause production or consumption to fall short of an optimum level. Thus they do not lead to maximum social welfare. Pigou's major contribution lies in studying the main causes leading to divergences between private and social costs and benefits and in suggesting measures for removing these divergences.

The private product diverges from the social product due to the existence of external economies or diseconomies thereby leading to divergenies between private and social costs and benefits. We analyse these external economies and diseconomies in the light of Pigou's analysis.

Positive Externalities of Production:

According to Pigou, when a firm renders a benefit or cost of a service to other firms without appropriating to itself all the benefits or costs of his service, it is an external economy of production. External economies of production accrue to one or more firms in the form of reduced average costs as a result of the activities of another firm.

In other words, these economies accrue to other firms in the industry with the expansion of a firm. They may be the result of reduced input costs which lead to pecuniary external economies. Whenever external economies exist, social marginal benefit will exceed private marginal benefit and private marginal cost will exceed social marginal cost.

This is illustrated in Figure 5.1, where PMC (or supply) is the private marginal cost curve of firms. The demand curve D intersects the PMC curve at point E and determines the competitive market price OP and output OQ.

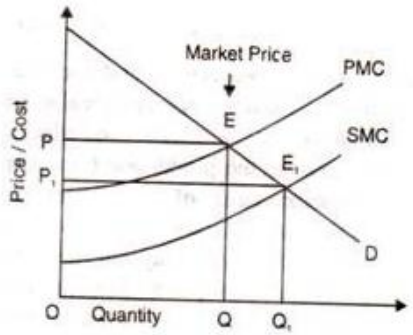


Fig.5.1

SMC is the social marginal cost curve which intersects the demand curve at point E_1 and determines the social optimum output level OQ_1, at price OP_1. Since for every unit of output between OQ and OQ_1, The social marginal cost (OP_1) is less than the competitive market price OP, its production involves a net social gain equal to QQ_1.

Negative Externalities of Production:

When the production of a commodity or service by a firm affects adversely other firms in the industry, social marginal cost is higher than social marginal benefit. Suppose a factory situated in a residential area emits smoke which affects adversely health and household articles of the residents.

In this case, the factory benefits at the expense of residents who have to incur extra expenses to keep themselves healthy and their households clean. These are social marginal costs because of harmful externalities which are higher than private marginal cost and also social marginal benefit.

This is illustrated in Fig.5.2, where the SMC curve is above the PMC curve which intersects the D curve at point E and determines the competitive price OP and output OQ. But the socially optimum output is OQ_1 and price the is OP_1 as determined by the intersection of SMC and D curves at point E_1.

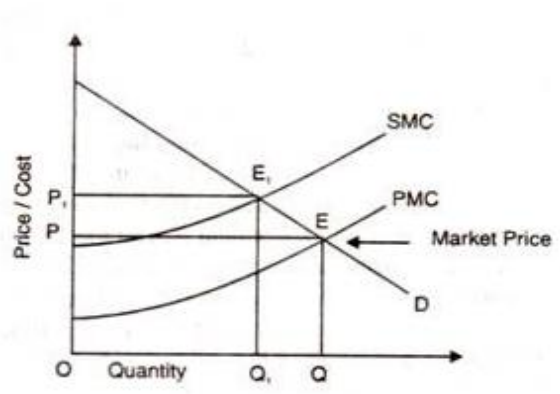


Fig.5.2

Thus the firms are producing Q_1 Q more than the social optimal output OQ_1. In this case, for every unit between Q_1 and Q, social marginal cost (SMC) is more than the competitive market price OP. Thus its production involves a social loss, i.e., $OQ - OQ_1$.

Positive Externalities in Consumption:

Externalities in consumption lead to non-attainment of Pareto optimality. External economies of consumption arise from non-market interdependencies of the satisfactions enjoyed by different consumers. An increase in the consumption of a good or service which affects favourably the consumption patterns and desires of other consumers is an external economy of consumption When an individual installs a TV set, the satisfaction of his neighbours increases because they can watch TV programmes free at his place.

Here social benefit is larger and social cost is lower than the private benefit and cost. But the TV owner is likely to use his TV set to a smaller extent than the interests of society require because of the inconvenience and nuisance caused by his neighbours to him.

Negative Externalities in Consumption:

Negative externalities in consumption arise when the consumption of a good or service by one consumer leads to reduced utility (dissatisfaction or loss of welfare) of other consumers. Negative externalities in consumption arise in the case of fashions and articles of conspicuous consumption which reduce their utility to some consumers.

For example, smokers cause disutility to non-smokers, and noise nuisance from stereo systems to neighbours etc. Such diseconomies of consumption prevent the attainment of Pareto optimality.

Suppose there are two room-mates A and B. Individual A likes to smoke while individual B likes clean air. Further, B's utility of consuming clean air is affected by individual A's smoking. This is explained in terms of Figure.5.3 (A) & (B).

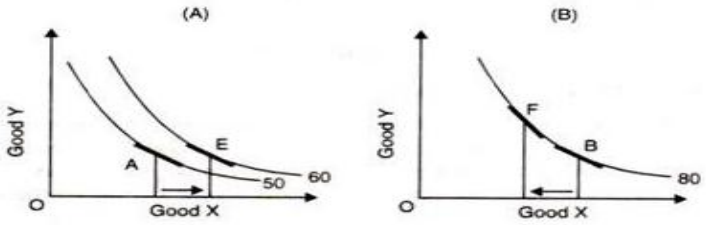


Fig.5.3

Initially, individual A's utility from smoking gives him 50 utilis at point A while, individual B's consumption of clean air gives him 80 utilis At point B. When there are no externalities in consumption, the tangent at point A and point B are parallel to each other.

If individual A smokes at his leisure then his utility increases to 60 utilis and he move to point E. The effect of individual A's smoking reduces the utility of clean air to individual B who moves from point B to point F on the same utility curve.

Individual A has moved on a higher utility curve from 50 to utility curve 60, but the non-smoker is on the same utility curve 80. Thus Pareto optimality is not attained because the utility of one consumer (smoker) A has increased whereas the utility level of the other consumer (non-smoker) B has been reduced.

Conclusion:

These drawbacks in the Pigovian analysis have led modern economists to expound the 'compensation principle' and the 'social welfare function' which are attempts at giving a new tinge to welfare economics.

Pareto Optimality and Economic Efficiency:

Economists defined social welfare as a sum total of cardinally measurable utilities of different members of the society. An optimum allocation of resources was one which maximised the social welfare in this sense. V. Pareto was the first to part with this traditional approach to social welfare in two important respects.

First, he rejected notion of cardinal utility and its additive nature and, second, he detached welfare economics from the inter-personal comparisons of utilities. Pareto's concept of maximum social welfare which is based upon ordinal utility and is also free from value judgements occupies a significant place in modern welfare economics.

Pareto optimum may not be sufficient condition" for attaining maximum social welfare but it is a necessary condition for it. To repeat, Pareto optimum (often called Economic Efficiency) is a position from which it is impossible to make anyone better off

without making someone worse off by any reallocation of resources or distribution of outputs.

Thus, in the Pareto optimum position the welfare of any individual of the society cannot be increased without decreasing the welfare of another member. Before explaining the conditions of achieving Pareto optimality, we shall explain Pareto criterion of evaluating changes in social welfare because the concept of Pareto optimality or maximum social welfare is based upon Pareto criterion of welfare.

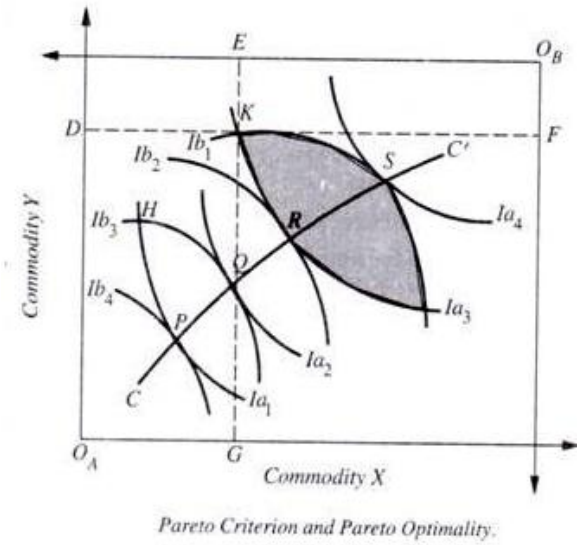
Pareto Criterion of Social Welfare:

The concept of Pareto optimum or economic efficiency stated above is based on a welfare criterion put forward by V. Pareto. Pareto criterion states that if any reorganisation of economic resources does not harm anybody and makes someone better off, it indicates an increase in social welfare. If any reorganisation or change makes everybody in a society better off, it will, according to Pareto, undoubtedly mean increase in social welfare.

Thus, in the words of Prof. Baumol "any change which harms no one and which makes some people better off (in their own estnuauon) must be considered to be an improvement." Parto criterion can be explained with the help of Edgeworth Box diagram which is based on the assumptions of ordinal utility and non-interpersonal comparison of utilities.

Suppose two persons A and B form the society and consume two goods X and Y. The various levels of their satisfaction by consuming various combinations of the two goods have been represented by their respective indifference curves.

In Figure 5.4, O_a and O_b are the origins for the utilities of two persons A and B respectively. I_{a1}, I_{a2}, I_{a3}, I_{a4} and I_{b1}, I_{b2}, I_{b3}, I_{b4} are their successively higher indifference curve. Suppose the initial distribution of goods X and Y between the members of the society, A and B, is represented by point- K in the Edgeworth Box.



Pareto Criterion and Pareto Optimality.

Fig.5.4

Accordingly, individual A consumes O_AG of X + GK of Y and is at the level of satisfaction represented by indifference curve I_{a3}. Similarly, individual B consumes KF of X+ KE of Y and gets the satisfaction represented by indifference curve Ib_1.

Thus the total given volume of goods X and Y is distributed between A and B. In this distribution, individual A consumes relatively large quantity of good Y and individual B of good X. Now, it can be shown with the aid of Pareto's welfare criterion that a movement from the point K to a point such as S or R or any other point in the shaded region will increase social welfare.

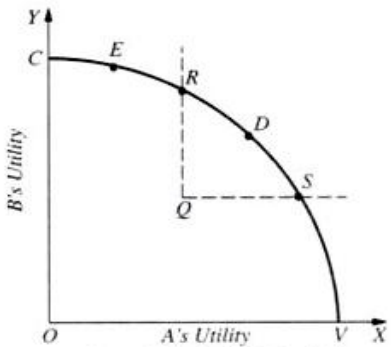
Any movement from K to S through redistribution of two goods between two individuals increases the level of satisfaction of A without any change in the satisfaction of B because as a result of this A moves to his higher indifference curve I_{a4}, and B remains on his same indifference curve I_{b1} (K and 5 lie on B's same indifference curve I_{b1}).

In other words, as a result of the movement from K to S, individual A has become better off whereas individual B is no worse off. Thus, according to Pareto criterion, social welfare has increased following the movement from K to S and therefore K is not the position of economic optimum.

Similarly, the movement from K to R is also desirable from the point of view of social welfare because in this individual B becomes better off without any change-in-the satisfaction of individual A. Therefore, both the positions S and R are better than K. The tangency points of the various indifference curves of the two individuals of the society are the Pareto optimum points and the locus of these points is called 'contract curve'.

Pareto criterion can also be explained with the help of Samuelson's utility possibility curve. Utility possibility curve is the locus of the various combinations of utilities obtained by two persons from the consumption of a particular bundle of goods.

In Figure 5.5, CV is a utility possibility curve which shows the various levels of utilities obtained by two individuals A and B of the society resulting from the redistribution of a fixed bundle of goods and its consumption by them.



Pareto Criterion Explained with Utility Possibility Curve

Fig.5.5

According to Pareto criterion, a movement from Q to R, or Q to D, or Q to S represents the increase in social welfare because in such movements the utility of either A or B or both increases. A movement from Q to R implies that the utility or welfare of B increases, while that of A remains the same.

On the other hand, a movement from Q to S implies that while A has become better off, B is no worse off. And a movement from Q to D or any other point on the segment between R and S will mean increase in welfare or utility of both the individuals. Thus points R, D and S are preferable to Q from the point of view of social welfare.

But unfortunately Pareto criterion does not help us in evaluating the changes in welfare if the movement as a result of redistribution is from the point Q to a point outside the segment RS; such as point E on the utility possibility curve CV. As a result of the movement from point Q to E, the utility of A decreases while that of B increases. In such circumstances, Pareto criterion cannot tell us as to whether social welfare increases or decreases.

Marginal Conditions of Pareto Optimality:

Pareto concluded from his criterion that competition leads the society to an optimum position but he had not given any mathematical proof of it, nor he derived the marginal conditions to be fulfilled for achievement of the optimum position. Later on, Lerner and

Hicks derived the marginal conditions which must be fulfilled for the attainment of Pareto optimum.

These marginal conditions are based on the following important assumptions:

1. Each individual has his own ordinal utility function and possesses definite amount of each product and factor.
2. Production function of every firm and the state of technology is given and remains constant.
3. Goods are perfectly divisible.
4. A producer tries to produce a given output with the least-cost combination of factors.
5. Every individual wants to maximise his satisfaction.
6. Every individual purchases some quantity of all goods.
7. All factors of production are perfectly mobile.

Given the above assumptions various marginal conditions (first-order conditions) required for the achievement of Pareto optimum or maximum social welfare are explained below:

The Optimum Distribution of Products among the Consumers: Efficiency in Exchange:

The first condition relates to the optimum distribution of the goods among the different consumers composing a society at a particular point of time. The condition says: "The marginal rate of substitution between any two goods must be the same for every individual who consumes them both."

The marginal rate of substitution of one good for another so as is the amount of one good necessary to compensate for the loss of a marginal unit of another so as to maintain a constant level of satisfaction. So long as the marginal rate of substitution (MRS) between two goods is not equal for any two consumers, they will enter into an exchange which would increase the satisfaction of both or of one without decreasing the satisfaction of the other.

This condition can be better explained with the help of the Edgeworth Box diagram. In Figure 5.6 goods X and Y, which are consumed by two individuals A and B composing a society are represented on the X and Taxes respectively. O_A and O_B are origins for A and B respectively.

la_1, Ia_2, Ia_3 and Ib_1, Ib_2, lb_3 are the indifference curves showing successively higher and higher satisfaction of consumers A and B respectively. CC is the contract curve passing through various tangency points Q, R, S of the indifference curves of A and B.

The marginal rates of substitution (MRS) between the two goods for individuals A and B are equal on the various points of the contract curve CC'. Any point outside the

contract curve does not represent the equality of MRS between the two goods for two individuals A and B of the society.

Let us consider point K where indifference curves I_{a1} and I_{b1} of individuals A and B respectively intersect each other instead of being tangential. Therefore, at point K marginal rate of substitution between two goods X and Y (MRS_{XY}) of individual A is not equal to that of B.

With the initial distribution of goods as represented by point K, it is possible to increase the satisfaction of one individual without any decrease in that of the other or to increase the satisfaction of both by redistribution of the two goods X and Y between them. A movement from K to S increases the satisfaction of A without any decrease in B's satisfaction.

Similarly, a movement from K to Q increases B's satisfaction without any decrease in A's satisfaction. The movement from K to R increases the satisfaction of both because both move to their higher indifference curves. Thus, movements from K to Q or to S or to any other point on the segment SQ of the contract curve will, according to Pareto criterion, increase the level of social welfare.

It follows that movement from any point away from the contract curve to a point on the relevant segment of the contract curve will mean increase in social welfare. At any point away from the contract curve in the Edgeworth box, the indifference curves of the two individuals will intersect which will mean that MRS_{xy} of two individuals is not the same.

And, as explained, this indicates that through exchange of some units of goods between them, they can move to some point on the contract curve where the social welfare (that is, welfare of two individuals taken together) will be higher.

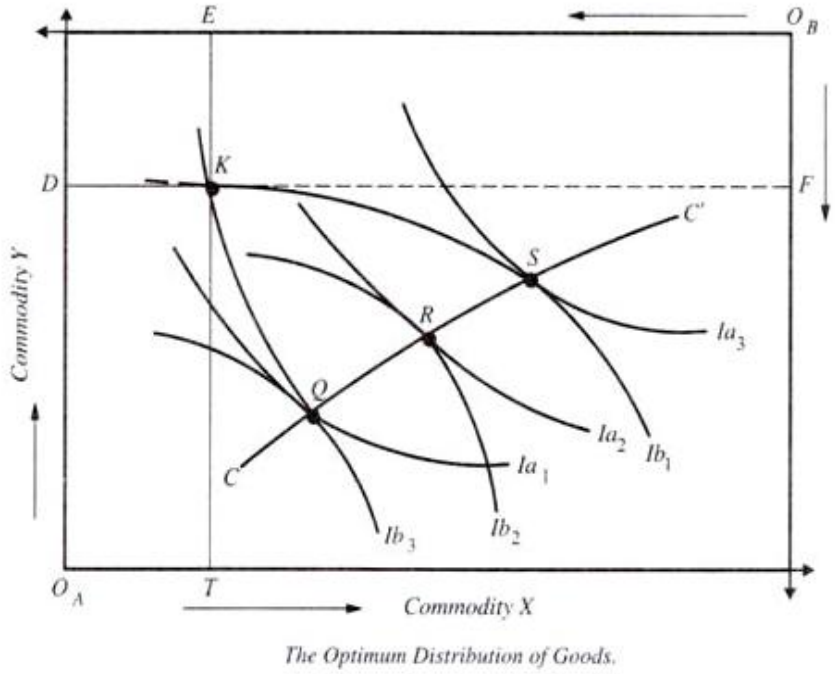


Fig.5.6

Since the slope of an indifference curve represents the marginal rate of substitution (MRS_{XY}) at any point of the contract curve, which represents tangency points of the indifference curves, MRS_{XY} of the two individuals are equal. Therefore, points on the contact curve represent the maximum social welfare.

However, a movement along the contract curve in either direction will make one individual better off and the other worse off since it will put one individual on his successively higher indifference curves and the other on his successively lower indifference curves. Thus, every point on the contract curve denotes maximum social welfare in the Paretian sense but we cannot say anything about the best of them with the help of Pareto criterion.

The Optimum Allocation of Factors: Pareto Efficiency in Production:

The second condition for Pareto optimum requires that the available factors of production should be utilised in the production of products in such a manner that it is impossible to increase the output of open firm without a decrease in the output of another or to increase the output of both the goods by any re-allocation of factors of production.

This situation would be achieved if the marginal technical rate of substitution between any pair of factors must be the same for any two firms producing two different products and using both the factors to produce the products.

This condition too can be explained with the help of Edgeworth Box diagram relating to production. This is depicted in Fig. 5.7. Let us assume two firms A and B producing the same product by using two factors labour and capital. The available quantities of labour and capital are represented on X and Faxes respectively. O_A and O_B are the origins for firms A and B respectively.

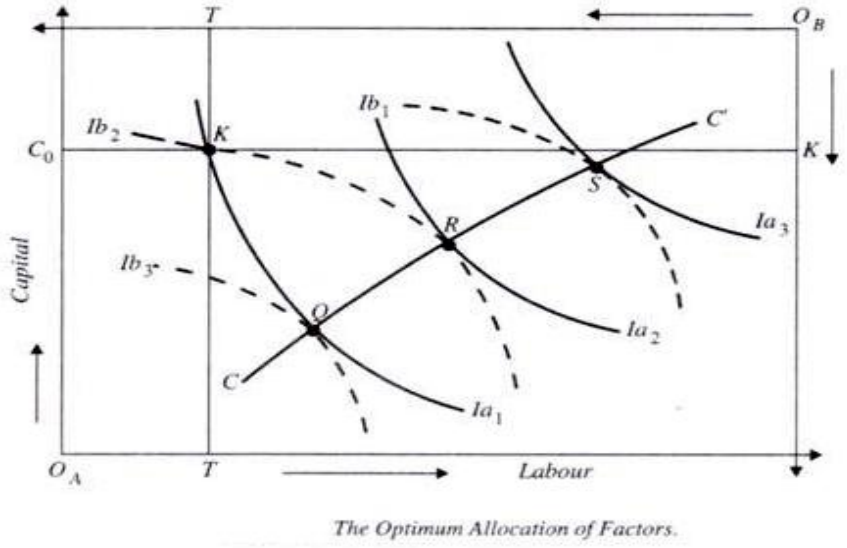


Fig.5.7

Isoquants I_{a1}, I_{a2}, I_{a3} and I_{b1}, I_{b2}, I_{b3} of firms A and B respectively represent successively higher and higher quantities of output which they can produce by different combinations of labour and capital. The slope of the isoquants, which are convex to the origin, represents the marginal rate of technical substitution (MRTS) between two factors.

MRTS of one factor for another is the amount of one factor necessary to compensate for the loss of the marginal unit of another so that the level of output remains the same. So long as the MRTS between two factors for two firms is not equal, total output of a product can be increased by transfer of factors from one firm to another.

In terms of the above diagram any movement from K to S or to Q raises the output of one firm without any decrease in the output of the other. The total output of the two firm's increases when through redistribution of factors between the two firms, a movement is made from the point K to the point Q or S on the contract curve.

A glance at Figure 5.7 will reveal that movement from point K outside the contract curve to the point R on the contract curve will raise the output of both the firms individually as well as collectively. Therefore, it follows that corresponding to a point outside the contract curve there will be some points on the contract curve production at which will ensure greater total output of the two firms.

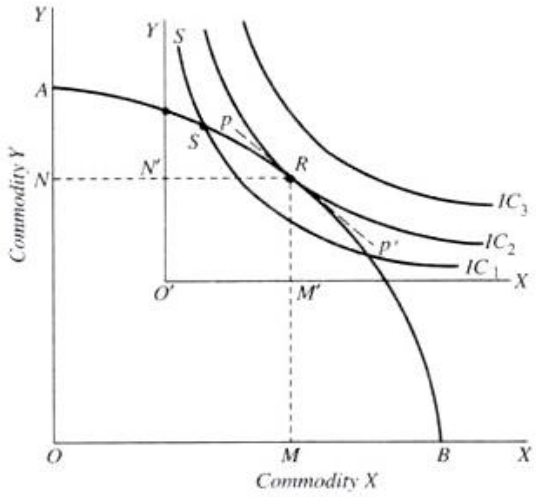
As the contract curve is the locus of the tangency points of the isoquants of two firms, the marginal rate of substitution of the two firms is the same at every point of the contract curve CC. It therefore, follows that on the contract curve at every point of which MRTS between the two factors of two firms is the sum, the allocation of factors between the two firms is optimum.

When the allocation of factors between the two firms is such that they are producing at a point on the contract curve, then no re-allocation of factors will increase the total output of the two firms taken together.

But it is worth mentioning that there are several points on the contract curve and each of them represents the optimum allocation of labour and capital as between the two firms. But which one of them is best cannot be said on the basis of Pareto criterion because movement along the contract curve in either direction represents such factor reallocation which increases the output of one and reduces the output of another firm.

3. Optimum Direction of Production: Efficiency in Product Mix:

This condition relates to the pattern of production. The fulfillment of this condition determines the optimum quantities of different commodities to be produced with the given factor endowments. This condition states that "the marginal rate of substitution between any pair of products for any person consuming both must be the same as the marginal rate of transformation (for the community) between them." According to this condition, for the attainment of maximum social welfare goods should be produced in accordance with consumer's preferences. Let us explain this with the help of Fig.5.8



Optimum Direction of Production :Optimum Product Mix.

Fig.5.8

In Fig.5.8 commodities X and Y have been represented on the X and Y axes respectively. AB is a community's transformation curve between any pair of goods X and Y. This curve represents the maximum amount of X that can be produced for any quantity of Y, given the amounts of other goods that are produced and fixed supplies of available resources.

IC_1 and IC_2 are the indifference curves of a consumer the slope of which at a point represents the marginal rate of substitution between the two goods of the consumer. The MRT' of the community and MRS of the consumer are equal to each other at point R at which the community's transformation curve is tangent to the indifference curve IC_2 of a representative consumers, Point R represents optimum composition of production in which commodities X and Y are being produced and consumed in OM and ON quantities.

This is because of all the points on the community's transformation curve, point R lies at the highest possible indifference curve IC_2 of the consumer. For instance, if a combination of goods X and Y represented by S is being produced and consumed, the consumer would be at a lower level of welfare because S lies on his lower indifference curve IC_1 which intersects the community's transformation curve instead of being tangential to it.

As a result, at point S, MRS_{XY} of the consumer is not equal to the MRT_{XY} of the community. With the situation at S there is a possibility of moving the consumers to a higher indifference curve by changing the direction (i.e. composition) of production i.e. by increasing the production of X and reducing the production of Y. Thus, the optimum direction of production is established at point R where community's transformation curve is tangent to the indifference curve of a consumer in the society.

The Second-Order and Total Conditions of Pareto Optimality:

The marginal or the first order conditions explained above are 'necessary' but not sufficient for the attainment of maximum social welfare because the marginal conditions by themselves do not guarantee maximum welfare.

The marginal conditions can be fulfilled even at the level of minimum welfare. To attain the maximum social welfare position second-order conditions together with the marginal conditions must be satisfied. The second order conditions require that all indifference curves must be convex to the origin and all transformation curves concave to it in the neighbourhood of any portion where marginal conditions are satisfied.

But even the satisfaction of both (first and second order conditions) does not ensure the largest maximum welfare because even when marginal conditions (first and second order) are fulfilled, it may still be possible to move to a position where social welfare is greater. To

attain the maximum social welfare, another set of conditions which are called by J.R. Hicks as the 'total conditions' must also be satisfied.

The total conditions state, "That if welfare is to be a maximum, it must be impossible to increase welfare by producing a product not otherwise produced or by using a factor not otherwise used." If it is possible to increase welfare by such activities the optimum position is not determined by marginal conditions alone.

Therefore, welfare will be really maximum if the marginal as well as total conditions are satisfied. But such a social optimum too is Hot a unique one. It is one of a large number of optima. The whole analysis of conditions of Pareto optimality assumes a given distribution of income.

With a change in the distribution of income Pareto optimality will be achieved with different output-mix of various products and different allocation of various factors among products. Thus, a new optimum will emerge due to redistribution of income and there are no criteria to judge whether the new optimum is better or worse than the previous social optimum. This can be known only with the help of some value judgements regarding income distribution which has been ruled out by the Pareto criterion.

A Critical Evaluation of Pareto Criterion and Pareto Optimality:

Pareto criterion and the concept of Pareto optimality and maximum social welfare based on it occupy a significant place in welfare economics. To judge the efficiency of an economic system, the notion of Pareto optimality has been used.

It has also been used to bring out the gains of trading or exchange of goods between individuals. But even Pareto criterion which rules out comparing those changes in policies which make some worse off has been a subject of controversy and has been criticised on several grounds.

First, it has been alleged that Pareto criterion is not completely free from value judgements. The supporters of Pareto criterion claim that it provides us with an 'objective' criterion of efficiency. However, this has been contested.

Against Pareto criterion it has been said that to say that a policy change which makes some better off without others being worse off increases social welfare is itself a value judgement. This is because we recommend such changes which pass Pareto criterion.

The implication of this assertion will become obvious when the persons who gain as a result of policy change are the rich and those who remain where they were before are poor. Therefore, to say on the basis of Pareto criterion that whenever any policy change which,

without harming anyone, benefits some people regardless of whoever they may be, increases social welfare is a value judgement which may not be accepted by all.

Second, an important limitation of Pareto criterion is that it cannot be applied to judge the social desirability of those policy proposals which benefit some and harm others. Such policy changes are quite rare which do not harm at least some individuals in the society.

Thus, Pareto criterion is of limited applicability as it cannot be used to pronounce judgements on a majority of policy proposals which involve a conflict of preferences of two individuals. Thus, according to Prasanta K. Patnaik, "Pareto criterion fails seriously when it comes to comparing alternatives. Whenever there is conflict of preferences of two individuals with respect to two alternatives, the criterion fails to rank those two alternatives no matter what the preferences of the rest of individuals in the society might be".

To evaluate social desirability of those policy changes which benefit some and harm others, we need to make interpersonal comparison of utility which Pareto criterion refuses to do. Thus, "Pareto criterion works by sidestepping the crucial issue of inter-personal comparison and income distribution, that is, by dealing only with cases where no one is harmed so that the problem does not arise".

Another shortcoming of Pareto criterion and notion of maximum social welfare based on it is that it leaves a considerable amount of indeterminacy in the welfare analysis since every point on the contract curve is Pareto-optimal.

For instance, in Fig. 5.4, every point such as P, Q, R, S on the contract curve is Pareto-superior to any point such as K and H which lies outside the contract curve. Movement from one point on the contract curve to another as a result of change in economic policy, that is, through re-allocation of resources that makes one individual better off and the other worse off, that is, one gains at the expense of the other.

This means that on the basis of Pareto criterion, social alternatives lying on the contract curve cannot be compared since with any movement on the contract curve one individual gains and the other loses, that is, it involves redistribution of income or welfare. Therefore, to compare various alternatives lying on the contract curve and to choose between them, inter-personal comparison and value judgements regarding proper distribution of income need to be made. However, Pareto refused to make value judgements and sought to put forward a value-free or objective criterion of welfare.

It, therefore, follows that on the basis of Pareto criterion where the change from an alternative lying outside the contract curve to an alternative on the contract curve is judged to

increase social welfare but this cannot be said of the change from one position on the contract curve to another on it. But as there are infinite numbers of points on the contract curve all of which are Pareto optimal, no choice can be made out of them on the basis of Pareto criterion.

To remove this indeterminacy and to choose among the alternatives lying on the contract curve one need to make some additional value judgements beyond what is implied in the Pareto criterion. Henderson and Quandt hold a similar view when they assert, "The analysis of welfare in terms of Pareto optimality leaves a considerable amount of indeterminacy in the solution there are infinite number of points which are Pareto optimal." They further remark that.

The indeterminacy is the consequence of considering an increase in welfare to be unambiguously defined only if an improvement in one individual's position is not accompanied by a deterioration of the position of another. The indeterminacy can only be removed by further value judgements."

Above all, a chief drawback of Pareto-optimality analysis is that it accepts the prevailing income distribution and no attempt is made to find an optimal distribution of income, since it is thought that there does not exist any objective, value-free and scientific way of finding optimal distribution of income.

Thus, Pareto optimality analysis remains either silent or biased in favour of status quo on the issue of income distribution. Further, Pareto optimality analysis may lead to recommend the prevailing income distribution where a majority of the population lives on the subsistence level or below the poverty line while a few live a life of affluence. Thus, "Ultimately, the Paretian approach can be considered the welfare economists' instrument par excellence for the circumvention of the issue of income distribution."

It may also be mentioned that for any initial distribution of income (that is, for any given distribution of goods) between the individuals, there will be several Pareto optimal positions. Consider Figure 5.4. Corresponding to point, K, the points on the segment RS on the contract curve CC' will all represent Pareto optimal positions.

Likewise, corresponding to a given distribution of income (i.e. distribution of goods) as represented by point H, the points on the segment PQ of the contract curve CC' will be Pareto-optimal. Thus corresponding to a different distribution of income, there will be different Pareto optima. In the Paretian analysis there is no way of evaluating whether one pattern of income distribution is better than the other.

Prof. Amartya Sen's Critique of Pareto Optimality:

Further, criticising Pareto criterion Prof. Amartya Sen has pointed out that the success that the criterion of Pareto optimality has achieved in judging the desirability of a social state or a policy change is very limited. To quote him, "A Social state is described as Pareto optimal if and only if no- one's utility can be raised without reducing the utility of someone else. This is a very limited kind of success and in itself may or may not guarantee much. A state can be Pareto optimal with some people in extreme misery and others rolling in luxury, so long as the miserable cannot be made better off without cutting into the luxury of the rich." So, according to him, this is not a good and adequate criterion for judging social welfare.

Further, Prof. Sen has criticised Pareto optimality on the basis that it identifies well-being with utility and captures the efficiency aspects only of utility-based accounting. It may be noted that utility is interpreted in two ways, Firstly, it is said to mean 'happiness'. Secondly, it is interpreted in the sense of 'desire-fulfilment'.

He is of the view that utility does not always reflect well-being. To quote him, "To judge the well-being of a person exclusively in metric of happiness or desire- fulfillment has some obvious limitations. These limitations are particularly damaging in the context of interpersonal comparison of well-being. Since the extent of happiness reflects what one would expect and how the social 'deal seems in comparison with that."

He is of the view that people living a life of great misfortune with little hope and opportunities may get more utility or happiness even from small gains. But that should not be interpreted that there is a significant improvement in their well-being.

The measure of utility in the sense of happiness may not reveal the true picture about the state of his deprivation. He thus writes The hopeless beggar, precarious landless labourers, the dominated housewife, the hardened unemployed or the over-exhausted coolie may all take pleasure in small mercies, and manage to suppress intense suffering for the necessity of continued survival, but it would be ethically deeply mistaken to attach correspondingly small value to the loss of their well-being because of their survival strategy.

According to Prof. Sen, even in case of desire-fulfillment, the same problem arises, because "the hopelessly deprived lack the courage to desire much, and their deprivations are muted and deadened in scale of desire-fulfillment." The sum and subsistence of Sen's criticism is that the concept of utility used in for adjudging Pareto optimality whether it is interpreted in terms of happiness or desire-fulfillment is seriously inadequate and insufficient for judging a person's well-being. To quote him, well-being is ultimately a matter of

valuation, and while happiness and fulfillment of desire may well be valuable for the person's well being, they cannot on their own or even together adequately reflect the value of well-being."

It is thus clear that welfare or well-being of individuals depends on a wide range of variables that those associated with utility which is derived from the consumption of goods and services and amount of leisure enjoyed. Welfare or well-being also depends on such variables as political and environmental factors, personal and political freedom individuals enjoy, disposition of their neighbours.

For an adequate measure of well-being these variables cannot be ignored. "In comparing different economic systems or comparing different ways of organising a given economy, the possibility that some of these variables might be affected cannot be ignored. Thus, a reorganisation that gives everyone more income and leisure might not improve the welfare of the community if at the same time it limits individual freedoms or requires the abandonment of cherished cultural traditions."

In the end, it may be pointed out that Pareto criterion is not altogether unless. It is useful in the sense that, "by throwing out the Pareto in optimal alternatives, it reduces the range within which socially best alternatives are to be looked for, and therefore does serve as a useful first step.

The trouble arises if one gets so fascinated with this first step that one does not try to go any further, but that can hardly be called a defect of Pareto criterion." Moreover, as has been pointed out above, Pareto analysis has been used to bring out the gains from trading or exchange of goods between the two individuals.

Perfectly Competitive Equilibrium and Pareto Optimality:

In our above analysis we have explained the various marginal conditions of attaining Pareto optimality or, in other words, optimum allocation of resources. It has been claimed by several economists that perfect competition is an ideal market form which ensures the attainment of Pareto optimality or maximum social welfare as it fulfills all the marginal conditions required for the purpose.

In what follows we shall show how perfectly competitive equilibrium satisfies all the marginal conditions required for the achievement of Pareto optimum. We shall further explain what are the major obstacles in the way of maximising social welfare or achieving Pareto optimality.

Perfect Competition and Optimal Distribution of Goods or Efficiency in Exchange:

The condition for Pareto optimality with regard to the distribution of goods among consumers requires that the marginal rate of substitution (MRS) between any two goods, say X and Y, must be the same for any pair of consumers. Let A and B be the two consumers between whom two goods X and Y are to be distributed.

Under perfect competition prices of all goods are given and same for every consumer. It is also assumed that consumers try to maximise their satisfaction subject to their budget constraint.

Now, given the prices of two goods, consumer A will maximise his satisfaction when he is buying the two goods X and Y in such amounts that:

$MRSA^{A}_{XY} = P_X/P_Y \ldots$ (i)

Likewise, the consumer B will also be in equilibrium (maximise his satisfaction) when he is purchasing and consuming the two goods X and Y in such amounts that:

$MRSA^{B}_{XY} = P_X/P_Y \ldots$ (ii)

Since this is essential condition of perfect competition that prices of goods are the same or uniform for all consumers, the price ratio of the two goods (P_X/P_Y) in equations (i) and (ii) above will be the same for consumers A and B. It, therefore, follows from equations (i) and (ii) above that under conditions of perfect competition marginal rate of substitution between two goods X and Y will be equal for the two consumers. That is,

$MRSA^{A}_{XY} = MRSA^{B}_{XY}$

This result will hold good between any pair of goods for any pair of consumers.

Perfect Competition and Optimal Allocation of Factors:

The second marginal condition for Pareto optimality relates to the optimal allocation of factors in the production of various goods. This condition requires that for the optimal allocation of factors marginal rate of technical substitution (MRTS) between any two factors, say labour and capital, must be the same in the production of any pair of products.

This condition is also satisfied by perfect competition. For a producer working under perfect competition prices of factors he employs are given and constant and he is in equilibrium (that is, minimises his cost for a given level of output) at the combination of factors where the given isoquant is tangent to an iso-cost line.

As is well known, the slope of the isoquant represents marginal rate of technical substitution between the two factors and the slope of the iso-cost line measures the ratio of the prices of two factors. Thus, under perfect competition, a cost-minimising producer

producing goods will equate MRTS between labour and capital with the price ratio of these two factors.

Thus under perfect competition:

$MRSA^{X}_{LK} = P_L/P_K \ldots$ (i)

Where P_L and P_K are the prices of labour and capital respectively and $MRTS^{X}_{LK}$ is the marginal rate of technical substitution between labour and capital in the production of good X. Similarly, producer B producing good T and working under perfect competition will also equate his marginal rate of technical substitution between the two factors with their price ratios. Thus

$MRSA^{Y}_{LK} = P_L/P_K \ldots$ (ii)

Since, under perfect competition, prices of factors are the same for all the producers, each producer will adjust the use of factors in such a way that his marginal rate of technical substitution (MRTS) between labour and capital in the production of goods is equal to the same factor price ratio.

In other words, (P_L/P_K) will be the same for all of them and to this $MRTS_{LK}$ of the producers will be made equal.

It, therefore, follows from (i) and (ii) above that under perfect competition:

$MRSA^{X}_{LK} = MRSA^{Y}_{LK}$

We thus see that perfect competition ensures optimal allocation of resources as between different firms using these resources for production of commodities.

Perfect Competition and Optimum Direction (i.e. Composition) of Production: Allocative Economic Efficiency:

The most important condition for the attainment of Pareto optimum is one which refers to the optimum direction or composition of production. In other words, this condition requires how much amounts of different goods should be produced and resources allocated accordingly.

This refers to the general condition for optimum allocation of resources which has also been called the condition for General Economic Efficiency and General Pareto Optimum. This condition states that marginal rate of substitution between any two commodities for any consumer should be the same as the marginal rate of transformation for the community between these two commodities.

Under conditions of perfect competition, each firm to be in equilibrium produces so much output of a commodity that its marginal cost is equal to the price of the commodity.

Thus, for firms in perfect competition, $MC_X = P_X$, $MC_Y = P_Y$, where MC_Z and MC_Y are marginal costs of production of commodities X and Y respectively and P_X and P_Y are prices of commodities X and Y. Therefore, it follows that firms working in perfect competition will be in equilibrium when they are producing commodities in such quantities that

$MC_X / MC_Y = P_X/P_Y$

The ratio of marginal costs of two commodities represents the marginal rate of transformation between them.

Therefore, for firms producing under perfect competition:

$MRT_{XY} = MC_X / MC_Y = P_X/P_Y$

When there prevails perfect competition on the buying side, each consumer maximises his satisfaction and is in equilibrium at the point where the given budget line is tangent to his indifference curve.

In other words, each consumer is in equilibrium when:

$MRS_{XY} = P_X/P_Y$

Since, under perfect competition, the ratio of prices of two commodities (P_X/P_Y) consumers and producers it follows from (i) and (ii) above that

$MRS_{XY} = MRT_{XY}$

Likewise, this will hold good for any other pair of commodities. Thus, perfect competition satisfies the marginal condition required for the Pareto optimal composition or direction of production. We thus see that all first order marginal conditions required for the attainment of Pareto-optimality or maximum social welfare are fulfilled under perfect competition. It is in this sense that perfect competition represents economic optimum from the viewpoint of social welfare.

Kaldor-Hicks Welfare Criterion: Compensation Principle:

Economists like Kaldor, Hicks and Scitovsky have made efforts to evaluate the changes in social welfare resulting from any economic reorganisation which harms somebody and benefits the others. These economists have sought to remove indeterminacy in the analysis of Pareto optimality.

They have put forward a criterion known as the ‘compensation principle’ on the basis of which they claim to evaluate those changes in economic policy or organisation which makes some individual better off and others worse off. The ‘compensation principle’ is based on the following assumptions

Assumptions:

1. The satisfaction of an individual is independent of the others and he is the best judge of his welfare.
2. There exist no externalities of consumption and production.
3. The tastes of the individuals remain constant.
4. The problems of production and exchange can be separated from the problems of distribution. Compensation principle accepts the level of social welfare to be a function of the level of production. Thus it ignores the effects of a change in distribution on social welfare.
5. Utility can be measured ordinally and interpersonal comparisons of utilities are not possible.

Given the above assumptions, a criterion of compensation principle can be discussed. Kaldor, Hicks and Scitovsky have claimed to formulate a value-free objective criterion of measuring the changes in social welfare with the help of the concept of 'compensating payments'.

Nicholas Kaldor was the first economist to give a welfare criterion based on compensating payments. Kaldor's criterion helps us to measure the welfare implications of a movement in either direction on the contract curve in terms of Edgeworth box diagram.

According to Kaldor's welfare criterion, if a certain change in economic organisation or policy makes some people better off and others worse off, then a change will increase social welfare if those who gain from the change could compensate the losers and still be better off than before. In the words of Prof. Baumol, "Kaldor's criterion states that a change is an improvement if those who gain evaluate their gains at a higher figure than the value which the losers set upon their losses."

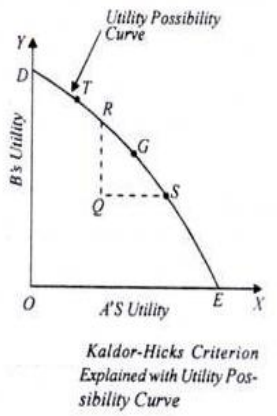
Thus, if any policy change benefits any one section of the society (gainers) to such an extent that it is better off even after the payment of compensation to the other sections of the society (losers) out of the benefits received, then that change leads to increase in social welfare. In Kaldor's own words, "In all cases…. where a certain policy leads to an increase in physical productivity and thus of aggregate real income… it is possible to make everybody better off without making anybody worse off. It is quite sufficient…. to show that even if all those who suffer as a result are fully compensated for their loss, the rest of the community will still be better off than before."

Prof. J.R. Hicks supported Kaldor for employing compensation principle to evaluate the change in social welfare resulting from any economic reorganisation that benefits some

people and harms the others. This criterion states that, "If A is made so much better by the change that he could compensate B for his loss and still have something left over, and then the reorganisation is unequivocal improvement."

In other words, a change is an improvement if the losers in the changed situation cannot profitably bribe the gainers not to change from the original situation. Hicks have given his criterion from the losers' point of view, while Kaldor had formulated his criterion from gainers' point of view. Thus the two criteria are really the same though they are clothed in different words. That is why they are generally called by a single name 'Kaldor-Hicks criterion'.

Kaldor-Hicks criterion can be explained with the help of the utility possibility curve. In Fig. 5.9, ordinal utility of two individuals A and B is shown on X and Y axis respectively. DE is the utility possibility curve which represents the various combinations of utilities obtained by individuals A and B. As we move downward on the curve DE, utility of A increases while that of B falls. On the other hand, if we move up on the utility curve ED, utility of B increases while that of A falls.



Kaldor-Hicks Criterion Explained with Utility Possibility Curve

Fig.5.9

Suppose the utilities obtained by A and B from the distribution of income or output between them is represented by point Q inside the utility possibility curve DE. Let us assume that as a result of some change in economic policy, the two individuals move from point Q to point T on the utility possibility curve DE.

As a result of this movement, utility of individual B has increased while the utility of A has declined, that is, B has been become better off and A has become worse off than before. Therefore, this movement from point Q to point T cannot be evaluated by means of Pareto criterion. Of course, points such as R, G, S or any other point on the segment RS of utility-possibility curve DE are socially preferable to point Q on the basis of Pareto criterion.

However, the compensation principle propounded by Kaldor-Hicks enables us to say whether or not social welfare has increased as a result of movement from Q to T. According to Kaldor-Hicks criterion, we have to see whether the individual A who gains with the movement from position Q to position T could compensate the individual A who is loser and still be better off than before.

Now, it will be seen from Figure 5.9 that utility possibility curve DE passes through points R, G and S. This means that by mere redistribution of income between the two individuals, that is, if individual B gives some compensation to individual A for the loss suffered, they can move from position T to the position R.

It is evident from the figure that at position R individual A is as well off as at the position Q but individual B is still better off as compared to the position Q. It means due to a policy change and consequent movement from position Q to position T, the gainer (individual B) could compensate the loser (individual A) and is still better off than at Q.

Therefore, according to Kaldor-Hicks criterion, social welfare increases with the movement from position Q to position T, because from T they could move to the position R through mere redistribution of income (i.e. compensation).

It is noteworthy that, according to Kaldor-Hicks criterion, compensation may not be actually paid to judge whether or not social welfare has increased. It is enough to know whether the gainer could compensate the loser for the loss of welfare and still be better off.

Whether redistribution of income (that is, payment of compensation) should be actually made following the change in policy is left for the Government to decide. If it is possible for the gainer to compensate the loser and still be better off, the economists can say that social welfare has increased.

It may be noted that gainer can compensate the losers and still be better off only when the change in economic policy leads to the increase in output or real income. That is why Kaldor and Hicks claim that they have been able to distinguish between changes in output from change in distribution.

When their criterion is satisfied by a change in the situation, it means that the economy has moved to a potentially more efficient position and as a result social welfare can be said to have increased. Now, whether redistribution of income is actually made through payment of compensation by the gainers to the losers, according to them, is a different matter.

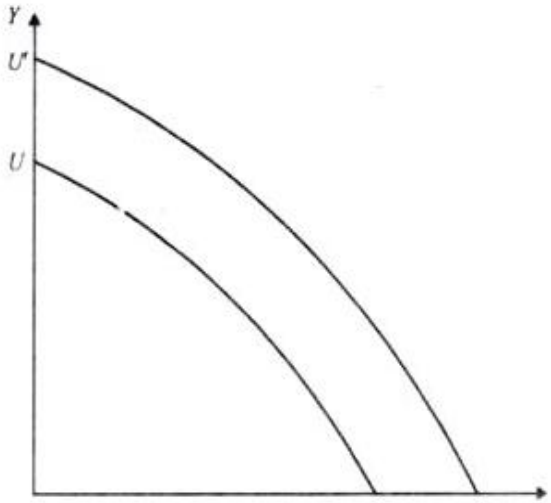
Now, the implications of Kaldor-Hicks criterion become more clear if through redistribution the position of the two individual changes from T to G (see Fig. 5.9). It is quite

manifest that at position G both the individuals A and B are better off than at the position Q. Thus, the position T to which the two individuals moved as result of a certain change in economic policy is superior to the initial position Q from the viewpoint of social welfare, since from position T movement can be made merely through redistribution of income to position G where both are better off as compared to the position Q.

It may be noted that in the situation depicted in Figure 5.9, the change in economic policy brings about a movement from a position inside the utility possibility curve to a point on it. Now let us see what happens to social welfare if as a result of the adoption of a certain economic policy the utility possibility curve moves outward and the two individuals move from a point on a lower utility possibility curve to a point on a higher utility possibility curve.

It can be shown that, according to Kaldor-Hicks criterion, such a movement causes an improvement in social welfare. Consider Figure 5.10. UV is the original utility possibility curve and Q represents the position at which the two individuals are initially placed. Now, suppose utility possibility curve shifts outward to the new position, U'V, and the two individuals are placed at point R on it.

In movement from Q on the utility possibility curve UV to point R on the utility possibility curve U 'V' the utility of A has increased and that of B has declined. But position R denotes greater social welfare on the basis-of Kaldor's criterion when compared to the position Q on the original utility possibility curve UV because with UV as the utility possibility curve it is possible to move through mere redistribution of income from position R to position S where the individual B has been fully compensated for his loss of utility, the individual A is still better off as compared to position Q. To conclude, any change in the economy that moves the individuals from a position on a lower utility possibility curve to a position on a higher utility possibility curve increases social welfare.



Kaldor-Hicks Welfare Criterion

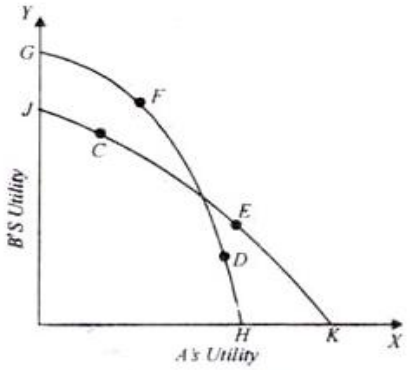
Fig.5.10

Scitovsky Paradox:

Scitovsky pointed out an important limitation of Kaldor-Hicks criterion that it might lead to contradictory results. He showed that, if in some situation, position B is shown to be an improvement over position A on Kaldor-Hicks criterion, it may be possible that position A is also shown to be an improvement over B on the basis of the same criterion.

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 same criterion. According to Scitovsky, Kaldor- Hicks criterion involves such contradictory and inconsistent results. Since Scitovsky was the first to point out this paradoxical result in Kaldor-Hicks criterion, it is known as 'Scitovsky Paradox'.

How Kaldor-Hicks criterion may lead to contradictory results in some situation is depicted in Figure 5.11. In this figure JK and GH are the two utility possibility curves which intersect each other. Now suppose that the initial position is at point C on JK.



Scitvosky Paradox

Fig.5.11

Further suppose that due to a certain policy change, utility possibility curve changes and takes the position GH and the two individuals find themselves at position D. 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 C. Thus movement from position C to position D satisfies Kaldor-Hicks criterion.

But, as has been pointed out by Scitovsky, reverse movement from position D on the new utility possibility curve GH to the position C on the old utility possibility curve JK also represents an improvement on Kaldor- Hicks criterion, that is, C is socially better than D on the basis of Kaldor- Hicks criterion.

This is because from position C movement can be made by mere redistribution of income to position E on the utility possibility curve JK on which position C lies and which also passes through the position E. And, as will be observed from Fig. 5.11, that at position E while A is as well of as at position D, the individual B is still better off than at D.

We thus see that the movement from position C to the position D due to a policy change is passed by the Kaldor-Hicks criterion and also the movement back from position D to position C is also passed by the Kaldor-Hicks criterion. This implies that D is socially better than C on this criterion and C is also socially better than D on the same criterion. So Kaldor-Hicks criterion leads us to contradictory and inconsistent results.

It is mention worthy that these contradictory results are obtained by Kaldor-Hicks criterion when following a policy change new utility possibility curve intersects the former utility possibility curve. After bringing out the possibility of contradictory results in Kaldor-Hicks criterion Scitovsky formulated his own criterion which is generally known as Scitovsky's Double Criterion.

Scitovsky's Double Criterion of Welfare:

To rule out the possibility of contradictory results in Kaldor-Hicks criterion Scitovsky formulated a double criterion which requires the fulfillment of Kaldor-Hicks criterion and also the fulfillment of the reversal test. It means that a change is an improvement if the gainers in the changed situation are able to persuade the losers to accept the change and simultaneously losers are not able to persuade the gainers to remain in the original situation.

Scitovsky's double criterion can also be explained with the help of utility possibility curve. In Figure 41.4, CD and EF are the two utility possibility curves which do not intersect

each other at any point. Suppose there is a change from position Q on the utility possibility curve CD to the position G on the utility possibility curve 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 utility possibility curve EF passing through point R. From the position G, movement can be made to the position R simply by redistributing income between the two individuals. And R is better than Q because the utility of both the individuals 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.

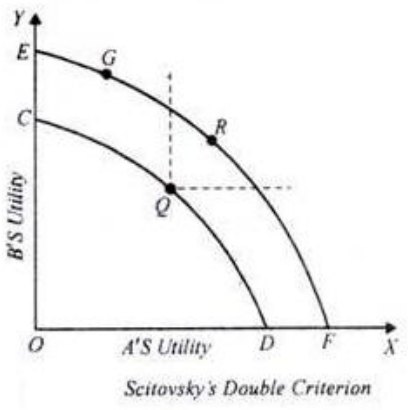


Fig.5.12

Now, let us see, what happens to the reversal test. It must also be satisfied, if the Scitovsky double test is to be fulfilled. That is, a movement from the position G back to the original position Q must not be passed by Kaldor-Hicks criterion if Scitovsky's reversal test is to be satisfied. It is evident from Figure 5.12 that from position R we cannot move to any position on the utility possibility curve CD merely through redistribution of income which is socially better than G (that is, which raises utility of either A or B, the utility of the other remaining constant or which raises the utility of both).

We thus see that while movement from position Q to G is passed by Kaldor-Hicks criterion, reverse movement from position G to position Q is not passed by Kaldor-Hicks criterion. Hence, in Figure 5.12 the movement from the position Q to G satisfies Scitovsky's criterion.

Thus when the two utility possibility curves are non-intersecting and change involves movement from a position on a lower utility possibility curve to a position on a higher utility possibility curve, the change raises social welfare on the basis of Kaldor-Hicks-Scitovsky

criterion. This happens only when a change brings about increase in aggregate output or real income.

A Critique of the Compensation Principle:

The compensation principle as developed by Kaldor, Hicks and Scitovsky, has been a topic of much discussion in welfare economics since 1939. Prof. Kaldor was the first to give a criterion to judge the changes in social welfare when an economic change benefits some people and harms the others.

Later Hicks also supported this criterion in 1940, though he put it in different words. Scitovsky tried to improve the Kaldor-Hicks criterion by formulating his own double criterion. These welfare economists have claimed that they have succeeded in developing a welfare criterion based on ordinal concept of utility and also which is free from any value judgements. But compensation principle has been bitterly criticised by the various welfare economists.

First, 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.

Thus, according to Little, Kaldor has given only a definition of 'increase in wealth' or 'increase in efficiency'. Kaldor himself has interpreted the compensation principle in this sense as he says that, "when the production of wealth goes up, some income distribution could be found which makes some people better off, and no one worse off than before". However, as desired income distribution via compensation is only hypothetical, therefore, according to little, it is not a welfare test but a definition of 'economic efficiency' in terms of over-compensation.

Second, compensation principle is not free value judgements as is claimed by its propounders. It involves implicit value judgements. Prof. Baumol and Little are of the opinion that the contention of Prof. Kaldor that the changes which enable the gainers to compensate the losers and still be better off are good changes is itself a value judgement.

According to little, to say that a policy which meets the Kaldor-Hicks criterion increases the output or "efficiency" of society is, in effect, to recommend it. According to him, Kaldor and Hicks have coined a definition of "efficiency" whose implicit ethical implications or value judgements will hardly find favour with many people.

Compensation is after all only hypothetical; it is consistent with making the poor get poorer. Thus, according to Little, if the value judgements implicit in Kaldor-Hicks criterion

are made explicit, then the claim of Kaldor and Hicks that they have discovered a criterion of detecting increases in wealth, production or efficiency free from value judgements is hardly acceptable.

Third, likewise, Baumol is also of the view that Kaldor-Hicks criterion is based upon unacceptable implicit value judgements. "By using a criterion involving potential money compensations, they set up a concealed interpersonal comparison on a money basis".

If an individual A evaluates his gain from a change worth Rs. 500 whereas another individual B evaluates his loss due to that economic change at Rs. 75, we cannot conclude that social welfare has increased; for if the loser is poor and the gainer a rich one, it may be possible that loss of satisfaction of the poor from Rs. 75 is far greater than the addition to the satisfaction of the rich by Rs. 500 because the marginal significance of one rupee to a poor is far greater than that of the rich.

Thus without actual compensation, the change would mean a major loss of welfare to the poor individual B and a trivial gain of welfare to the rich individual A even if it passes the Kaldor criterion with flying colours. To quote Baumol again, "The Kaldor and Scitovsky criteria have thus ducked the basic problem of the interpersonal comparison required to evaluate a policy change which harms X but aids Y. They duck it by saying implicitly that the recommendation should be based on X's and Y's relative willingness and ability to pay for what they want"?

Fourth, Kaldor-Hicks have claimed that through compensation principle they have been able to separate a production change from the distribution change by which it is accompanied. For instance, as a result of a policy change output of Coca-Cola increases and that of whisky decreases.

Now, if individual X prefers Cola Cola but Y prefers whisky, the question whether there has occurred any increase in. production is inseparably connected with the distribution of these beverages between X and Y. In many cases it is, therefore, difficult to say whether or not production has increased without considering how the output or real income is being distributed.

Moreover, Kaldor and Hicks think that the level of production is the main determinant of social welfare and the distribution a secondary one. But this is quite untenable. A lower total output equitably distributed ensures greater social welfare than larger output, inequitably distributed. They essentially accept the existing distribution of income and wealth and ignore its impact on individual utilities and well-being.

Fifth, Prof. Baumol, Little and Arrow point out another major flaw in compensation principle that it does not envisage social welfare. This principle proves the social desirability of change in the social state on the basis of the criterion that gainers could compensate the losers and still be better off than before.

These critics are of the opinion that policy changes which would increase social welfare when accompanied by actual compensation need not lead to improvement in social welfare if compensation is not actually made.

Dr. Rothenberg has given a very good example to illustrate this. He supposes an initial social state in which a firm adopts a new invention and as a result the cost of production of the firm is reduced but it throws the competitors out of industry and the workers become unemployed.

Let us suppose that the gainer firm from the invention can compensate the losers out of its increased income and still be better off. If the compensation is not actually made in the changed situation, social welfare will decrease as the welfare loss suffered by the workers rendered unemployed will be very large indeed.

As a matter of fact, there is no guarantee that compensation will be actually made in such cases. Thus, so long as compensation is hypothetical a change might make the rich richer and the poor get poorer and therefore reduce social welfare.

It follows from above that a basic flaw in Kaldor-Hicks compensation principle is that it refers to potential welfare rather than actual welfare since it does not envisage that compensation should be actually made.

In the absence of actual compensation one cannot saw whether or not actual social welfare has increased as a result of a certain policy change unless one is prepared to make some value judgements. Therefore, making value judgements, especially that concerning distribution of income or welfare, is quite indispensable in welfare economics. And economists should not fight shy of making those value judgements which are widely accepted by the people.

It may also be noted that if compensation is actually made then Kaldor-Hicks criterion is quite unnecessary, for in that case only Pareto criterion will be sufficient to judge the effect of a policy change on social welfare.

Sixth, compensation principle does not take into account the external effects on consumption and production. The exponents of compensation principle are of the opinion that

an individual's welfare depends solely upon his own level of production and consumption and is not affected by the production and consumption activities of the others.

But this is not a realistic assumption because a person's level of satisfaction (or dissatisfaction) depends to a large extent upon the consumption of goods and services by other persons. A person is more satisfied as his relative economic position in the society is improved.

Thus, if an economic change leaves a person where he was before but makes some other individuals better off, he will not feel as well off as in the original situation, that is, his level of welfare will fall. The gains by some individuals from a policy change have usually unfavourable external effects on the welfare position of those whose position is said to have remained unchanged.

Bergson Social Welfare Criterion:

Bergson pointed out that for comparing utility levels of different individuals resulting from changes in economic policy, making interpersonal comparison of utility cannot be avoided. But he stressed that these comparison of interpersonal utilities should be made explicit.

Thus Bergson put forward a social welfare criterion in which he incorporates explicitly such comparison as utility of individual A from Re 1= utility of individual B from Re 0.20. On the basis of such explicit value judgements or interpersonal comparison of utilities he gives ranking to alternative social states from the viewpoint of social welfare.

Bergson along with Samuelson developed the concept of social welfare function which incorporates explicit value judgements for evaluating the welfare implications of policy changes and also finding out a unique social optimum.

'Arrow's Impossibility Theorem'

The Arrow's impossibility theorem is a social-choice paradox illustrating the impossibility of having an ideal voting structure that is reflective of specific fairness criteria, such as Pareto efficiency. Arrow's impossibility theorem states that a clear order of preferences cannot be determined while adhering to mandatory principles of fair voting procedures.

Breaking down 'Arrow's Impossibility Theorem'

For example, the following shows the type of problem typical of an election. Consider the following example, where voters are asked to rank their preference of candidates A, B and C:

45 votes A > B > C (45 people prefer A over B and prefer B over C)
40 votes B > C > A (40 people prefer B over C and prefer C over A)
30 votes C > A > B (30 people prefer C over A and prefer A over B)

Candidate A has the most votes, so he/she would be the winner. However, if B was not running, C would be the winner, as more people prefer C over A. (A would have 45 votes and C would have 70). This result is a demonstration of Arrow's theorem.

Cumulative voting is the procedure of voting for a company's directors; each shareholder is entitled one vote per share multiplied by the number of directors to be elected. This is sometimes known as proportional voting. This is advantageous for individual investors, because they can apply all of their votes toward one candidate.

Breaking down 'Cumulative Voting'

Cumulative voting is a voting system used by organizations that allow shareholders to vote proportionately to the number of shares they hold. This allows a shareholder with 100 shares to cast the equivalent of 100 votes towards any single issue. In cases where multiple candidates are being considered for multiple positions, such as board seats, each shareholder has the option of placing all of his votes toward one seat during elections, or towards one choice when voting on others matters, but he can also choose to split his votes across multiple options.

For example, if the aforementioned shareholder is participating in a vote for two open board seats, with Candidates A and B running for the first seat and Candidates C and D running for the second seat, he would possess a total of 200 votes. He could choose to participate only in the first seat vote, sending all 200 votes toward the candidate of his choice, Candidate A. He could also vote solely on the second seat, placing all 200 votes on Candidate C. If he wants to vote in both, he can split his votes equally, giving 100 to Candidate A and 100 to Candidate C, or he can direct his votes in an alternate proportion, such as 150 votes for Candidate A and 50 votes to Candidate C.

Macro theories of Distribution

The Ricardian or Classical Theory

The Ricardian theory makes use of two principles in income distribution viz. the ‘marginal principle’ and ‘surplus principle’. With the marginal principle, the theory explains the determination of the share of rent in the aggregate national output and with the surplus principle the shares of wages and profits in the remaining national income. As Kaldor

observes, rent is the difference between the product of labour on the marginal land and the product of average land.

In the Ricardian Theory, the marginal product is assumed to be equal to the sum of wages and profits. The wage rate is said to be determined by the minimum subsistence level. The balance of the marginal product (i.e. after deducting the subsistence wage) is the share of profit. Thus, profits are a residual income (i.e. what is left over after paying rent and wages).

With economic growth the share of wages increases at the expense of profits. Thus, in the Ricardian Macro economic model, there is continuous tendency towards a declining rate of profit with growth in output and employment.

The Marxian Theory

As Prof. Kaldor observes, the Marxian theory is an adaptation of Ricardo's surplus theory. According to Marx, the value of a commodity is determined by the labour-time necessary for its production. But labour produces more than the value of its labour power more than what is necessary for maintaining the minimum subsistence standard. Hence, a surplus emerges which is expropriated by the capitalists in the form of profits. This is Marx's Theory of Surplus Value. Thus, according to Marx profits represent exploitation of labour by the capitalists. The capitalists are able to exploit labour through the ownership of non-human means of production (i.e machines)

The Marxian Theory of Profit Karl Marx referred his surplus value as nothing but the profits earned by the capitalists. According to him, Use Value – Exchange Value = Surplus Value of Profit.

U – E=S

Use Value is nothing but the total value (utility created of produced)

Exchange Value is the amount of money paid to workers as wages or (subsistence) and the surplus value is the profit pocketed by the capitalists. Marx said the capitalists exploits the labour by 100% in other words, if the total value created is Rs.100 then Rs.50 is paid as exchange value and Rs.50 The capitalists profit.

100 -50=50

Marx took the surplus value (Profit) as the indicator of the capitalists progress and decline. Out of the main factor responsible for the collapse of capitalism is due to the falling rate of profit and this happens as the capitalists increases the fixed capital to replace the labour force

The falling rate of profit as one of the indicator of the collapse of capitalism. The notion of input value is crucial to the Marxian theory. It is the amount of this surplus that determines the relative share of profits in the national income. Wages remaining constant at the subsistence level, the difference between total output and the subsistence output (i.e, the rate of exploitation) increases with technical progress. Hence under the capitalistic system the relative share of wages in the national income declines and the share of profits increases (quite the opposite of what Ricardo thought), though the rate of profit will go down as a result of capital accumulation. This is what calls immoderation of the pro-letarist.

Total 'C' Fixed capital	Variable capital V = wages	Surplus value profit = S	$\text{Profit} = \frac{S}{C+V}$
50	50	50	$\frac{50}{100} = 50\%$
100	50	50	$\frac{50}{150} = 33.33\%$
150	50	50	$\frac{50}{200} = 25\%$
200	50	50	$\frac{50}{250} = 20\%$

(Here the assumption of constant V and S is taken into account)

But Marx has proved to be a bad prophet of future events. Worker's living conditions have continued to improve and the rate of profit has not fallen. Labour theory of values on which he based his theory has been rejected.

Kalecki's Degree of Monopoly Theory

According to Kalecki, the distribution of national income into profits and wages depends upon the degree of monopoly in the economy. The degree of monopoly of a firm is measured by (p-a) which is the difference between the price of the product and the average cost on national labour and the materials per unit of output. This difference is made up of entrepreneurial profits, interest, depreciation and salaries and thus represents gross capitalist

income (inclusive of salaries) per unit of output. We can get his total income by multiplying it by the total number of units produced. Also, in order to get the gross capitalist income of the economy as a whole, we have to sum up the gross capitalist incomes of all firms which may be represented by the formula Ex (p-a). If we divide it by 1 (aggregate turn over)

We get

$$\text{Macro - degree} = \frac{\text{Gross capitalist income}}{\text{Aggregate turn over}}$$

of monopoly

Kelecki shows the dependence of labour's share in national income on the macro-degree of monopoly power. The relative share of wages in national income is given by the formula Wages National Income. This share is inversely related to the degree of monopoly power. In other words, increase in the degree of monopoly power will reduce the relative share of wages (i.e. manual labour's share)

Labour's share in national income has remained constant by and large. This, according to Kalecki is due to the fact that increasing degree of monopoly power has been counter-balanced by a fall in the price of basic raw materials

Kaldor's or Keynesian Theory of Income Distribution

Kaldor has called his Income Distribution theory as Keynesian Theory since he employs Keynesian theoretical framework. In its elucidation, Kaldor also divides the national Income into two parts viz. wages and profits. Since profits are defined as the income of property owners, it includes wage and interest besides ordinary profits.

In Kaldor's model, the share of profits in the national income is a function of investment-income ratio (I/Y) the greater this ratio, the greater the share of profits and vice versa. According to Kaldor's theory, a shift in the distribution of income in favour of capitalist class is essential in full employment equilibrium is to be maintained.

Modern Concept of Profit

1. **Economists Concept:** An Economists takes into account the gross profit as the

TR = OP Price Per unit x OQ quantity sold

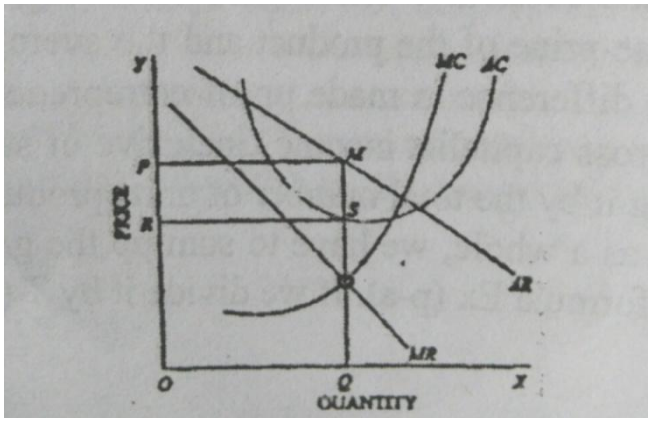
TC = OR Cost per unit x OQ quantity sold

Where TR = Total revenue and

TC = Total Cost

TR – TC =

□ OPMQ - □ ORSQ□ PMSR



(An example of a monopolistic firm or monopolist)

AR = Average Revenue

MR = Marginal revenue

AC = Average cost

MC = Marginal cost

2. A Corporate sector concept

Total revenue – Total cost = Gross Profit

Total Revenue is nothing but price or quantity x, whereas total cost is nothing but the expenses incurred for raw material wages, manufacturing cost, advertisement, wages, salaries and so on

Gross Profit

$$\text{Gross Profit} \quad \left\{ \begin{array}{l} \text{(a) Reserved Fund +} \\ \text{(b) Depreciation +} \\ \text{(c) Corporate taxes} \end{array} \right\} = \text{Net Profit}$$

Summary

The Ricardian theory makes use of two principles in income distribution viz., the 'marginal principle' and 'surplus principle'. According to Marx, the value of a commodity is determined by the labour-time necessary for its production. But labour produces more than the value of its labour power more than what is necessary for maintaining the minimum subsistence standard. According to Kalecki, the distribution of national income into profits and wages depends upon the degree of monopoly in the economy. According to Neo-classical

theory, the marginal principle can be applied to all factors of production by taking them as variable factors, and their rewards are equal to their marginal products.

Prepared by
Dr. K.S. MONY
Associate Professor of Economics
Ambai Arts College, Ambasamudram.