M.A. (ECONOMICS) PART-I **PAPER-I**

SEMESTER–I **ADVANCED ECONOMIC THEORY-I**

UNIT NO. 2

Department of Distance Education
Punjabi University, Patiala

SECTION- B

LESSON NO.

2.1 : Theory of Production–I

2.2 : Theory of Production–II

2.3 : Theories of Costs

2.4 : Price-output Decisions Under Perfect Competition

2.5 : Price-Output Decisions Under Monopoly and Price Discriminating Monopoly

LESSON NO. 2.1 **AUTHOR : DR. BALBIR SINGH**

THEORY OF PRODUCTION-I

2.1 Introduction
2.2 Objectives
2.3 The concept of Production Function
2.3.1 A Production Function
2.3.2 B Laws of Production
2..4 Summary

2.1 Introduction

Production is the most important activity that directly affects an economy. Therefore, the decisions about what to produce and how to produce are vital to any economic system. The cost of production of a commodity is determined mainly by the way in which the resources are combined in a production process, the productivity of resources in various combinations and the prices of the resources involved in the production process. The method of production to be adopted may be determined by the quantity of the output to be produced. This method, in turn, influences the distribution of income and consequently affects the choice of goods to be produced. The process of production itself is capable of influencing consumer's perferences and market. Therefore, the theory of production is one of the most important topics to be studied in economics.

2.2 Objectives

In this lesson we shall discuss (A) the concept of production function and (B) the laws of production,

2.3. The concept of Production Function

Production function shows the functional relationship between physical inputs and physical output. It defines the maximum quality of physical output available from a given set of inputs or the minimum amount of inputs necessary to produce a given amount of output. According to Koutsoyiannis, "The production function is a purely technical relation which connects factors of inputs and output." As Stigler puts it, "The production function is the name given to the relationship between the rates of inputs of productive services and the rate of output of product." Whereas Koutsoyiannis defines production function as a technical relationship between inputs and output. Brown says that the production function also specifies the relation between the inputs themselves. Ferguson views production as "a schedule… showing the maximum amount of output that can be produced from any specified set of inputs, given the existing technology of state of art." Output is a dependent variable in a production function and is measured as a flow of goods over a specific period of time. Inputs are considered as independent variables.

2.3.1 A Production Function

The general mathematical form of production function is :

$Y = f (L, K, R, S, Re, E_f)$

Where

Y = Output

L = Labour input

K = Capital input

R = Raw Material

S = Land input
Re = return to scale
E_f = Efficiency Parameter

For a simplicity, only the inputs of labour and capital are considered as independent variables in a production function. Both labour and capital enter the production function explicitly. Normally land does not explicitly enter the production function. It is because of the implicit assumption that land does not impose any restriction on production. A simple form of a production function is.

Y = f (K, L)

Where Y is the output per unit of time; K and L the quantities of the services of capital and labour per unit of time and it is the functional relation between inputs and the output.

Assumptions

There are implicit assumptions in the concept of production function. First, the inputs and the output are non-negative. Second, the key independent variables are assumed to be adequate to explain the changes in the output and so only these key variables are included in the production function. Third, an implicit assumption underlying a production function is that the technology does not change. Fourth, the external effects are not considered. That is, the production of a firm is a relation between its output and the inputs it uses. So, it is assumed that the output of the other firms or the inputs used by these firms do not matter.

Types

There are two types of production functions. The fixed proportion production function with either single or multiple processes, and the variable proportion production function:

a) FIXED PROPORTIONS PRODUCTION PROCESS

(i) SINGLE PROCESS

A Fixed proportion production function is one in which the technology requires a unique combination of inputs to produce each level of output. In this process the possibility of substitution between inputs is ruled out. It means that to produce a given level of output efficiently, there exists only one way in which inputs may be combined to produce efficiently a given level of output.

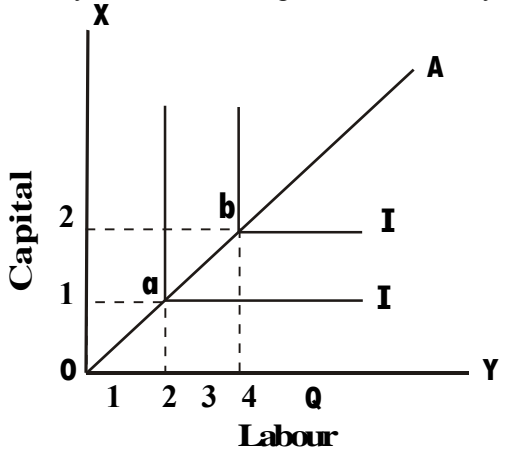


Fig. No. 1

Figure 1 shows the fixed proportion production process. We assume that there are only two inputs, Capital (K) and Labour (L) and the services of these two inputs are combined in the specific proportion 1 : 2 respectively, to produce one unit of the output. Isoquants are L shaped. Let us assume that at a, the output is one unit. Output of one unit can be produced (at 'a') using 1 unit of capital and 2 units of labour. However, if we increase the amount of labour input and keep fixed the quantity of capital at 1, the output will not increase. Similarly, if we increase the amount of capital input and keep the labour input fixed at 2 units the output would not increase. So any increase in an input without the proportionate increase in the quantity of the accompanying input will not increase the output, we will have to double quantity of both the inputs subject to the condition that constant returns to scale is in operation. In the figure, if we want to double the quantity of the output (i.e. a to b) we have to increase both the inputs in a fixed proportion. This has been done by increasing the amount of capital to 2 units and that of labour to 4 units.

(ii) Multiple Process

A variant of the above discussed function in a fixed proportion production function with multiple number of processes. Each process is defined by a line from the origin along which the ratio of inputs is fixed. Within one productive process no factor substitution is possible. But in different processes there are various factors in different quantities, since they involve different fixed factors ratios. Such a product of commodity for which four fixed proportions processes are available is shown in figure No. 2. In this figure four fixed capital labour ratios have been drawn and all yield 100 units of commodity. OA, OB, OC and OD are the process rays whose slopes represent different capital-labour ratio. We get a linked line ABCD by joining point A, B, C and D. Each of these points on

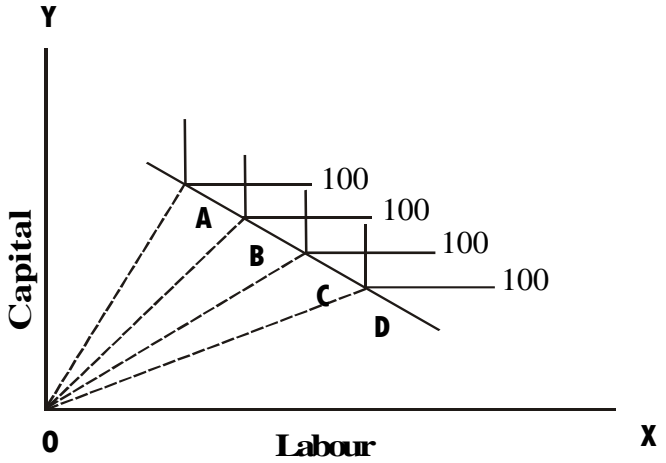


Fig. No. 2

the line represents a factor combination which can produce 100 units of the commodity. But one thing to note is that all points on this linked isoquant line are not feasible factor combinations and can not directly produce 100 units of output. The factor combinations lying between A and B, B and C, and C and D on the linked line are not feasible factor combinations, for it has been assumed by us that only four factor combinations A, B, C and D corresponding to four availble processes are feasible factor combinations capable

of directly producing 100 units of output.

(b) Variable Proportions Production Function

This is the standard neo-classical production function. The function is continuous because it is assumed that the production processes are infinite, which means inputs can be combined in any number of ways. In a two input model it is possible to substitute continuously one input for another, while keeping the level of output unchanged.

Figure No. 3 shows an isoquant map. An isoquant or equal product curve is obtained from an infinite number of combinations of the services of labour and capital. An isoquant map consists of infinite number isoquants. This is because the response of output to infinitesimal changes in inputs is assumed to be continuous. The isoquants do not intersect each other because of the assumption that the production is single valued, i.e. at any given combination of inputs there is unique level of output.

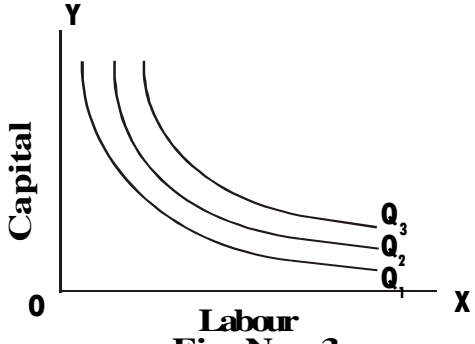


Fig. No. 3

A variable production function may be written as :

$Q = f(K, L)$

The isoquant, which is the locus of efficient points of inputs combinations to produce a given level of output, is continuous and smooth. An isoquant is assumed to be convex to the origin. The marginal product of labour and also of capital is positive. The marginal rate of technical substitution diminishes as one input is substituted for another.

2.3.1 B Laws of Production

The laws of production refer to the various technically possible ways of increasing the level of output. The level of output can be changed by changing the factor input combinations. We can increase output by changing all the factors of production and this is possible in the long run only. When we change output by changing the amount of all the factors, it is studied under the law known as law of returns to scale. On the other hand, we can increase output by using more of the variable

Self-Check Exercise-I

Q.1 What do you mean by Production Function?

Ans. .....

.....

.....

factor(s) while other factors are kept constant. The increase in output with at least one factor constant is studied under the law known as the law of variable proportions. Thus we have two laws of production i.e. the law of returns to scale and the law of variable proportions. We propose to discuss these two laws in detail one by one and we take up first the law of variable proportions.

1. The law of Variable Proportions

In the short period, one or more than one factor input remains fixed, while one or more than one factor input can be varied with respect to the fixed factor(s). The law of proportions explains the behaviour of output as the quantity of one factor is increased, keeping the quantity of other fixed. This law states that as equal increments of input are added, the inputs of other productive services being held constant, beyond a certain point, the resulting increment of product will decrease. It means that the marginal and average output would eventually decline.

This law is based mainly on three assumptions. First, that the state of technology remains constant. Secondly, that of all the inputs used, the quantity of some (at least one) of the factor input must be kept constant. Thirdly, there is possibility of varying proportion in which the various factor inputs are combined in production.

The following table would be helpful to explain the law :

TABLE NO. 1

No. Of workers	Total Product (in Quintals)	Marginal Product (in Quintals)	Average Product (in Quintals)	
1	2	3	4	
1.	5	5	5	
2.	12	7	6	1st Stage
3.	24	12	8	
4.	32	8	8	
5.	37	5	7.40	
6.	39	2	6.49	2nd Stage
7.	39	0	5.57	
8.	37	-2	4.62	3rd Stage
9.	32	-5	3.54	

Note :- In this table only the labour input is allowed to vary, while all other factor inputs are assumed to be held constant.

It is clear from the above table that as the number of workers is increased, the marginal product of labour is increasing up to the employment of the 3rd worker and when more workers are employed, the MP of labours starts declining and becomes zero as the 7th worker is employed. If we continue to employ more workers the marginal product of the successive workers would be negative as is shown in the table no.1. The total product increases at in-

creasing rate upto the point where marginal product is maximum and after this the total product would continue to increase but a decreasing rate till the marginal product becomes zero, when the marginal product is zero the total product would be maximum. In the above table when 8th worker is employed the marginal product of the worker is negative and so the total product would also decrease. The law has three stages. The first stage is characterised by increasing average product and in the second stage both average product and the marginal product fall, but are positive. The third stage is reached when the marginal product becomes negative.

The diagrammatic representation of the law would make it more clear.

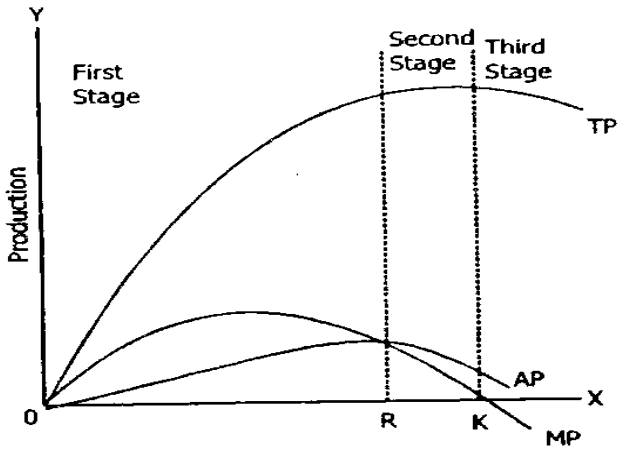


Fig. No. 4

The diagram no.4 makes it clear that in the first stage the marginal product is higher than the average product. The first stage ends where the average product reaches its maximum (point R). But the marginal product reaches its maximum and starts declining even before the stage ends. The total product continues to increase up to the point where second stage ends (point K).

No producer would like to produce in the third stage because the marginal product of the variable factors is negative. A rational producer would not produce in the first stage the marginal product of the variable factor is more than average product, and moreover, average product continues to increase with the increase in the variable factor. Economically, the second stage is the most important region where the average product is greater than marginal product and both AP and MP are still positive. Moreover, the marginal product of the fixed factor in the first stage would be negative. If a producer decides to operate in the first stage, it would mean that he is not utilising fully the opportunity of increasing production by increasing quantity of the variable factors whose average product continues

Self-Check Exercise-II

Q.2 Give the mathematical form of Production Function.

Ans. .....

.....

.....

to increase throughout the first stage. Thus, it would be profitable for the producer to operate in the second stage where the law of diminishing returns is in operation. In the second stage the relavant ratio of the variable factor to the fixed factor, which a firm should use would depend upon the competitive cost of the variable factors and fixed factors.

Now we shall discuss the reasons responsible for the operation of the above mentioned three stages of production. The first, second and third stages are also known as the law of increasing returns, the law of diminishing returns and the law of negative returns, respectively.

The law of increasing returns occurs mainly due to two reasons. Firstly, when we add more units of the variable factors to the constant quantity of the fixed factor then the fixed factor is more effectively utilised and it means the efficiency of the fixed factor increases. Generally, the fixed factor is indivisible and in the beginning, the variable factor is relatively smaller in quantity, some amount of the fixed factor may remain unutilised. When the quantity of the variable factor is increased, full utilization of the fixed factor becomes possible and it results in increasing returns.

Secondly, as the more units of the variable factor are added, the efficiency of the variable factor itself increases. It is because when the quantity of the variable factor becomes greater, the scope for specialisation of division of labour increases.

The law of diminishing returns starts operating only after certain amount of the variable factor has been added to the fixed quantity of the other factor. The per unit output made by the variable factor after a point becomes less and less because the additional units of the variable factor have less and less of the fixed factor to work with. Now the fixed factor is overused and the variable factor is partially used. It happens mainly due to the indivisibility of the fixed factor. Thus, the law of diminishing returns comes into force when we continue to increase the quantitiy of the variable factor even after the optimum combination of the variable factor with the fixed factor has been achieved. Mrs. Joan Robinson says that the diminishing returns occur because the factors of production are imperfect substitutes for one another. Because we can not use one factor in place of the other to infinite level, the law of diminishing returns operate.

The third stage or the law of negative returns operates when the number of the variable factors becomes too excessive relative to the fixed factor that they start getting in each other's way and it results in the fall of the total output. Moreover the excessive number of the variable factor also impairs the efficiency of the fixed factor. It is generally said that "Too many cooks spoiled the broth." So all this leads to the negative marginal product of the variable in the third stage.

2. Returns to Scale

We have explained above the law of variable proportions which presumes that one factor in the combination of factors is fixed and another is variable. Here we will explain another type of production known as the returns to scale. In case of returns to scale all

inputs of factors are considered to be variable. In this type of production function we try to find out the behaviour of output when it is possible to change the size of all the factors. It means we attempt to explain the behaviour of output in response to changes in these scales. Any change in the scale means that all inputs or factors are changed in the same proportion. The following table would be helpful to explain the returns to scale.

TABLE – 2

Units of all the factors used	Total Returns (quintals)	Average Returns (quintals)	Marginal Returns (quintals)	
One	10	10	10	
Two	22	11	12	Increasing
Three	35	11.66	13	
Four	60	15	25	
Five	85	17	25	Constant
Six	102	17	17	
Seven	114	16.28	12	Decreasing

From the table it is clear that when the units of all the factors are increased in the same proportion, the total output increases at the different rates i.e. increasing, constant and decreasing. Now we will discuss in some details the increasing returns to scale, constant returns to scale and decreasing returns to scale.

Increasing Returns to Scale

It means that the output increases at a higher rate than the increase in factors of production employed. We get the increasing returns, generally, in the initial stages of production. Chamberlin is of the view that increasing returns to scale are obtained due to specialisation. As the size of the concern increases division of labour becomes possible and it leads to the increasing returns. The increasing returns to scale are also available to those industries where capacity varies roughly with the cube of the dimensions whereas the material required to construct the capital goods varies roughly with the square dimensions. A 2x2x2 box holds two times much raw materials as 1x1x1 box, but contains in four times much material (capacity).

Constant Returns to Scale

It means the total output increases at the same rate at which all the factors of production are increased. If the amount of all the factors is doubled, total output will also be doubled; and if the factors are tripled, the output will also be tripled. In our table given above the marginal returns are constant when 4 and 5 doses are applied. In mathematics this law of constant returns is known as linear and homogeneous production of the

first degree. There are evidences which show that production function for the economy as a whole is not too far from being homogeneous of the first degree. The view that in the production function for an individual firm, there is a long phase of constant returns to scale, is also supported by empirical evidence.

Decreasing Returns to Scale

Decreasing returns to scale means that proportionate increase in inputs of all factors used in production will give rise to a less than proportionate increase in output. It occurs mainly because of the increasing complexities of large scale management. When the size of a firm gets very large the administrative organisation becomes extensive. Management also becomes unable to adopt itself rapidly to changing condition of demand and cost. With the expansion of the factors after a point diseconomies of scale take place and cause decreasing returns to scale. There is a difference between decreasing returns to scale and diminishing marginal returns, because the former arises from the increasing complexities of the organisation, while the later is caused by the application of increasing units of a variable factor to fixed factors. The decreasing returns to scale, in our table starts operating only after the sixth dose of inputs has been applied.

The following figure (No. 5) shows the increasing, constant and decreasing returns to scale :

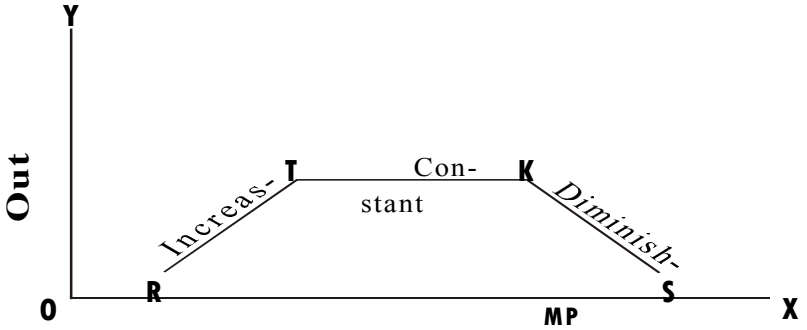


Fig. No. 5

The figure shows that when the units of factors of production are increased, the marginal product increases in the initial stage. As more and more units are put into production process the marginal product becomes constant. If we increase further the units of inputs the output will start declining. In this figure (No. 5) the product curve shows increasing returns from R to T, constant returns from T to K and diminishing returns from K to S.

Application of the Law of Diminishing Returns

The law of variable proportions states that the marginal physical product of the variable factors increases in beginning but starts declining after a certain amount of the variable factor relative to fixed factor has been used. At the time of Marshall it was generally thought that the law of diminishing returns applied to agriculture and manufacturing industries were subjected to increasing or constant returns. But now it is regarded that the law of diminishing returns applies both to agriculture and industry. Marshall is of

the view that this law operates quickly where nature plays an important role as compared to man. He says that where nature plays relatively more important role, the law of increasing returns comes into force. The role played by nature is more in case of agriculture as compared to industry. That is why it is said that this law applies more quickly to agriculture than to industry. It may be due to the following reasons : - First, specialisation and division of labour are relatively easy in industry than in agriculture. Secondly, technological innovations are slow in agriculture in comparison to industry. Thirdly, the role of nature is more in agriculture than in industry. Fourthly, human supervision is relatively difficult in agriculture as compared to industry because agriculture is spread over vast areas.

This law of diminishing returns is based on certain assumptions. First, the techniques of production are assumed to be constant. Secondly, this law starts operating only after a certain stage in production is achieved. Thirdly, the soil to be used for cultivation must not be new.

Production Funtion: The funtional relationship between physical inputs and physical output.

Isoquant: It is the locus of efficient points of inputs combination to produce a given level of output. Its curvature depends on the substitutability of factor input.

Law of variable proportion: In this case, one or more than one factor input remains fixed and with the increase in quantity of any other factor, the behaviour of output is studied.

Returns to scale: It studies the behaviour of output when it is possible to change the size of all the factor inputs.

2.4 Summary: In this lesson, you have read about the concept of production funtion and the laws of production. Production funtions shows the funtional relationship between the input and output. There are two types of production funtions. The fixed probortion production function and the variable proportion production funtion. The laws of production refere to the various technically possible ways to increase the level of output. There are two laws of production: the law of returns to scale and the law of various propostion.

The theory of production, as stated in the beginning, is very important in economics. The cost of production of a product is determined by the laws of production in operation. These costs through the supply of the product influence the price of the product. The laws of production are helpful in taking decision about whether there is a need of providing subsidy or imposing taxes on an industry. But the law of diminishing returns is the most fundamental law. This law was helpful to Malthus in formation of his theory of population. Ricardo also based his theory of rent on this law of diminishing returns. The necessity for improvement in the techniques of production to increase output is explained by this law of diminishing returns.

Long Questions:

Q. Explain the Law of Variable Proportions?

Q. Explain the Law of Returns to Scale?

Short Questions:

Q. Increasing Returns to Scale

Q. Describe Returns to Scale

Q. Constant Returns to Scale

Q. Variable Proportions Production Functions

LESSON NO. 2.2 **AUTHOR : DR. BALBIR SINGH**

THEORY OF PRODUCTION-II

2.2.1 Introduction
2.2.2 Objectives
2.3.3 Cobb-Douglas Production Function
2.2.4 Constant Elasticity of Substitution (CES) Production Function
2.2.5 Technological Progress and the Production Function
2.2.6 Translog Production Function
2.2.7 Elasticity of Technical Substitution.

2.2.1 Introduction

In the modern world different quantitative models have been framed by various scholars to know the optimal relationships between inputs and the output. The specific combinations of inputs and their effect on the quantity/quality of the output are discussed.

2.2.2 Objectives

An analysis in this regard would be taken up in this lesson in the context of two types of production function and also elasticity of substitution.

After having gone through this lesson, you would able to understand:

-Meaning and properties of Cobb-Douglas Production Function (CDPF)

-Meaning and properties of Constant Elasticity of Substitution (CES)

-Meaning and types of Technological Progress and the Production Functions

-The concept of Translog Production Function

-The concept of Elasticity of Technical Substitution between two factors.

2.2.3 Cobb-Douglas Production Function

A number of production functions have been given by economists to measure relations between changes in physical inputs and physical outputs. An important production function is the one associated with the names, C.W. Cobb and P.H. Douglas and is known as Cobb-Douglas production function. Originally, this function was applied not to production process of an individual firm but to the whole of manufacturing industry. A number of studies in 1920's and 1930's by Cobb-Douglas and others were conducted mainly in U.S.A. which showed the Cobb-Douglas production function was found almost correct. In this function, generally two inputs are taken (though in the recent past, it has been worked out by applying more than two inputs). This function is linear and homogeneous. This function can be written as :

$Q = KL\alpha\ C\beta$

Where

Q = Quantity of the output

L = Labour input

C = Capital input

K = Positive constant

α = exponent of labour

β = exponent of capital

Properties of the Function

1. Since the function is having log-linear form, it is simple to handle. In logarithmic form, the function is : $\log Q = K + \alpha \log L + \beta \log C$
2. The function is mostly used in form :

   $Q = K L^{\alpha} C^{1-\alpha}$ $(\alpha < 1)$

In this special case, where $\alpha + \beta = 1$, the function shows constant returns to scale. That is if the inputs of labour and capital are increased by constant g, then the quantity of output will be increased to :

$K (gL)^{\alpha} (gc)^{1-\alpha} = Kg^{\alpha} L^{\alpha} C^{1-\alpha}$

But because $g^{\alpha} g^{1-\alpha} . = g$, therefore

$= g K C^{1-\alpha}$

$= g QL^{\alpha}$

Thus the output has also increased by g because the inputs were increased by g. It is not necessary that $\alpha + \beta = 1$ in Cobb-Douglas production function. If the function is homogeneous of degree 1 there are constant returns to scale. If in the above equation, g cannot be factored out (i.e. taken common) then the production function is not homogeneous, i.e. it is non-homogeneous. A homogeneous function is a function such that if each of the inputs is multiplied by g, then g can be completely factored out. If it is of degree less than one, decreasing returns to scale is said to be existing. Similarly if it is of degree greater than one, there is an increasing returns to scale.

3. The function yields diminishing returns to each input. This can be easily shown, consider input L.

$Q = KL^{\alpha} C^{\beta}$

$$\frac{\partial Q}{\partial L} = \propto KL^{\alpha-1} C^{\beta}$$

Second order partial derivative is.

$$\frac{\partial^2 Q}{\partial L^2} = \alpha (\alpha - 1) KL^{\alpha-2} C^{\alpha} > 0$$

Since $\propto$ is a positive fraction, $(\alpha-1)$ is negative. Thus the rate of change of marginal product of input L is negative, and MP_L declines. Similarly, it can be shown for input C.

4. α and β show the output elasticity coefficient for input L and C. The output elasticity of an input is defined as the ratio of the relative change in output over a relative change in the input. In symbols, the output elasticity of input L is written as :

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

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

$$= \frac{\alpha K L^{\alpha-1} C^{\beta} L}{Q}$$

$$= \frac{\alpha K L^{\alpha-1} C^{\beta} L}{K L^{\alpha} C^{\beta}} = \alpha$$

Similarly, it can be shown that β is the output elasticity of input C.

5. In Cobb -Douglas production function, factor intensity is measured by the ratio $\frac{\alpha}{\beta}$. The higher this ratio, the more labour intensive the technique is. On the other hand, the lower ratio, $\frac{\alpha}{\beta}$ shows that the technique is more capital intensive.

6. α and β show the relative distributive shares of inputs L and C. The relative distributive share of input L is given by

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

Substituting the value of $\frac{\partial Q}{\partial L}$ and Q into this the expression, we get

$$\frac{\frac{\partial Q}{\partial L} L}{Q} = \frac{\alpha K L^{\alpha-1} C^{\beta} L}{K L^{\alpha} C^{\beta}} = \propto$$

It is clear that $\propto$ is the relative distributive share of factor L.

7. The function can also show the relative efficiency of the firms. The coefficient K measures the efficiency in the organisation of factors of production. If the two firms have the same $L^{\alpha} . C^{\beta}$ and still produce different quantities of output, the difference can be due to the superior organisation and entrepreneurship of one of the firms which results in different efficiencies. If firm A has a larger K than that of firm B, then firm A is said to be more efficient than firm B.

Douglas and others made time-series and cross section studies for the manufacturing industry in many countries such as U.S.A., Canada etc. Douglas found almost similar results in these countries. The studies revealed that labour exponent was two-thirds and the capital exponent was one-third. In majority of the cases, the exponent seemed to be unity. It imples from Cobb-Douglas function that labour's share in national product tends to be constant under competitive market conditions.

The function has certain limitations. Firstly, the total output has been considered as a function of labour and capital only. The role of raw materials was not properly recognised in this function. But now, the economists have worked out this production function successfully by including raw materials along with labour and capital in this function. Secondly, all units of labour have been considered homogeneous in this function, actually, they are not homogeneous in the real life. Thirdly, this function mainly perceive constant returns or at the most declining returns in industry, but increasing returns has equal probability of its application in industry. Fourthly, in the inter-industrial samples, we face the problem of non-constant technology while using this function, and this also happens in case of time-series samples. But it can be said that despite its limitations, the function is very beneficial in research work especially because it is relatively simple to handle.

2.2.4 Constant Elasticity of Substitution (CES)

For a long time, the theory of production had focused on the convenient values for the elasticity of substitution $\sigma = 1$ and $\sigma = 0$. In the Cob-Douglas production function with constant returns to scale, the value of the elasticity is one. On the other extreme, in the Leontief-type input-output production function, the value of elasticity is zero. A linear production function implies infinite elasticity of substitution.

Studies made by different writers for empirical evidence about the elasticity of substitution turned out to be inconsistent with any of the three values of the elasticity of substitution. Particularly, in a study made by Arrow, Chenery, Minhas

Self-Check Exercise-I

Q.1 Write the mathematical form of Cobb-Douglas Production Function?

Ans. .....

.....

.....

and Solow, capital-labour proportion were found to vary among countries much more in some sectors than in others. They could conclude that the values of the elasticity of substitution suggested by their study were significantly different from either zero or one. This led these writers to their path breaking formation of CES production function. The substitution is between two factors of production i.e. Labour and capital. The elasticity of substitution is constant and it is not necessarily equal to zero or one. This production function takes the following form.

$$Y = \gamma \left[\delta K^{-\rho} + (1-\delta) L^{-\rho} \right]^{-1/\rho}$$

Where γ is the efficiency parameter, δ (delta) is the distribution parameter taking the value between zero and one, and ρ is the substitution parameter whose value lies between –1 and infinity.

Properties of CES Production Function

Firstly, it can be shown that ρ specifies the elasticity of substitution, since $\sigma = \frac{1}{1+\rho}$.

Although this specification restrict σ to constancy it permits a much wider choice among alternative values. Since $-1 < \rho < \infty$ (infinity), the range for the e_s becomes $\infty > \sigma > 0$.

Secondly, it is a generalised production function. The three examples of production function mentioned are special limiting cases of the CES, as one can see by substituting the appropriate values.

(i) If $\rho = -1$, $\sigma = \infty$ and the production function becomes the linear production function;

(ii) If $\rho = \infty$, $\sigma = 0$ and the production function takes on the Leontief form with fixed factor proportions.

(iii) If $\rho = 0$ then $\sigma = 1$ and we have the Cobb-Douglas production function.

Thirdly, by allowing values of σ other than one, the CES production function admits the possibility that a certain change in the rate of relative factor prices will induce a change in the relative factor utilization ratios which is not necessarily equi-proportional in the opposite direction.

Fourth, the marginal production of capital in the production function are

$$\frac{\partial Y}{\partial R} = \delta \gamma^{-\rho} \left(\frac{Y}{K} \right)^{1+\rho}$$

$$\frac{\partial Y}{\partial R} = (1-\delta) \gamma^{-\rho} \left(\frac{Y}{L} \right)^{1+\rho}$$

Since the parameters δ and γ and the variables Y, K and L are all positive, the marginal product of any factor, i is positive and decreasing with the level of factor i

but increasing with the level of factor j where $j \neq i$.

Fifth, this factor is homogeneous of degree one, the function can be rewritten as

$$Y = \gamma \left[\delta K^{-\rho} + (1-\delta) L^{-\rho} \right]^{-v/\rho}$$

Where special parameter v measures return to scale.

By increasing the inputs K and L m-fold we get,

$$\gamma \left[\delta(mK)^{-\rho} + (1-\delta)(mL)^{-\rho} \right]^{-1/\rho}$$

$$\gamma \left(m^{-\rho} [\delta K^{-\rho} + (1-\delta) L^{-\rho}] \right]^{-1/\rho}$$

$$= (m^{-\rho}) . \gamma \left[\delta K - \rho + (1-\delta) L^{-\rho} \right]^{-1/\rho}$$

$$= m.y$$

The CES production function thus possesses the properties of a linearly homogeneous function.

Sixth, computation of the e_s is not difficult in this production function. The CES production function can be estimated either directly by using maximum likelihood techniques or indirectly by utilizing the relationship between the average productivity of labour and the wage rate, which yields the value of e_s as the coefficient for the wage rate.

2.2.5 Technological Progress and the Production Function

As we know that with the passage of time new and more efficient methods of production are developed, e.g. technological changes occur in the production related activities. Moreover, the coming up of new inventions may result in the increase of the efficiency of all methods of production. At the same time, some techniques may become out-dated and so are to be left away from the production function. These changes in technology are referred to as the technological progress. The effect of innovations in this process has been shown with an upward shift of production function in Fig.1. On the other hand, a downward movement of production isoquant as is shown in Fig. 2. The shift shows that the same level of output may be produced by less factor inputs or more outputs may be with the same inputs.

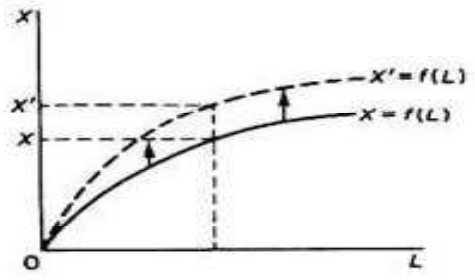


Fig. 1

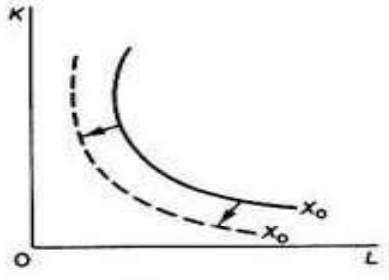


Fig. 2

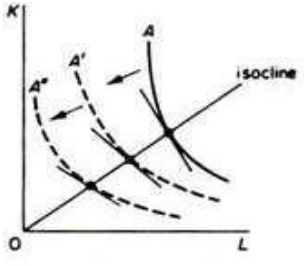
The technical progress may also change the shape as well as place of the isoquant. There are three types of technical progress as suggested by J.R. Hicks, depending upon its effects on the rate of substitutions of the factors of production, as given below:

Capital Intensive Technical Progress :

Technical progress is capital intensive if along a line on which capital labour (K/L) ratio is constant, the MRS_{LK} increases. It means that technical progress increases the marginal product of capital by more than the product of labour. This ratio of marginal products (which is the MRS_{LK}) decreases in absolute value and the slope of the isoquant turns negative, this sort of technical progress increases MRS_{LK}. So, the scope of the shifting isoquant becomes less steep along any given radius. This capital intensive technical progress has been depicted in Fig. 3.1.

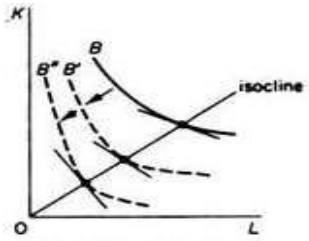
Labour Intensive Technical Progress:

Labour intensive technical progress if, along a radius through the origin (with constant capital labour ratio (i.e. K/L), the MRS_{LK} increases. It means that technical progress increases the marginal product of labour MP_L faster than MP_k. So, the MRS_{LK} can be described as the ratio of marginal products $[(\partial X/\partial L)]/[\partial X/\partial K)]$, increases in absolute value (but decreases if the minus sign is taken into account). The downward shifting isoquant becomes steeper along any given radius through the origin. It is depicted in Fig. 3.2.



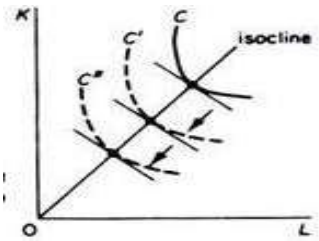
Capital deepening technical progress

Fig. 3.1



Labour deepening technical progress

Fig. 3.2



Neutral technical progress

Fig. 3.3

Neutral-Technical Progress :

If the technical progress is neutral, it increases the marginal product of both factors by the same percentage, i.e. the MRS_{LK} (along any radius) remains constant. The isoquant shifts downwards parallel to itself, has been shown in Fig. 3.3.

Self-Check Exercise-II

Q.2 What is the relationship between technological progress and Production Function?

Ans. .....

.....

.....

2.2.6 Translog Production Function

One of the important production functions that we are discussing in this chapter is the Translog Production Function estimation. A noteworthy feature of this function is that when the number of production factors increases, the number of parameters exploes. This production function considers the problem of occurence of collinearity. This production function has some properties that are given below :

(i) Brief historical aspect of this production function ;

These production functions take place in the context of researches related to the discovery and definition of new flexible forms of production functions. The first form of translog production function was given in 1967 by J. Kmenta for the approximation CES production function with the second order Taylor series when the elasticity of substitution is very close to the unitary value, which is the case of Cobb Douglas Production Function. The form of the above mentioned production function can be written as :

$$\text{In } Y = \text{In } A_3 + \delta_3 . \text{ In } K + \beta_3 . \text{ In } L + \chi_3 . \text{ In}^2 (K/L) \ldots\ldots\ldots (1)$$

Where :

In = natural logarithm

Y = Output (Gross Domestic Product)

K = Fixed capital

L = Employed population

A_3, δ_3, β_3, χ_3 are parameters to be estimated.

In 1971, some new forms of production function were proposed. The first one was found by imposing the condition of $\alpha + \beta = 1$. So, the production function became a labour productivity function.

$$\text{In } (Y/L) = \text{In } A_2 + \delta_2 . \text{ In } (K/L) + \chi_2 . \text{ In}^2 (K/L) \ldots\ldots (II)$$

The above mentioned production function is one of a second order polynomial in the logarithms of the single input taken, capital-labour ratio, respectively. The second form of production function was defined in conditions of relaxing the constraints imposed to the parameters in the Kmenta function, in order to test the homotheticity assumptions, as written below :

$$\text{In } Y = \text{In } A_{KL} + \delta_K . \text{ In } K + \alpha_L . \text{ In } L + \beta_K{}^2 \text{ In } K + \beta_L{}^2 . \text{ In}^2 L + \beta_{KL} . \text{ In } K . \text{In } L \ldots (III)$$

It is important to mention that the term "translog production function", is the abridged form of transcendental logarithmic production function which dealt with the problems of strong separability and homogeneity of Cobb Dougles and CES production functions and their implications for the production frontier. The generalised form of translog production function, which takes into account a number of inputs, is given as:

$$\text{In } Y = \text{In } A_{\delta_i \beta_{ij}} + \sum_{i=1}^{n} \delta_i . \text{In } X_i + \left(\frac{1}{2}\right) . \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} . \text{In} X_i . \text{In} X_j \ldots\ldots (IV)$$

This function explains a class of flexible functional forms for the production

functions. An important benefit of this function is that unlike Cobb-Douglas Production function, it does not assume rigid premises e.g. perfect or smooth substitution between factors of production or perfect competition in the perfect factors market. Moreover the concept of translog production allows to pass from a linear relationship between output and the factors of production, which are taken into account to a non-linear one. Due to its properties the translog production function can be used for the second order approximation of linear homogenous production frontier or for the measurement of the total factor productivity dynamics.

At the second place the main indicators and constraints in the estimation of translog production functions parameters are given below.

The Marginal Product $\left(\frac{\partial Y}{\partial X_i}\right)$ in a translog production function is equal to

$$\frac{\partial Y}{\partial X_i} = a_i + \sum_{j=1}^{n} \beta_{ij}.\text{In}\, X_j$$

It is important to mention that the marginal product of a translog production function is formally akin to a Cobb-Douglas Production Function. Taking the marginal product in consideration it can also determine the marginal rate of transformation between two factors of production i.e. $\left(\frac{\partial X_j}{\partial X_i}\right)$

$$\frac{\partial X_j}{\partial X_i} = \frac{\partial_i + \sum_{i=1}^{n} \beta_{ij}.\text{In}X_i}{\partial_j + \sum_{j=1}^{n} \beta_{ij}.\text{In}X_j}$$

C.E. Ferguson stated that the marginal product is equal to the elasticity of scale.

As explained earlier the Translog Production Function has a number of benefits in the research of the economic activity from the theoretical view point. A large number of that are to be estimated for making operational the concept of "translog production function" there are hard constraints on the result feasibility, as the occurrence of an extended collinearity is there. Actually the number of Parameters explodes as the number of production factors, which are taken into account increase. Suppose if the number of production factors that are under consideration is equal

to n, the number of estimated parameters becomes equal to $\frac{n.(n+3)}{2}$

If in estimation, ordinary least square (OLS) method is used, even if three production factors are considered, the chances of occurrence of the harmful collinearity is very high. For this harmful collinearity a solution is used in order to solve the difficulty generated by a collinearity is the ridge regression, that theoretically permits to take estimations which are not distroted by the high degree of collinearity and especially by the harmful one. However this kind of ridge regression has also a shortcoming, i.e. the ridge (Correction) parameter applied in order to diminish the impact of collinearity parameter is in fact subjectively chosen (J.Klacek, J. Vopravil 2008). So when the ridge regression is used, one is not sure whether the solution obtained is an optimal one. Besides, the deviation of results obtained the ridge regression tends to be greater and greater in the context of as the number of factors of production is higher and higher.

In an other solution to obtain feasible results with translog production is the limitation of the number of production factors to those which can be really considered for the behaviour of output. Because the collinearity is a cummulative phenomenon, the first test for introduction in the translog production function of a specific production factor is to estimate, by the OLS method, of the translog production function which is related only to the analysed factor. In case, that the results obtained in estimation are considered feasible, the respective factor of production may be introduced into an extended translog production function.

To sum up, the estimation of the parameters of translog production with a single production factor allows to enlarge the vision related to the relationship between the output and the analyzed production factor from a linear one to a non-linear one. In this context we can stress on the role of the acceleration of the dynamics of the production factors in order to increase the output level.

In the context of weak and degrading collinearity and accelerated dynamics of the production factor determines a translog multiplier less than 1, but a less accelerated dynamics determines a translog multiplier more than 1. In other words it can be said that the elasticity of scale tends to decrease as compared to the output elasticity as the dynamics of the production factor becomes more and more non-linear.

Even in the situation of single production factor, the problem of collinearity in the translog production function is n controlled. The values of R(InX; In^2X) are quite high, so the incidence of harmful collinearity may be considerable, on the other hand, if the harmful collinearity does not occur, the most frequent feature of the respective phenomenon is the degrading case.

2.2.7 Elasticity of Technical Substitution

The elasticity of substitution between two goods, is a measure of the case with which one can be substituted for the other. The elasticity depends on the substitution effect and so can be

measured from the indifference curves. Stonier and Hague say that "elasticity of substitution, like the substitution effect upon which it depends, can be measured at any point on any indifference curve. It measures the extent to which goods can be substituted for one another, if the consumer is to retain on that given indifference curve."

This notion also applies in case of factors of production, where it is known as elasticity of technical substitution. J.R. Hicks observes that the elasticity of substitutions is the measurement of case with which the variable factor can be substituted for other. Mrs. John Robinson says that the elasticity of substitution is "the proportionate change in the ratio of amounts of the factors divided by the proportionate change in the ratio of their marginal physical productivities." The elasticity of substitution between two factors would be 'infinite' if one can be substituted for another without any trouble at all. On the other hand the elasticity of substitution would be, 'zero', if one factor can not be substituted for another at all. When it is difficult to substitute one factor for another, then a small change in proportion of the two factors will bring about a large change in the marginal rate of technical substitution between the two factors. But when the substitution between two factors is easy then a small change in the proportion of two factors of production would not bring much change in the marginal rate of technical substitution-between the factors. So from the change in proportion of two factors and the resultant change in the marginal rate of technical substitution, we can know the elasticity of technical substitution :-

$$\text{Elasticity of Substitution} = \frac{\text{Proportionate change in the amount of x with respect to y}}{\text{Propotionate change in the marginal rate of technical Substitution of x for y}}$$

$$es = \frac{\frac{x}{y}}{\frac{d(MRTS)}{MRTS}} = \frac{d\left(\frac{x}{y}\right)}{d\,(MRTS)} \cdot \frac{(MRTS)}{\left(\frac{x}{y}\right)}$$

Similarly, elasticity of demand can be known in case of demand for consumer goods. We have to replace factors and MRS in place of commodities and MRTS, respectively in the above formula.

We shall use the following two diagrams to easily understand the concept of elasticity of substitution : In Figure 7.1, an isoquant curve between two close substitutes has been drawn. Due to the close substitution between factors and isoquant is near to be the straight line. In figure 7.2, an isoquant between two complementary factors has been drawn. The convexity between two factors is very large, because the substitution between two complementary factors is difficult.

We know that the marginal rate of technical substitution at a point on the isoquant curve can be known from the slope of the tangent drawn at the point. In the following two

figures on points A and B of the two isoquant curves, two tangents PQ and RS have been drawn, their slope indicates marginal rate of technical substitution at them. In figure 7.1 and 7.2 the change in marginal rate of technical substitution between points A and B, respectively is equal because the corresponding tangents in the two figures are parallel to each other i.e. PQ in figure 7.1 is parallel to PQ of figure 7.2 and RS of figure 7.1 is parallel to RS of figure 7.2. It means when we move from A to B the change in the marginal rate of technical substitution (MRTS) is the same in both the figures. Whereas in both the figures fall in MRTSxy between A to B is the same, the increase in the quantity of x on isoquant curve of figure 7.2. It is clear from the two figures that the distance $M_1 M_2$ in figure 7.1 is much greater than the distance M_1M_2 in figure 7.2. Thus, in figure 7.2 the same relative change in MRTS as that in figure 7.1 brings about a small increase in quantity demanded of factor x i.e. elasticity of substitution between two factors shown in figure 7.2, is very low. But in figure 7.1 with the same relative change as in figure 7.2, the increase in quantity of factor x is very large, i.e. elasticity of substitution between two factors is very large.

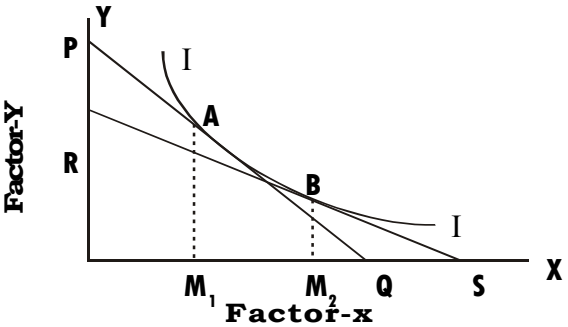


Fig. No. 7.1

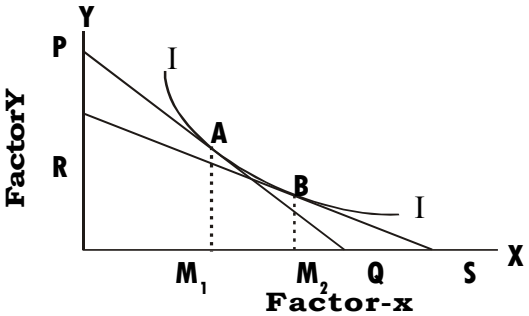


Fig. No. 7.2

The factors, which are perfect complementary to each other are used in fixed proportions and no substitution between them is possible. So the marginal rate of technical substituion between perfect complementary factors is zero. But in the real world, perfect complements rarely exist because some substitution is always possible between the factors. On the other hand, when the two factors are perfect substitutes of each other, then the proportion between them can be increased infinitely without any change in the MRTS between them. So the elasticity of substitution between such factors would be infinite. But in the real life such factors rarely exist. So, it can be said that we mostly come across such situations in the real world where elasticity of substitution between factors (or goods) ranges between zero and one.

Cobb-Douglas production function: It establishes the relationship between its inputs (generally two inputs i.e. labour and capital) and output. It studies impact of changes in the inputs, factor intensity and relative efficiency of the firms.

Constant Elasticity of Substitution: It shows the production technology has a constant percentage change in factor proprtion due to %age change in MRTS i.e. Marginal rate of technical substitution.

Long Questions:

Q.1 Explain Cobb-Douglas Production Function along with its Properties?

Q.2 Explain Constant Elasticity of Substitution (CES) Production Function along with its properties?

Q.3 Briefly explain Technological Progress and the Production Function?

Q.4 Discuss about the translog production funtion and elasticity of technical substitution.

Short Questions:

Q.1 Write two properties of the Cobb Dougles Production Function.

Q.2 Describe about Elasticity of Technical Substitution?

Q.3 Capital Intensive Technical Progress?

Q.4 Labour Intensive Technical Progress?

Q.5 Neutral Technical Progress?

LESSON NO. 2.3 **AUTHOR : DR. BALBIR SINGH**
INDERJEET SINGH

THEORIES OF COSTS

2.3.1 Introduction
2.3.2 Objectives
2.3.3 Theories of cost
2.3.3.1 Traditional Theory of costs
2.3.3.2 Modern Therories of Costs

2.3.1 Introduction

The significance of costs can not be ignored especially in the modern world. Costs are one of the important factors that determine the size and quality of the product. The costs influence both the supply and demand in a society. The cost of production in relation to price advises a firm to produce or not to produce and determine the level of production or service to be provided to the customers.

2.3.2 Objectives

"In this lesson we shall discuss (A) traditional as well as (B) modern theory of costs.

2.3.3 Theories of cost

The costs bold very importance in the business/production activities. In this lesson 8.3.1 traditional as well as 8.3.2 modern theories of costs, will be discussed.

2.3.3.1 Traditional Theory of Costs

In the short run, almost all the factors of production are fixed in quantity and the total fixed costs (TFC) include the expenditures of the firm per unit of time for all the fixed inputs. Similarly, the total variable costs (TVC) is the total expenditure incurved by the firm per unit of time for the all variable inputs. Total costs equals total fixed costs plus total variable costs. Some other concepts of costs include explicit costs and implicit costs. Explicit cost include the value of actual inputs used by the firm to produce the product. The value inputs owned by the firm should be inputed or estimated from what they could earn in their best alternative use.

The sum of total fixed cost and total variable costs is called the total cost of production. The total fixed cost curve is horizontal straingth line to the OX axis which shows that whatever the quantity of output the same i.e. constant. The total variable cost curve is sloping upward rising from zero output initially gradually and later at a fast speed. Besides this, the total cost of the vertical addition of total fixed cost (TFC) and total variable cost (TVC). Since the TFC is constant the difference between TC and TFC will always continue to be the same.

One or more factors of production are fixed quantity in the short run. Total fixed cost (TFC), total variable cost (TVC) and total cost (TC) have been shown in fig. no. 1

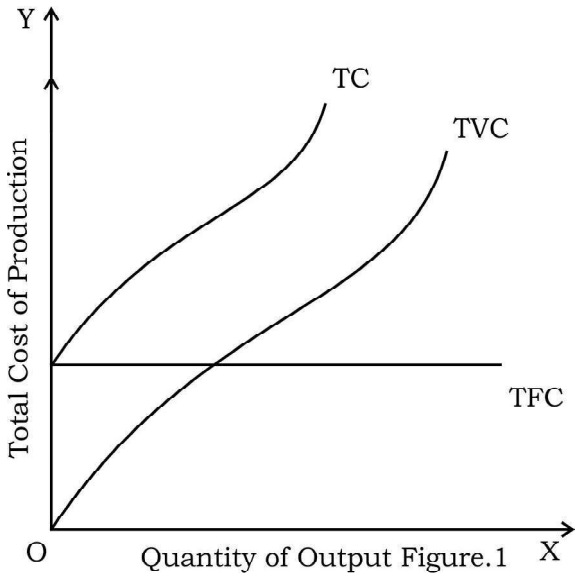


Figure.1

The Short-run Cost Curves

The short-run cost curves include the average fixed cost, the average variable, the average cost, and the marginal cost curves. Average fixed cost is equal to the total fixed cost. Average variable cost equals total variable cost divided by output. Similarly, average cost can beknown by dividing it with total output. Average cost is equal to average fixed cost plus average variable cost. Marginal cost is the change on total cost per unit change in output.

In figure No. 2 we observe that average fixed cost contnue to decline as the output increases. As the output multiplied by the average fixed cost is the same, the total fixed cost remains the same. But average variable cost, initially, decrease upto a certain point, then starts to increase.

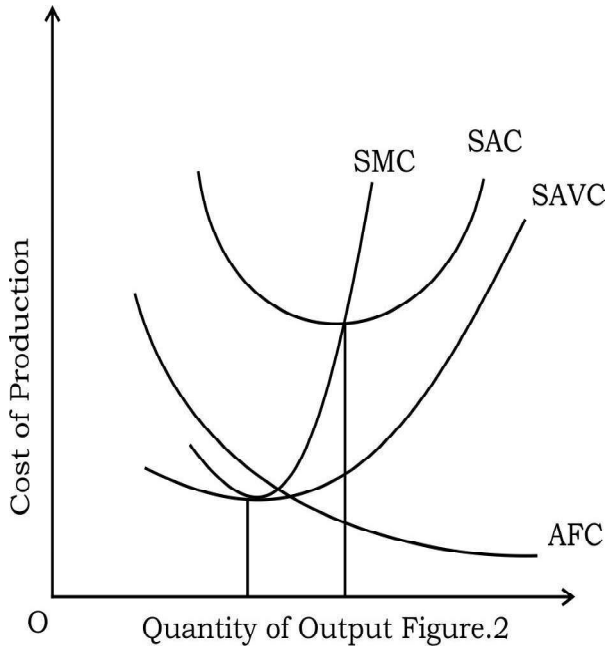
Self-Check Exercise

Q. What do you mean by total cost of Production? Draw its diagram?

Ans. .....

.....

.....



It is because of due to the application of the law of the diminishing returns (or increasing costs) that the AVC curve rises. That is why it becomes U shaped. The marginal cost is the net addition to the total cost as one more unit of production i.e., additional cost made in producing the additional unit.

Long Period Cost Curves

But, what about the long-period cost curves of a firm ? As already observed, all factor inputs are assumed to be freely variable in the long period. There is a view represented by economists like Knight, John Robinson and Kaldor that if all factors are freely variable and the commodity is perfectly divisible, a large firm has no advantage over a small firm under perfect competition, for this view inputs all economies and diseconomies in production to the existence of indivisible factors and assumes that there are no independent economies or diseconomies of scale. The scale of a firm is increased accordingly as it increases all of its factors inputs together in the same proportion so that the proportion in which they are combined remains the same. Naturally, then the long period cost curves of the firm (LAC and LMC) will be horizontal as shown in Fig. 3 below :

Self-Check Exercise

Q. Define marginal cost? Also tell the reason why marginal cost is U-Shaped?

Ans. .....

.....

.....

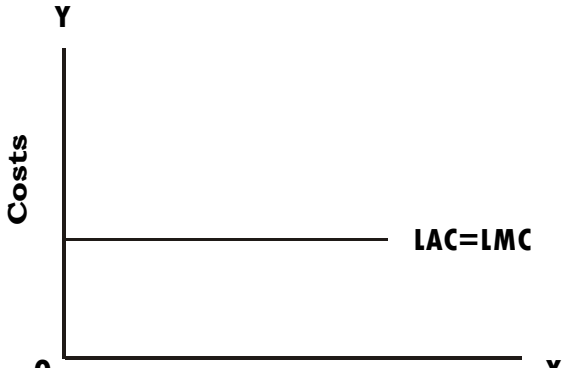


Figure 3

But there is an opposite view represented by Chamberlin, according to whom there are independent economies as well as diseconomies of scale. If we agree with this view, the long period cost curves will also be U-shaped as shown in Fig. 4 below:

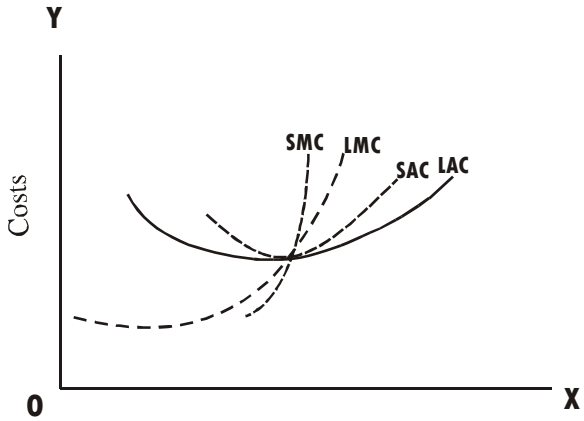


Figure 4

However, they will be flatter than the short-period cost curves for the obvious reason that in the long period it is possible to adjust the factor proportion adequately and thus to reduce the rate at which total cost tends to change in the firms's output.

2.3.3.2 Modern Theories of Costs

The traditional theory of cost has been criticised by a number of scholars and and they have given new theories of costs. The writers of this new approach were mainly G. Stigler, Sargent Florence, C.A. Smith, K.J. Arrow etc., It was in 1939 that G. Stigler advanced the idea that average variable cost in the short-run has a flat stretch over a period of time. It is also called saucer-shaped short run average variable cost (SAVC). It is also argued that this flatness is mainly due to the fact

that firm helps some provision for additional productive capacity which is known as the 'reserve capacity' to increase the output in the context of abrupt demand of its product.

The Average Fixed Cost

Some changes have been introduced regarding the average fixed cost (AFC). Basically, this cost comprises of the cost of physical and personal organisation of the firm i.e., the following one;

(1) The salaries of staff employed directly in production and paid on a fixed term basis.
(2) The wear and tear of machinery.
(3) The salaries and expenditures of administrative staff.
(4) The costs on maintenance/repair of buildings.
(5) Expenses on the maintenance of land.

The importance of AFC is that it mainly determines the size of the plant. The entreprenur would like to plan the level of output that he can produce efficiently and flexibly. Moreover such a plant will have the capacity more than the 'expected average' level of sales. It is because of the businessman prefers to have some 'reserve capacity' due to some reasons. Actually, he wants to meet seasonal and cyclical fluctuations in his demand. Secondly, the reserve capacity would provide more freedom to increase his output in case of abrupt increase in demand for his product. Thirdly, the reserve capacity is required to have time for repairs without affecting the running of the plant continuosly. At fourth place, some kind of technology needs reserve capacity because it is very difficult to install and operate time and again. Even at the organisational and administrative level some reserve capacity is needed for employing the administrative staff to allow some increase in the production operations of the firm.

The figure 5 shows that the firm has some largest-capacity units of machinery which causes setting an absolute limit B in case of short run increase in output. Besides, the firm has small unit machinery which limit expansion of output to boundary (A) given in the diagram. But this is not an absolute boundary as the firm can expand its output in the short-run to absolute limit B. This can be done either by paying overtime to direct labour for working longer hours. In this context AFC is depicted by dotted line in the figure or this limit can be increased by purchasing some additional small units types of machinery. In this context AFC curve shifts upwards but starts declining again.

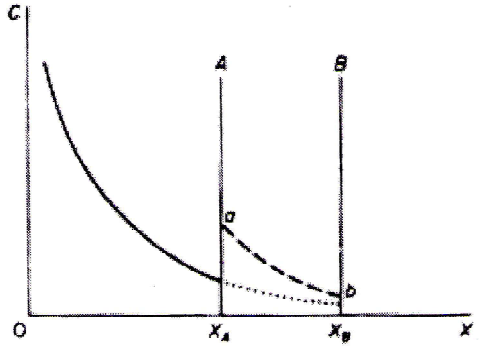


Figure 5

The average variable cost

In the modern theories of cost, P.W.S. Androw's theory come to be one of the most recognised theories regarding average variable cost. In his theory cost consists of (i) raw materials (ii) direct labour that changes with output and (iii) running costs of machinery.

In the short run, average variable cost is almost like U shaped but having a flat stretch over reasonable range of production as is shown in the figure 6. This flat stretch shows that the SAVC coincides with MC which remains per unit of output. Whereas to the left of this range of stretch MC remains below SAVC but on the right side MC remains above the SAVC. The declining portion of the SAVC indicates fall in costs because of better utilisation of the fixed factors. It is mainly because of this both skills and productivity increase of variable factor i.e. labour. This better skill further reduces the cost of production by making the proper use of raw materials etc.

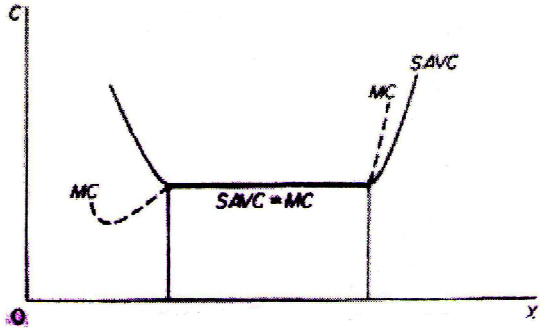


Figure 6

But the rising part of the SAVC indicates the decling productivity of labour due to frequent breakdown of machinery because of its overrise, the increasing cost of labour due to overtime payments, etc. The importance of SAVC with having a certain

as flat portion that represents a reserve capacity of the firm necessary for fulfilling the urgent demand of its product. The reserve capacity has been shown in figure 8.

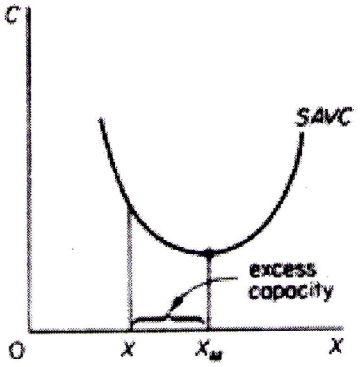


Figure 7

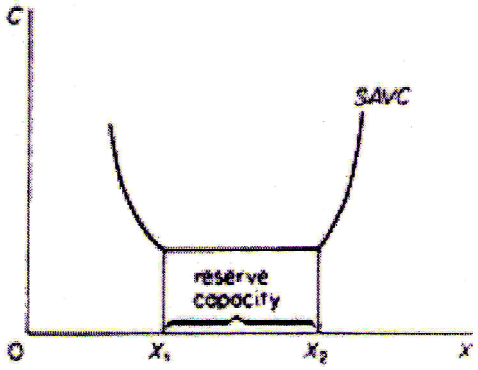


Figure 8

But it is different from the excess capacity that increases with the U shaped cost curves as given in the traditional theory of firms. This traditional theory in a way take the assumption that each plant is desined without any flexibility. It is designed in such a way that optimal production take place only in a single level of output (x_m in the figure 7). If a firm produces output x which is smaller than x_m $x_1$there is excess capacity i.e. equal to the difference between x-x_1. This excess capacity is considered undesirable as it leads to higher units costs.

In the modern theory of costs the range of output x_1x_2 as shown in figure 8 reflects the planned reserve capacity which does not lead to increase in costs. The firm anticipates using its plant sometimes closer to x_1 and other closer to x_2. Usually on an average an entrepreneur wants to operate his plant within the range of x_1 x_2. It is said that generally firms consider the 'normal' level of utilization of the plant which may be somewhere between two-third and three-quarters of their capacity. It

has been shown at a point closer to X_2 and X_1.

The average cost total (ATC) curve

In the modern theory, it becomes desirable to know the shape of the average total cost curve especially when the variable cost curve has a saucer type shape. This has been depicted in figure no. 9.

The ATC curve continues to decline upto certain level of output (x_A) and at this point the reserve capacity ends. At this level of output the MC intersects the average total cost curve at its minimum level. This takes place to the right of the level of output x_A, where the flat portion of SAVC comes to its end.

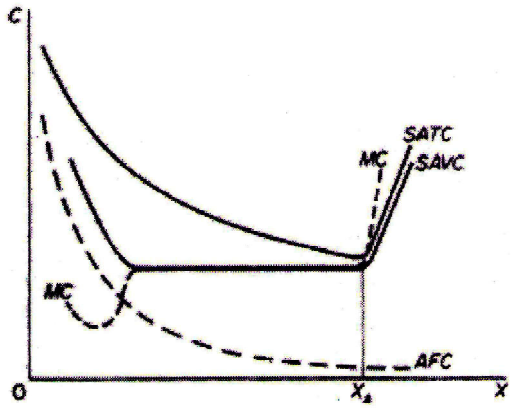


Figure 9

Long-run costs in Modern Theory : L-shaped cost curve

In the long-run all costs are considered variable costs and these costs make the long-run cost L-shaped. These costs have been further divided into production costs and managerial costs. At the large scale of production managerial costs may increase but on the other hand, the production costs decline more than the increase in managerial costs. It is due to this fact that the total LAC curve declines with the increase in scale.

Production costs decline steeply in the begining and then gradually. As the scale of production increases, the L-shape of the production cost curve is described by the technical economies of the large scale production. If new techniques are introduced by the large firms they must be cheaper to operate. Any how even with the existing known technology some economies can always be achived at the larger output. It may be because of further decentreaisation and skill improvement secondly, less repair costs as firm reaches a certain size and thirdly a multiproduct firm may well undertake itself in the production of some of the material or equipment which it requires rather buying from other firms managerial costs. There exists various

level of management each with its appropriate kind of management techniques. There are small scale and large scale organisational techniques, having different techniques of management that fall upto certain plant size but at very large scale of production managerial costs may rise but at a very slow rate.

In brief, production cost falls smoothly at large scales but managerial cost may rise at a slow rate at very large scale.

Today, modern scholar's seem to accept the decline in technical costs more than that of the probable rise of managerial costs, that is why the LRAC curve falls smoothly or remains constant at a very large scale output. LAC may be drawn implying that each short period SRAC which includes production costs, administrative costs, other fixed costs and an allowance for normal profit. LAC curve may be drawn by joining the points on the SATC curve corresponding to the two thirds of the full capacity of every plant size. This has been shown in figure 10.

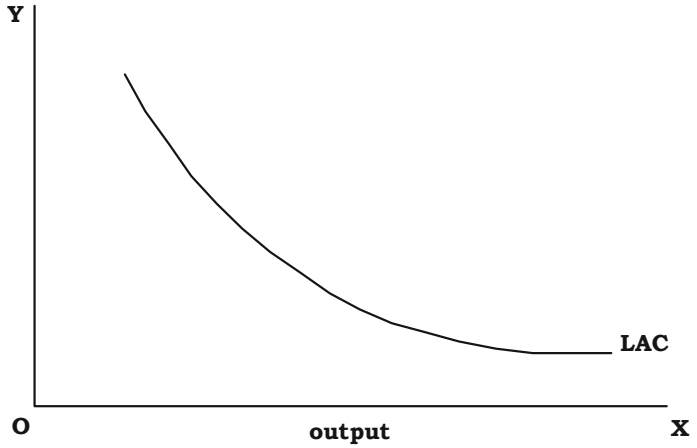


Figure. 10

If LAC declines continuously the LMC will be below the LAC at all scales as has been shown in figure 11.

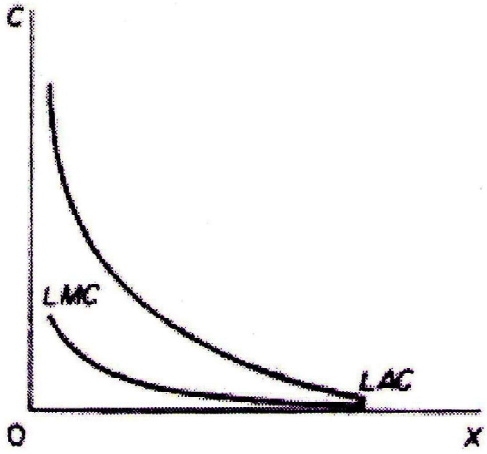


Figure 11

There is a minimum optimal scale of plant (in figure 11) at which all possible scales of economies are enjoyed, but beyond that scale the LAC remains constant. Here the LMC lies below the LAC until the minimum optimal scale is achieved and coincides with the LAC beyond that level of output as shown in figure 12.

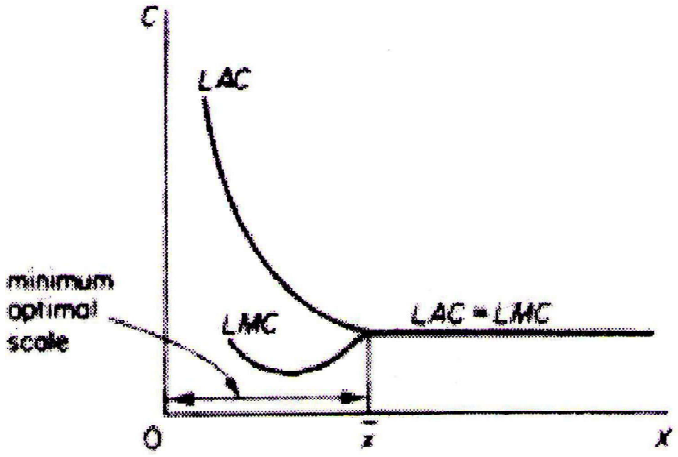


Figure 12

The above shapes of costs are more realistic than the traditional theory using U shaped cost curves.

Long Questions

Q. Explain Short-run and long run cost curves of Traditional theory of costs?

Q. Explain in detail Modern Theories of costs?

Short Questions

Q. Average Fixed

Q. Average variable cost

Q. Average Total cost

LESSON No. : 2.4 **AUTHOR : PROF. UPINDERJIT KAUR**

PRICE-OUTPUT DECISIONS UNDER PERFECT COMPETITION
(With Marginal Approach)

2.4.1 Introduction
2.4.2 Objectives
2.4.3 Market Forms
2.4.4 Price decisions under perfect competition
2.4.5 Equilibrium and Increasing Returns

2.4.1 Introduction

It is mentioned in the preceding lesson that the main task of theory of production is to discover the nature of the supply function of an industry. The supply function or curve of an industry is made up of the individual supply curves of the firms which make up the given industry. Hence we shall, first, enquire into the nature of an individual firm's supply curve and on the basis of it we shall reach at certain conclusions with regard to the nature of the supply curve of an industry under condition of pure competition. But we must first be clear about the meaning of the term "pure competition".

2.4.2 Objectives

The main objectives of this lesson are : - to explain the various types of market forms. - to discuss price-output decisions under perfect competition.

2.4.3 The Market Forms

There are different forms of markets that can be observed in a capitalist economy. These can broadly be classified into (a) perfect markets; and (b) imperfect markets, though the latter type may further be classified into (i) monopoly, (ii) monopolistic competition, (iii) oligopoly of which duopoly is only a special case, (iv) bilateral monopoly.

A perfect market is one in which there prevails perfect competition. Competition in a market is regarded to be perfect, if the following conditions are satisfied. First, the good must be homogenous, that is, consumers or the buyers regard all the individual units of it as equal in all respects regardless of whether a unit of the goods is being supplied by this firm or that firm. Second, the number of firms selling the particular good is so large that any change in the supply of individual firms has no significant effect on the total demand for the good. Third, there is perfect mobility of buyers and sellers as well as of all the factors of production. This implies that buyers can freely move from one part of the market to another in order to buy from the cheapest seller; and the seller can freely move from one part of the market to another in order to sell at the highest possible price. Similarly, factors of production can also move from a low

paid to high paid industry. Fourth, there is free entry into and a free exit from the industry, that is, an individual firm is free to enter or leave the industry, without any restriction. Fifth, there should be perfect knowledge on the part of the sellers as well as the buyers with regard to price being paid and charged. Sixth, there should be absence of cost of movement or transport as between different parts of the market. Lastly, absence of irrational preferences on the part of buyers is also assumed.

If all the above conditions are satisfied, there will tend to be one and only one price of a given good at any given time and the market will be said to be a perfect market.

However, it is well high impossible to come across such a market in real life. In this sense, it is only an abstraction and a logical concept. Chamberlin has given a less strangent concept of "pure competition". Pure competition is said to exist in a market, if only the following conditions are satisfied;

(i) the goods are homogenous;

(ii) the number of sellers as well as buyers is very large and there is free competition among them;

(iii) the entry and exit are free : These are the maximum conditions necessary for there to be one and only one price of a good at any given time.

The important point to note is that in a perfectly, even purely competitive market an individual seller cannot influence the price of the good by his own individual action; he is not a price maker but is only a price taker.

In principle, when competition in a market is imperfect in some way or the other, the market is said to be imperfect. Some important forms of imperfect markets are as follows :

(i) Monopoly : When there is only one firm in the given industry producing and selling a good which has no close substitutes, but the number of buyers is large and there is free competition among them.

(ii) Monopolistic Competition : When there is a large number of freely competing firms producing, and selling products, though not identical, but similar and therefore, very close substitutes of one another.

(iii) Oligopoly : When there are few competing firms so that a change in the output and price of any of them has a significant influence on the total supply of the goods as well as price of the good. When the number of firms is so small as two only, the market form is described as a duopoly.

(iv) Monopsony : When there is only one buyer in the market where the sellers are large in number and are freely competing among themselves.

(v) Bilateral Monopoly : When there is only one seller and also only one buyer of given good so that there is monopoly on both sides.

However, in the present lesson we shall be discussing price-output decisions on the assumption of perfect competition only.

2.4.4 Price Decisions under perfect competition

We would like to stress the point in the very beginning that under pure or perfect competition, an individual firm does not fix its price. The reason is that the elasticity of demand for the good of the individual firm under pure or perfect competition is infinite. This means that it can sell any quantity at the ruling price in the market. Therefore, it should have no motive to sell at a price lower than the ruling price. On the other hand, if it raises its price, the products of the rival firms being perfect substitutes of it (the good being homogeneous); none will buy from it, all buyers will shift to the rival firms. Hence a competitive firm will not experiment with price-making. It will only try to adjust its output to the prevailing price situation in order to maximise profits or to minimise losses, if it is running at a loss. In short, it is not a price maker but quantity adjuster. It takes the price as given, which has been determined by the interaction forces of aggregate demand and aggregate supply of the commodity.

In the short period a firm can adjust its output without changing the size of its fixed plant. As a matter of fact, short period is defined as a period which is so short that it is not possible for an existing firm in the industry to change the size of its fixed plant or to leave the industry and it is not possible for a new firm to enter the industry. So we can say that in the short period, the number and the size of the individual firms in the industry remain constant and the individual firm can change its output only by changing the quantities of the variable factors used with the fixed factors of the firm.

Now the question is how an individual firm varies its output in response to changes in the price of the good. For this will provide the clue to the nature of the firm's supply curve in the short period. Since we assume that the objective of a firm is to maximise its profit or to minimise its losses in the short period, we must know the condition of the maximisation of profits. The necessary or the "first order" condition of it is that the firm's marginal cost must equal its marginal revenue. If the firm's revenue is greater than its marginal cost, it means that the firm can add more to its total revenue than it will be adding to its total costs by increasing its output, and this will have increased profits. Hence, the firm will not be in equilibrium and will tend to expand its output. On the other hand, if the marginal cost is greater than the marginal revenue, it means that the additional output adds more to the total cost than to the total revenue of firm and thus reduces its profits. Hence the firm will again be not in equilibrium and will tend to reduce its output. The profits are maximum or losses are minimum when the firm's marginal cost equals its marginal revenue. However, this is only the necessary but not sufficient condition. The additional or "Second order" condition is that to the right of equilibrium point, the marginal cost must be greater than the marginal revenue. If it is otherwise, as is the case at Point P_1 in Fig. 1 the firm's profit will be increased by an increase in its output. But both, the "first order" and the "second order" conditions are satisfied at P1. Therefore the firm's profits are maximised at output Q_2 and not at Q_1 in Fig. 1.

Self-Check Exercise-I

Q. Define oligopoly?

Ans. .....

.....

.....

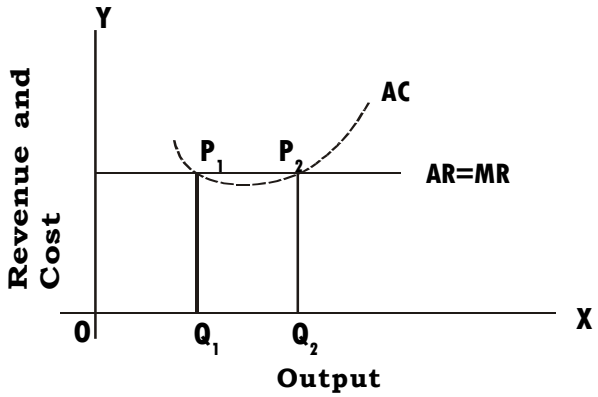


Fig No. 1

Keeping in view the argument of the above paragraph, we can now explain the output decisions of the individual firm in the short period with the help of the following diagram. Under pure competition, the demand for the product of an individual firm is perfectly elastic. Therefore, the average revenue of the firm will remain constant with changes in its output as shown by the horizontal curve in Fig. 2. Therefore, the marginal revenue of the firm, as it was explained in the previous lesson, will equal its average revenue, and will also be constant. Thus the various horizontal lines like AR1=MR1 AR2 MR2... in Fig. 2 represent the average revenue as well as the marginal revenue of the firm at different levels of its output. Each AR = MR line refers to a particular price level such as $P_1P_2P_3$.....

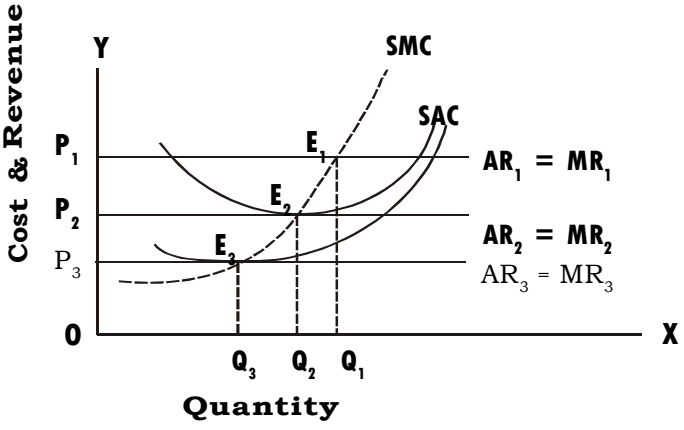


Fig. No. 2

Self-Check Exercise-II

Q. Define Bilateral Monopoly?

Ans. .....

.....

.....

Let us suppose that the ruling price is P_1, $AR_1 = MR_1$ line will then represent the average as well as the marginal revenue of the firm. AVC curve represent the behaviour of the average variable cost of the firm in the short period, while SAC and SMC represent the behaviour of the short period average costs and short-period marginal cost of the firm respectively. When they behave in the manner shown in Fig. 2 which was explained to you in the preceding lesson. The rising SMC cuts the MR_1 line at E_1, where the short-period marginal cost equals the marginal revenue of the firm and to the right of which the marginal cost is greater than the marginal revenue. Hence both the conditions of firm's equilibrium are satisfied at E_1, therefore, at price P_1, the firm will be in equilibrium at E_1 and it will produce Q_1 output. It is earning here super normal profits. If the price now falls to P_2 the $AR_1 = MR_1$ horizontal line will also fall to the level AR2 = MR2. Now the condition of equilibrium will be satisfied at E_2. Therefore, the firm will reduce its output to Q_2. It can be seen in Fig. 2 that at Q_3 the average revenue of the firm is less than its average cost by SE_2 amount. Hence the firm will be running at a loss, though loss will be the minimum at output Q_3. But you may be tempted to ask why the firm does not shut down its plant and stop producing. The answer to this question is that the firm cannot leave the industry in the short period and hence it has to incur the fixed costs even when it stop production. But by producing Q_1 output, it is able to recover at least a part of its fixed costs and thus its losses are less than they would be, if it stopped production completely, if the price falls still further to P_3, then $AR_3 = MR_3$ Line will represent the average and the marginal revenue of the firm. The equilibrium of the firm now takes place at E_3 and the firm decides to cut back its output to Q_3. Here the price just equals the average variable costs. If there is a further fall in the price, it can be seen by imagining a MR line below MR_3 . Line that the firm will produce and supply nothing at a price less than P_3 in Fig. 2 above which equals the average variables cost. Price P_3 is known as "Shut down" price and point E_3 is known as "Shut down" point of the firm.

The above analysis leads us to the following conclusions : (i) The individual firm increases its output and supply with increase in the price of good concerned and reduce its output and supply with a decrease in the price which implies that the firm's short period supply curve is positively sloped, that is rises upwards to the right; (ii) Since all points of equilibrium lie on the rising portion of the firm's short period marginal cost curve, it is this portion of the SMC curve which lies behind the short period supply curve of the firm and explains the positive slope of this supply curve; (iii) Since the firm stops producing at a price lower than the average variable costs, therefore the short period supply curve of the firm does not extend below the level of the minimum average variable costs as shown in Fig. 2.

The supply curve of the firm is portrayed below.

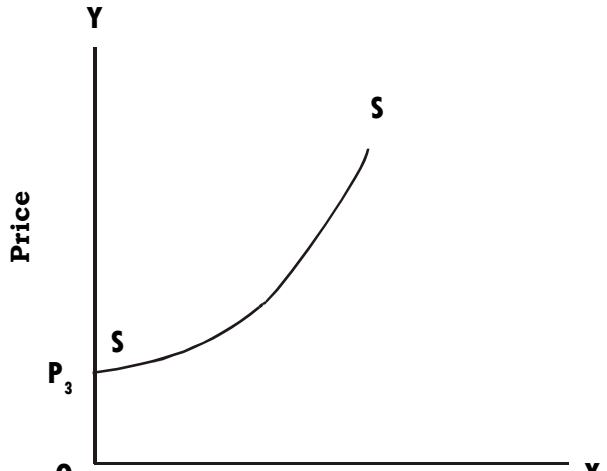


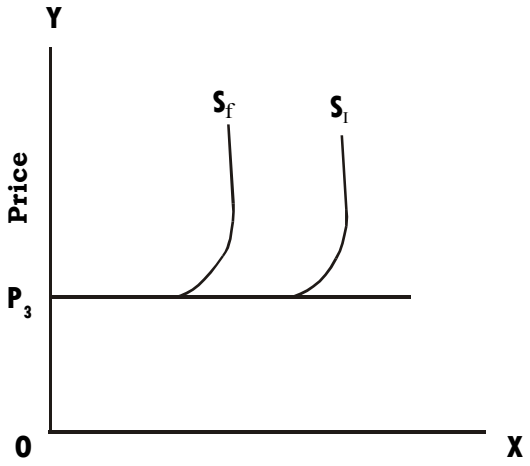
Figure No. 3

The Short-Period Supply Curve of the Industry

An industry under pure competition is made up of a large number of firms. If for the sake of simplicity of analysis, we suppose, that all the individual firms employ identical factors of production and technology, all of them will have identical cost curves. Therefore, at any given price, every firm will be producing the same output as the other. In terms of Fig. 2 at price P_1 every firm will be producing the same output as the other. In terms of Fig. 2 at price P_1 every firm will be producing OQ_1 output. If there are, say x firms in the industry, the industry's output at price P_1 , will be OQ_1 n. Similarly, at price P_2 the industry's output will be OQ_2, and so on. Since at a price less than P_3, that is at a price less than the minimum average variable costs, no firm will produce any positive output, the industry output, will be zero at any price less than P_3.

What has been said above implies that the industry's short-period supply curve is only a magnified form of the individual short-period supply curve. It is in fact, derived by the process of lateral summation of individual short period supply curves of all the firms in the industry or by multiplying the individual firm's supply curve by the total number of firms in the industry. However, it should be noted that the latter statement is true only if all the firms have identical cost curves. This naturally implies that the industry's short period curve is also positively sloped like the short-period supply curve of the individual firm. In Fig. 4 below, Sf represents the short period supply curve of an individual firm while SI represents the supply curve of the industry. Both have positive slope as both are rising upward to the right. As in Fig. 2, price P_3 in Fig. 4 represent the minimum average variable costs of the individual firm. Like the firms the industry's supply curve is also cut off at the level as shown in the figure. The industry's supply curve is to the right of the

firm's supply curve, at any price, the industry's supply must be greater than the supply of an individual firm in the industry.

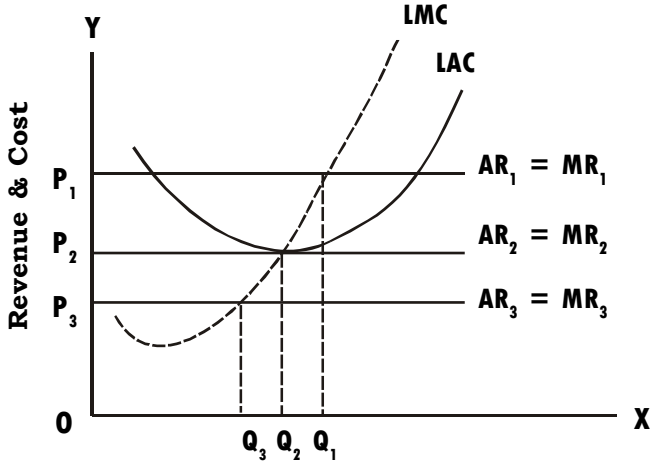


Supply
Figure No. 4

A final point to be remembered regarding the short period supply curve of the individual firm as well as that of the industry is that they are relatively inelastic because in the short period, a firm can change its output by varying only the variable factors, while the other factors remain fixed. Full adjustment is not possible in the short period, because the scale of the plant remains fixed.

Price Output Decisions in the Long Period

Long period is a period, which is long enough to enable the individual firm to change the scale of its operations by changing all the factors involved and enable the existing firm to leave and new firms to enter the industry. In other words long period is a period in which the size as well as the number of individual firms can change.



Quantity
Figure No. 5

If the demand conditions for the product of the industry are such that an individual firm is earning super-normal profit, it will try to increase its scale. On the other hand, if the conditions are such that it is not possible to recover even costs, the firm will seek to decrease the scale of its operations. A long period is indeed made up of a number of short periods, each short period associated with a definite scale of the plant. If we assume that there are non proportionate returns to scale too, as are the non-proportionate returns to factors proportions, the long period cost curves of the firm will also be U-shaped, as it was observed in the preceding lesson. The only difference compared to the short period cost curve, is that these are relatively shallow as there are no fixed costs in the short period, for they are ruled out by definition. A firm's long period costs curve is like that depicted in Fig. 5 above. It should be noted that the long period costs of a firm include its normal profits that is minimum amount of profits which the firm must earn in the long period, if it has to continue to be in the given industry. When the firm fails to earn this minimum profits it leaves the industry.

If the price of the good happens to be P_1, the horizontal $AR_1 = MR_1$ represents the firms' average revenue as well as marginal revenue. The firm will tend to produce that output at which its rising marginal cost equals its marginal revenue, for as already explained, this is the condition of profit maximisation. This, the profit maximising output will be Q_1. But at this output, the firms average revenue is much above the average cost. Hence, the firm will be earning supernormal profit which will attract new firms into the industry. The competition will increase and consequently the price will come down to P_2. Now the profit maximisation output of the firm will be Q_2 at which the firm's average revenue just equals its average costs and the firm just manages to earn normal profit. Assuming all firms to be identical, each firm will be in equilibrium at Q_2 output. If the price happens to be less than P_2 the firm will fail to recover even its costs as would be the case at price P_3 in Fig. 5. The firm fails to earn even normal profits. Hence, some of the individual firms will begin to leave the industry. The total output and supply of the industry will decrease and the price will ultimately rise to P_2 so that the remaining firms are able to earn just normal profit. Thus, the point E_2 in Fig. 5 above represents the long period equilibrium of the firm and its long period equilibrium output is Q_2.

If we focus our attention on the point E_2 of long period equilibrium of the firm, we can specify the conditions of the long period equilibrium of the firms which are as follows : (i) the firm's long period marginal cost must equal its marginal revenue, (ii) the marginal cost must be rising and (iii) the average cost of the firm equals its average revenue. The last two conditions contain with in themselves the

following conditions also; (a) the firm's total costs equal its total revenue and therefore (b) the firm earns just normal profits neither more nor less. Moreover, in the long period, the firm's price or average revenue equals the average cost as well as the marginal cost; in other words, in the firm's long period equilibrium under pure competition, can be written as :

AR = MR = MC = AC

One thing must be noted with regard to long run supply curve of the industry that it is not derived by the lateral summation of the long run marginal cost curves of a given number of firms. It is so because the longrun supply of a firm is determined when the price of the product is equal to both the marginal cost as well as the minimum average cost (Price = LMC = LAC). Secondly, the number of firms in the industry goes on changing as the price of product changes. Therefore, the quantity supplied by an industry in the long-run at a given price will be determined by the optimum output of a firm in the long run multiplied by the number of firms in the industry.

2.4.4 Equilibrium and Increasing Returns

While discussing the conditions of a firms' equilibrium in section 2 above we have specified two conditions for a firm to be in equilibrium. First, its marginal cost must equal its marginal revenue. Second, to the right of the point of equality between the MC and MR (i.e., the point of equilibrium), the MC must be greater than the MR.

In the light of the above conditions, it can be seen that when the firm is producing its output under the law of increasing returns, its equilibrium under conditions of pure or perfect competition shall be in doubt. When the production of the firm is subject to increasing returns, the marginal cost of the firm will go on falling with increasing output as shown by the MC curve in Fig. 5 (i) below. But if the firm is operating under conditions of pure competition, its average revenue and marginal revenue will remain constant with increase in its output as indicated by the horizontal P=AR=MR line in Fig. 6. (i) The firm's MC and MR are equal at output Q. But this is not the point of equilibrium output of the firm and R is not the point of equilibrium of the firm. The reason is that at the point of intersection between the firm's MC and MR curves (i.e. at R), the second order condition of the firm's equilibrium is not satisfied. To the right of R, the firm's MC is not greater than its MR. On the contrary, its MR is greater than its MC. The firm can very well see that it can increase its profits by expanding its output. But, if the increasing returns are not exhaustible the firm's marginal cost will go on falling, while its marginal revenue will remain the same. Hence the equilibrium will not take place under pure competition if it continues to enjoy increasing returns. However, it may happen

that as a firm goes on expanding, it may become so large that it may begin to control a very substantial portion of the total supply of the industry. As soon as this happens, there will be no longer pure competition in the market. The market will become imperfect in some form or the other; it may be transformed into a monopoly or a duopoly or an oligopoly. In an imperfect market the AR and MR curves of a firms fall downwards to the right. If the firm's MR is falling more rapidly than its MC the firm can attain a state of equilibrium under increasing returns. But there is no longer pure competition. Hence, perfect as well as pure competition and state equilibrium are said to be incompatible under increasing returns. To quote Prof. Kaldor, "long period stable equilibrium and perfect competition are incompatible assumptions."

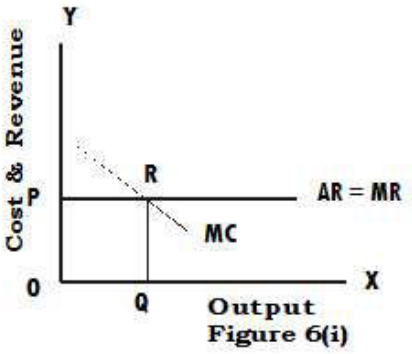


Figure 6(i)

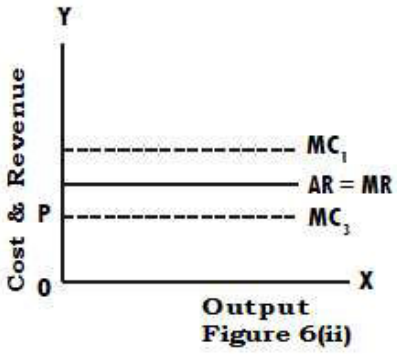


Figure 6(ii)

It may also be appreciated that the firm will not be in equilibrium at any positive output even when the firm is producing under constant returns. This can be illustrated with reference to Fig. 6 (ii) above. When there are constant returns, the firm's marginal costs remain constant. If the firm's marginal costs are represented by the horizontal line MC_1 , no output is profitable. So the firm will not produce any output. If the marginal costs are represented by the horizontal line MC_2 the firms profits will go on increasing idefinitely with increase in its output. Hence the firms' equilibrium will be indeterminate as long as there is perfect competition. If the firm's MC line happens to coincide with its MR line, the equilibrium is again indeterminate, for it can take place anywhere on the MC=MR line.

But under perfect competition, equilibrium, cost is rising so that MC should cut the horizontal marginal revenue curve from below at the point of equalities

between MR and MC as is at point E in fig. The equilibrium of the firm is established at point E or OM output at which MC is equal to MR and also MC is rising. It is not profitable to expand output beyond OM (or point E) because after if marginal cost is greater than price and marginal revenue. So equilibrium of a firm under perfect competition is possible under decreasing returns to scale or increasing cost.

But the equilibrium of the firm under perfect competition is possible only if the marginal cost is rising and that it cuts marginal revenue curve from below. In fig 6(iii) this point is equilibrium has been depicted at E where MC curve cuts the MR from below. It is not desirable to increase the output level after this point of equilibrium due to reason that marginal cost becomes greater than price and marginal revenue. Thus, the equilibrium of a firm under perfect competition is possible under decreasing returns to scale or increasing costs. Such a type of situation does not emerge otherwise the firm would become a monopolist firm and content of perfect competition would likely to lose.

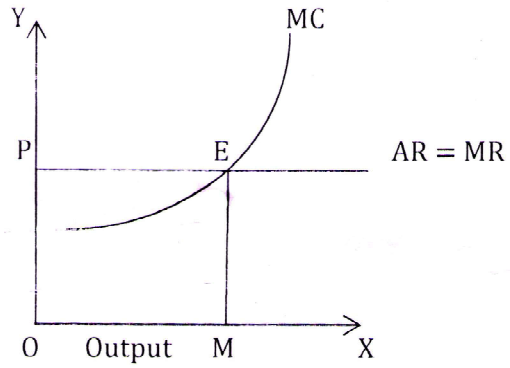


Figure 6 (iii)
Firm's equilibrium under perfect compeition with rising cost

However, it may be kept in mind that equilibrium of the firm under perfect competition will take place if it is the industry and not the individual firm which is operating under increasing returns. This may happen when the growth of the industry generates external economies for the individual firms.

Perfect Competition: This is an ideal situation of market with homogeneous products. All producers are price takes. Buyers and sellers have perfect knowledge about the price.

Shut down price: A firm will produce till the price is equal to average variable cost (AR=AVC) in the short run and an industry need atleast normal profit in the long run.

Long Question

Q. Explain how price and output Decisions are taken under perfect competition?

Q. Explain about Equilibrium and Increasing Returns?

Short Question

Q. The Short- Period Supply curve of the Industry.

LESSON NO. 2.5 AUTHOR : PROF. UPINDERJIT KAUR

PRICE-OUTPUT DECISIONS UNDER MONOPOLY AND PRICE-DISCRIMINATING MONOPOLY

2.5.1 Introduction
2.5.2 Objectives
2.5.3 Monopoly-main features
2.5.3.1 Equilibrium under Monopoly
2.5.3.2 Price- Discriminating Monopoly
2.5.4 Price- Discriminating Monopoly

2.5.1 Introduction

The two limiting cases of market situation are those of pure competition and pure monopoly. These are rarely found in the market. We have already discussed how price and output are determined in a perfectly competitive situation. We are now to study the way in which price and output are determined under monopoly and price discriminating monopoly.

The monopoly situation is an extreme form of imperfect competition. Monopoly can take the form of private monopoly or public monopoly. Similarly, there can be simple monopoly or discriminating monopoly. Under monopoly, the monopolist charges uniform price from all the customers of his product. But under discriminating monopoly he charges different prices from different customers for the same product or charges different prices for different units of the product from the same customer. Price discrimination occurs when different units of the same commodity are sold for different prices for reasons not associated with differences in costs. The essence of monopoly is the firm's power to influence the price of a product.

It was A. Cournot who developed the theory of simple monopoly in his book, *Researches into the Mathematical Principles of the Theory of Wealth (1838)*

2.5.2 Objective: After having gone through this lesson you would be able to:

- Understand meaning and main features of monopoly.
- Conditions of equlibrium under monopoly and price discriminating monopoly.

2.5.3 Monopoly – Main Features

The monopolist is the sole producer or supplier of the product. Under perfect competition there are numerous firms in an industry producing homogeneous product. No single firm exercises any influence over the price of the commodity. All the firms are merely price-takers. They are not price makers. They adjust their output to the level of prevailing price, determined by market forces. But under monopoly there is only one seller or supplier of the commodity. The monopolist thus constitutes both firm and industry. He has full control over the supply of the commodity. But on the side of demand there is perfect competition among the buyers as they happen to be in large numbers. The monopoly is on the supply side only. The control over supply of the product enables the monopoly firm to acquire exclusive power to fix any price for its product or to put any

output in the market for sale leaving it to consumers to decide about the market clearing price. The monopolist thus comes to exercise control over the price of his product. In other words, he is a price-maker and not a price taker. He acts on the belief that the amount which he can sell depends on the price he would charge.

Secondly, the product of the monopolist has no close substitutes. If close substitutes are available, then the monopolist does not enjoy monopoly power. It is then quite likely that he may incur heavy losses if he raises the price of the product beyond a certain limit. But it also does not mean that there are no substitutes at all. Substitutes are there but they are remote substitutes. Consequently the monopolist faces very small competition that can be ignored. Thus, the cross elasticity of demand for the monopolist's product is low or zero.

Thirdly, the monopoly can remain in existence only if strong barriers to the entry of new firms into the industry exist. Unlike perfect competition where there is free entry and existet, there are barriers to entry in pure monopoly. Among the barriers are found patents and licensing by government agencies, raw material control, the establishment of brand names, pricing policy designed to keep rivals out of the industry, large capital investment necessary to enter the industry, and the size of the market. The entry of new firms is thus barred. Potential competitions can also be intimidated by threats of sabotage. Thus various natural, legal and technological factors are responsible for the emergence and growth of monopolies. New firms cannot enter and compete with the monopoly firm. The monopoly firm thus continues to enjoy its privileged position even in the long run. A firm working under perfect competition can earn only normal profits in the long run because supernormal profits are eliminated with free entry of firms. On the other hand, the monopoly firm protects its supernormal profits by restricting the entry of prospective competitors in the market. The monopoly firm thus continues to enjoy certain privileges which are denied to a competitive firm. The monopolist can increase the sales of his product by adjusting the price downward.

Fourthly, like a producer working under perfect competition, the monopolist also aims at maximisation of his profits. Rather he is in a privileged position. Being a sole producer in the market he is able to fix any price for his product. He will be in a position to fix the price at a level where MC = MR and thus obtain maximum net profits. "The prima facie interest of the owner of a monopoly is clear to adjust the supply to the demand not in such a way that the price at which he can sell his commodity shall just cover his expenses but in such a way as to afford him the greater possible revenues" – Marshall. The monopolist thus keeps this aim in view while fixing the price of his product.

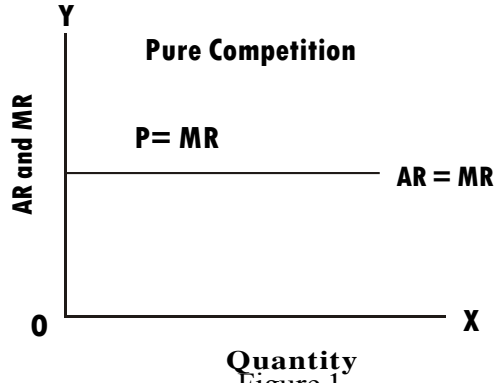


Figure 1

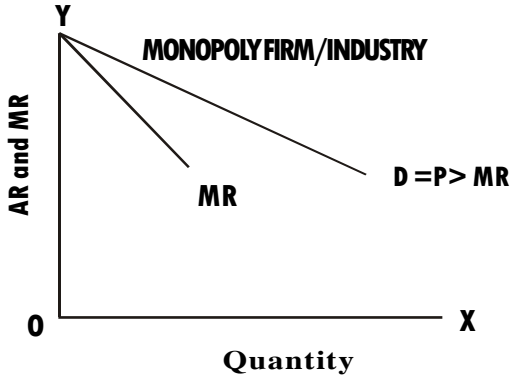


Figure 2

Fifthly, under monopoly, the single monopoly firm constitutes the whole industry. The demand curve for the monopolist's product is also market demand curve for the product since by definition the monopolist is the only supplier in the industry. As the demand of the consumer for a product slopes downwards, the monopolist faces a downward sloping demand curve. It thus implies that the monopolist can expand the demand for his product by lowering the price. The market demand curve facing the monopolist will be his average.
revenue curve since price is identical with average revenue. Unlike perfect competition, where average revenue curve is a horizontal straight line, average curve of the monopolist is downward sloping. It shows the sales he would be able to have at different prices. Since average revenue curve slopes downward throughout its length, marginal revenue curve will also slope downwards and lie below average revenue curve. It is because marginal revenue falls at a much faster rate than the average revenue.

It implies that whenever the monopolist sells a large quantity, the price of his product falls, hence marginal revenue must be less than the price.

The average revenue and marginal revenue are related to each other through elasticity of AR curve. MR can be known with the help of the following formula :

$$MR = AR \left(\frac{e - 1}{e} \right)$$

Here AR stands for average revenue and e for elasticity.
AR can also be known by the following formula :

$$\text{Price} \quad AR = MR \left(\frac{e}{e - 1} \right)$$

In other words price is equal to marginal revenue multiplied by $\frac{e}{e - 1}$

Since $\frac{e - 1}{e}$ will be less than unity, MR will be less than price. The extent to

which MR lies below AR depends upon the value of the fraction $\frac{e - 1}{e}$.

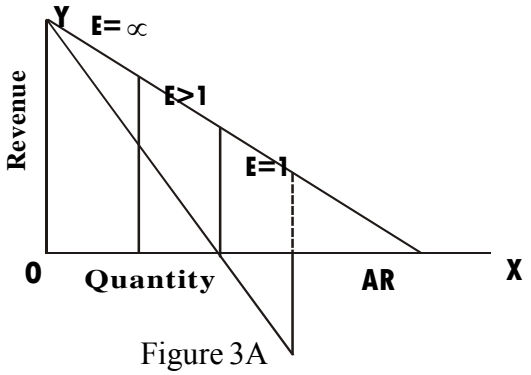


Figure 3A

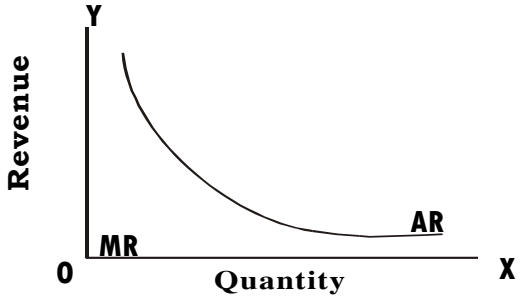


Figure 3B

It becomes evident from the diagram given above 3 (a) that where elasticity or AR curve is greater than unity, MR is always positive. Where it is equal to unity, MR is always zero. In case the elasticity of the AR curve is unity throughout its length like a rectangular hyperbola, the curve will coincide with X axis as shown with dotted line in figure 3 (b).

Where the elasticity of AR curve is less than unity, MR is negative and where elasticity of AR curve is zero, the gap between AR and MR curve becomes wider and MR curve lies much below the x-axis as is shown in figure 3 (A).

2.5.3.1 Equilibrium under monopoly

The monopolist wants to maximise his profits. How does a profit maximising monopolist select the level of output and prices at which to produce ? The features of cost curves under monopoly are generally the same as those under pure competition. There is a U-shaped average cost curve with a marginal cost curve that intersects it from below and passes through its lowest point AR and MR curve slopes downwards under monopoly. A profit maximising producer will produce that output at which marginal cost equals marginal revenue. The monopolist will fix that price at which the excess of gross receipts or revenue over costs is maximum. In the words of Mrs. Jon Robinson, "He will achieve this if he regulates his output in such a way that the addition to his total revenue from selling an additional unit is exactly equal to the addition to his costs caused by producing that unit. If he sold one unit more, he would lose more of revenue that he saved of cost and if he produces one unit less he would incur more of cost than he gained of revenue." In other words, he will always tend to equate marginal revenue with marginal cost. The price at which this quantity can be sold is seen from the demand curve. This is illustrated in the diagram given below :

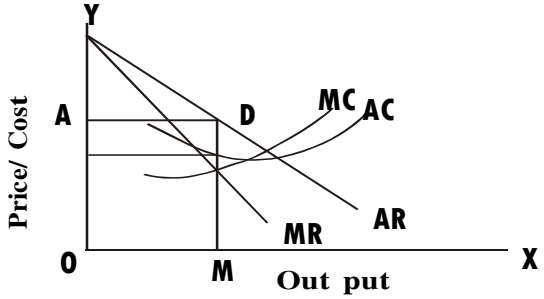


Figure 4 M.A. (Econ.) Paper I, Lesson No. 10

Here OM output represents profit maximising output. Price is represented by PM

Self-Check Exercise-I

Q. Describe the formula which shows the relationship between AR and MR?

Ans. .....

.....

.....

(or OA). ABCP rectangle represents the monopolist's profits. As is evident, price under monopoly is higher than marginal cost, because marginal cost is higher than marginal revenue.

The monopoly firm is earning supernormal profits in the short run (Fig. 4) The monopoly firm can continue earning supernormal profits even in the long run, provided it

is able to block the entry of other firms. The long-run equilibrium of a monopoly firm resembles the short-run, except that the plant size that is selected by the monopolist will be the one at which marginal revenue equals long-run marginal cost. The figure given below illustrates it.

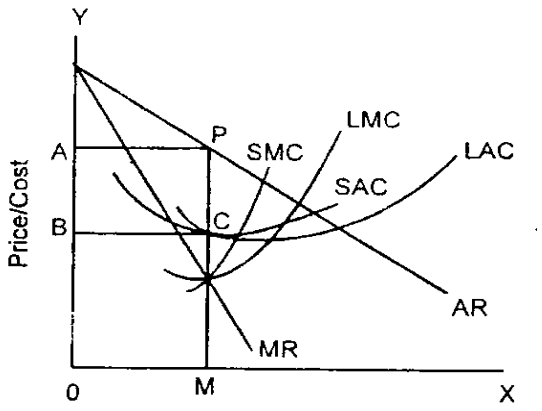


Fig. 5 M.A.(Econ.) Paper I, Lesson NO. 10

Rectangle ABCP represents the monopolist's supernormal profits in the long run. Price charged is PM (or OA) and equilibrium output is OM. This profit maximising output is decided at MR = LMC (Fig. 5).

Monopoly price is a function of the marginal cost of production and the elasticity of demand. The monopolist reaches equilibrium when MR = MC. This has four implications :

1. At equilibrium, elasticity of demand is always more than one. A profit maximising producer will always produce that level of output that falls within the elastic range of demand curve (positive marginal revenue (Fig. 3A). The equilibrium output is decided by MR=MC. Negative MC will imply that it costs less to produce a bigger quantity than smaller one. It will thus be profitable for the producer to reach his equilibrium only in that range where MR is positive or when elasticity of demand is greater than unity. However, in case of free goods, at equilibrium point, the elasticity of demand can be equal to one (where MR is zero). The general conclusion is that a monopolist would be in equilibrium only when the elasticity of demand for his product is greater than one.

2. Monopoly price is always greater than marginal cost.

(P > MC = MR).

$$\text{Price} = \text{MR}\left(\frac{e}{e-1}\right)$$

Since $\left(\frac{e}{e-1}\right)$ will be more than unity, Price will be higher than marginal cost.

3. Excess of monopoly price over marginal cost measures the degree of monopoly power. The degree of monopoly power is measured by

$$MR = \frac{P - MC}{P}$$

Where P stands for price and MC for marginal cost. The large the difference between marginal cost and price, the greater would be firm's monopoly power. If marginal cost is equal to zero, the monopoly power is absolute indicating the monopolist's ability to charge a price for a free good as in the case of mineral water spring. If on the other hand, marginal cost is equal to price, the monopoly power is zero. It thus indicated the competitive situation.

4. Under monopoly, equilibrium is possible with rising, constant as well as falling marginal and average cost curves. All that is needed is that at equilibrium point the marginal curve must cut the marginal revenue curve from below. It is unlike perfect competition where equilibrium position requires rising, marginal and average cost curves. The monopolist's equilibrium with rising, constant and falling marginal and price cost is illustrated in the figure 6 (a), 6 (b) and 6 (c).

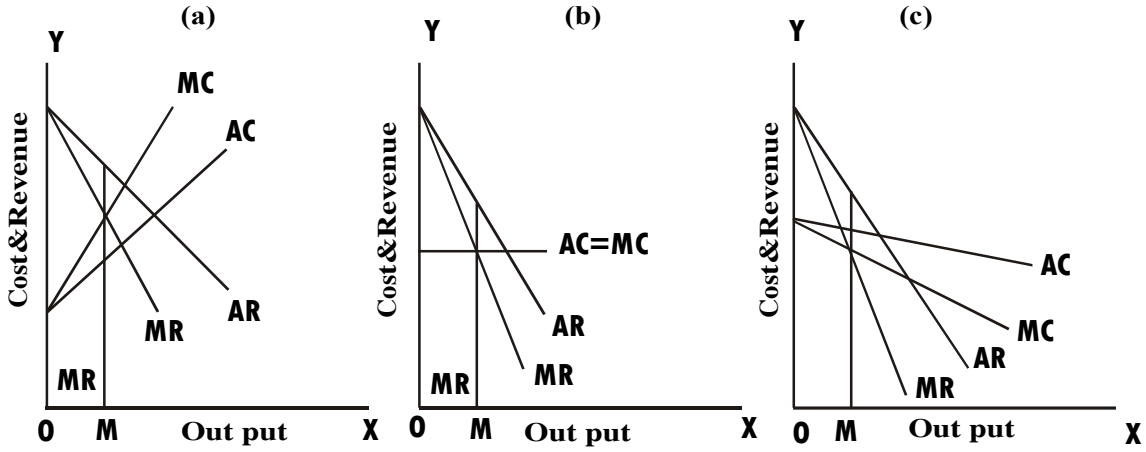


Figure 6

10.3.2 Price-Discriminating Monopoly

A monopolist fixes price of his product in such a way so as yield him the maximum net revenue and the price is fixed at the point where his marginal revenue equals marginal cost. Under simple monopoly, the monopolist charges one uniform price from all his customers for a single commodity. But this is not so under price-discriminating monopoly. Since the monopolist has a full control over the supply of a commodity he is in a position to charge different prices from different customers for it. It often happens that a monopolist finds it

possible and profitable to sell a single commodity at different prices to different buyers. This can occur when goods which are sold in the cheaper market cannot be brought from the monopolist and resold in the dearer market cannot transfer themselves into the cheaper market to get the benefit of the lower price. *The act of selling the same article under a single control, at different-prices to different buyers is known as price discrimination.*

"Under conditions of perfect competition price-discrimination could not exist even if the market could be easily divided into separate parts. In each section of the market the demand would be perfectly elastic, and every seller would prefer to sell his whole output in that section of the market in which he would obtain the highest price. The attempt to do so, of course, would drive the price down to the competitive level, and there would be only one price throughout the whole market."..... (John Robinson)

Price discriminating monopoly is likely to occur only when market is imperfect. The degree of price discrimination practised depends upon the degree of imperfection in the market. Price discrimination is more likely to occur when there is monopoly of the product by a single seller or group of sellers.

Conditions for Discrimination

Price discrimination becomes possible only when (i) no unit of demand can be transferred from one market to another and (ii) no unit of supply can be transferred from one market to another. These two conditions are two different sides of the same thing. Various examples can be given to elaborate the conditions further.

Discrimination can occur when the markets in which a monopolist is selling are divided from each other geographically. High transport costs may prevent resale of goods in a dearer market, bought from the cheaper market. Tariff barriers provide another example. They facilitate splitting of markets. A monopolist may be selling at a lower price in export market (in order to capture it) and at a higher price in home market (in order to make up the losses). This is known as dumping. This is made possible by high tariffs because goods cannot be bought from the foreign (cheaper) market and resold in home (dearer) market. Discrimination is most often practised in the sale of direct personal services, where there is no possibility of transfer from one market to the another. For example, doctor charges different fees from their patients according to level of income and wealth of their patients. Discrimination is again possible when the same service is required in connection with clearly differentiated products. For example, railways can charges different rates for the transport of cotton goods and of coal without any fear that sales of cotton will be turned into loads of coal in order to enjoy a cheaper rate. Legal sanction also facilitates price discrimination for instance different rates are charged for the use of electricity of industrial and lighting purpose. Discrmination may occur again when buyers are ignorant. For example, in case of goods sold on special orders, the individual buyers has no means of knowing what price is being charged from other buyers for a similar commodity. Another factor that facilitates price discrimination is product differentiation. Various brands of a certain article (which are almost alike) may be sold as

different qualities under different names at several different prices to different buyers. The monopolist is thus able to break up his market and sell the so-called superior varieties to the rich people at the higher prices and so-called inferior to the poor people. "In this way the market is split up, and the monopolist can sell what is substantially the same thing at several prices. The device of making the same thing appears in different guesses will also serve to save the monopolist from the reproaches of injustice between customers which some times put difficulties in the way of price discrimination." (John Robinson). Then there are those cases in which there is actual difference in the various varieties of the product and different prices are charged for different varieties. In such, cases price discrimination is said to be practised if extra price charged for the so-called superior varieties are not proportional to the extra cost incurred on them.

It is clear from above that price discrimination becomes possible only when no unit of the commodity sold in one market can be transferred to another, and when no buyer is able to transfer himself from dearer market to cheaper market to buy the commodity or service at the lower price. Price discrimination depends upon the ability of monopolist to split up and keep the market separate.

When is Price Discrimination Profitable ?

We have seen above the conditions which make price discrimination possible. It is quite likely that the monopolist may be able to charge different price in separate markets, yet it is not profitable for him to do so. Price discrimination becomes profitable only when elasticity of demand for the product is different in different markets at the single monopoly price. If elasticity of demand is the same in both markets it would not pay the monopolist to charge different prices in these markets, marginal revenue in the two markets at every price of the product will also be the same. It is so because:

$$MR = AR\left(\frac{e-1}{e}\right)$$

When marginal revenues at every price of product are the same in both markets, it will not pay the monopolist to transfer any quantity of the product from one market to another and charge different prices for the same product in these markets. When elasticity of demand is different in both markets at the same monopoly price, only then it is to the advantage of the monopolist to charge different prices in different markets. It is then in this interest that he must discriminate in prices if he wants to maximise his profits. If he does not do so and charge a single monopoly price (on the basis of aggregate marginal revenue and marginal cost of the output) in both the markets where elasticities of demand are different, he is not maximising his profits. "For if he charges the same price in each market he will find that, at that price, the marginal revenue obtained by selling an increment of output in each market separately is greater in some markets than in others. He can therefore increase his profit by selling less in those markets where elasticity of demand is less and the marginal revenue smaller, and selling more in those markets where

elasticity of demand is higher and the marginal revenue greater.

Now the question is how long will the monopolist go on transferring the output from the market with low elasticity of demand to the market with high elasticity of demand ? He will go on being so till marginal revenues in both markets are equalized. To quote Mrs. Joan Robinson, the monopolist will "adjust his sales in such a way that the marginal revenue obtained from selling an additional unit of output in any one market is the same for all markets. And his profits will be at a maximum when the marginal revenue in each market is equal to the marginal cost of the whole output." It follows from above that when marginal revenues are equalized in two markets (having different elasticities of demand at the single monopoly price) prices charged would be different. The monopolist would charge a higher price in market with low elasticity of demand and a low price in market with high elasticity of demand.

Conditions for Equilibrium under Perfect Discriminating Monopoly

Under simple monopoly a producer reaches his equilibrium when his marginal revenue equals his marginal cost of output. A price discriminating monopolist has to sell in different markets, having different elasticities of demand. He is to take two decisions : (i) how much total output should be produced and (ii) how much output should be sold in each market, and at what price should be sell in each market? The answer is provided by Mrs. Joan Robinson, "The monopoly output under price discrimination is determined by the intersection of the monopolist's marginal cost curve with the aggregate marginal revenue curve. This total output is made up of the amount sold in the two markets, in each of which marginal revenue is equal to the marginal cost of the whole output. The price in each market will be the demand price for the amount of output sold there."

The price discriminating monopolist will thus produce that level of output at which aggregate marginal revenue equals marginal cost of the total output. He will consider marginal cost of the whole output. In other words, price discriminating monopolist will reach his equilibrium when:

(i) Aggregate marginal Revenue (AMR) = Marginal Cost (MC)

(ii) $MR_1 = MR_2 = MC$

Degrees of Price discrimination

Price discrimination can be classified according to the extent to which it is practised. A monopolist may practice three categories of price discrimination:

First Degree Price discrimination

Second Degree Price Discrimination

Third Degree Price Discrimination

Self-Check Exercise-II

Q. Give the equilibrium conditions of price discriminating monopoly?

Ans. .....

.....

.....

Prof. A. C. Pigou in his book, *The Economics of Welfare* has given a good account of the degrees of price discrimination. While explaining them he says : "A first degree would involve the charge of a different price against all different units of commodity, in such a way that the price expected for each was equal to the demand price for it, and no consumer's surplus was left to the buyers. A second degree would be obtained if a monopolist were able to make separate prices, in such a way that all units with a demand price greater than x were sold at price x and greater than y at price y, and so on. A third degree would be obtained if the monopolist were able to distinguish among his customers, in different groups, separated from one another more or less by some practicable mark, and could charge a separate monopoly price to the members of each group. This degree, it will be noticed, differs fundamentally from either of the preceding degrees, in that it may involved the refusal to satisfy in one market, demand represented by demand prices excess of those which in another market, are satisfied."

We now discuss them in more detail.

(i) First Degree Price Discrimination

In first degree discrimination, the monopolist charges from a consumer a different price for each unit of the good. He charges each buyer the highest price that he will pay for each unit of the good he gets. The monopolist thus leaves no consumer surplus with the consumer.

As illustrated in Fig. 7, the monopolist has sold three units of the commodity at three different prices P_2 M_2, P_1 M_1, PM. He has not sold all three units at the uniform price of PM which he would have charged under simple monopoly. He has obtained from each buyer the higher price which he was willing to pay rather than go without the thing, "Under the first degree price discrimination, the monopolists realizes the entire utility obtained by consumers and leaves no consumer surplus with them. Again, under the first degree price discrimination, average revenue and marginal revenues curves coincide with each other as demand curve of the buyer becomes marginal revenue curve for the seller. This is unlike simple monopoly where marginal revenue curve is rising upward for the seller. This is unlike simple monopoly where marginal revenue lies below average revenues curve.

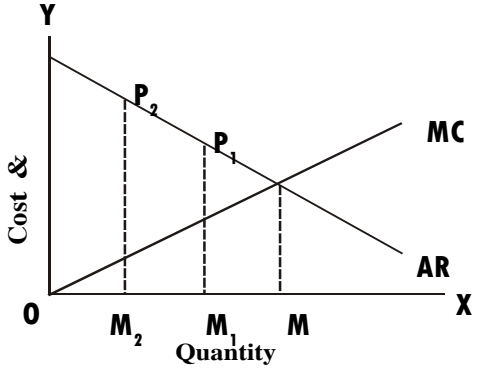


Figure 7 M.A. (Econ.) Paper I, Lesson No. 10

(ii) Second degree Price Discrimination

In second degree price discrimination, buyers are divided into different groups or classes and from each group or class a different price is charged which is the lowest demand price of that group or class. In other words, the monopolist charges from each group that price which a marginal individual of that group is just willing to pay. According to Prof. Pigou, the monopolist will charge separate price in such a way that all units of the commodity with a demand price greater than, say x, are sold at a price x, all units with a demand price greater than y but less than x at price y and so on. Thus in Fig. 8 OM block of units is sold at PM price. M^{th} unit has a demand price PM and earlier units have demand a price greater than PM as shown by the DP position of demand curve DD. Under second degree price discrimination OM units are sold at PM price. Thus consumer enjoys no consumer's surplus

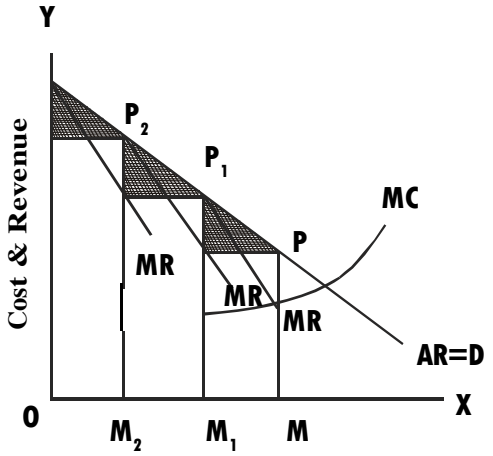


Figure 8 M.A. (Econ.) Paper I, Lesson No. 10

on Mth Unit, but he enjoys consumer surplus on earlier units as shown by the shaded area, Similarly, MM_1, block of units are sold at P_1 M_1 price, and M_1 M_2 block of units at P_2 M_2 price. Consumer enjoy surplus satisfaction to the extent of shaded areas shown in these respective block units.

(iii) Third Degree Price Discrimination (Two Cases)

The third degree price discrimination is the most common form of price discrimination. It is practised when the monopolist can divide his buyers into two or more markets and charge

different prices in different markets or sub-markets. This becomes possible when the elasticity of demand in different sub-markets is different and when each market is isolated from other market so that no consumer buying in the low prices market can re-sell the product in the high price market.

First Case

The monopolist divides his total market into two sub-markets on the basis of elasticity of demand. Suppose he divides his total market into sub-market I and market II taking elasticity of demand into consideration. As illustrated, in Fig., 9 D_1 and D_2 are demand curves of two markets. They both slope downwards to the right. MR_1 and MR_2 are the corresponding marginal revenue curves AMR is aggregate marginal revenues which has been derived by adding up laterally MR_1 at MR_2.

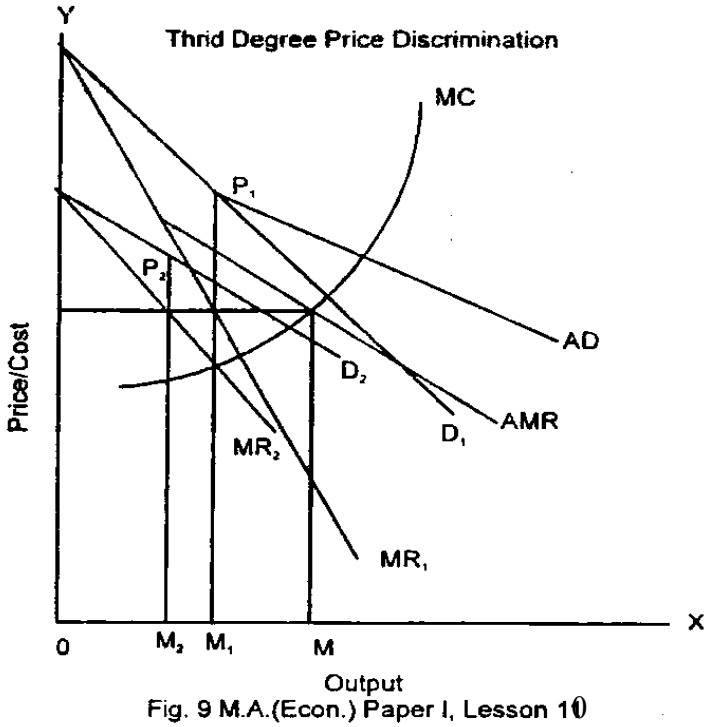


Fig. 9 M.A.(Econ.) Paper I, Lesson 10

AD is aggregate demand curve which is derived by lateral summation of D_1 and D_2 curves. Aggregate demand curve shows the total amount that would be sold at each price if the prices were the same in both markets, and aggregate marginal revenue show the amount of sales that would correspond to each value of the marginal revenues if the marginal revenues were the same in both markets. This curve will show the marginal revenues obtained by the discriminating monopolist. The monopolist will maximize his profits by producing that level of output at which marginal cost curve intersects the aggregate marginal revenue curve. This is represented by OM output in the figure. Thus total output OM is to be sold in two markets in such a way that marginal revenues in two markets are equal to each other and also to marginal cost of the whole output. OM in this figure is the total

output and equal to $OM_1 + OM_2$. MC is the marginal cost of the output OM. OM_1 is sold at the price $M_1 P_1$ in market 1, and OM_2 is sold at the price $M_2 P_2$ in market II. The area lying under aggregate marginal revenues curve (total revenues) minus the area lying under the marginal cost curve (total cost), shows the monopolist's profit.

Second Case : Equilibrium in the Dumping Case

A special case of price discrimination occurs when a producer is selling in two markets, one of which is perfectly competitive, while in the other he has a monopoly. This can often be noticed in international markets. Such situation arises when a producer sells his product in his home country where he has a monopoly and also in the world market which is perfectly competitive. In the home market, he confronts a downward slopping demand curve and in the world market, he faces a perfectly elastic demand curve. (fig.10)

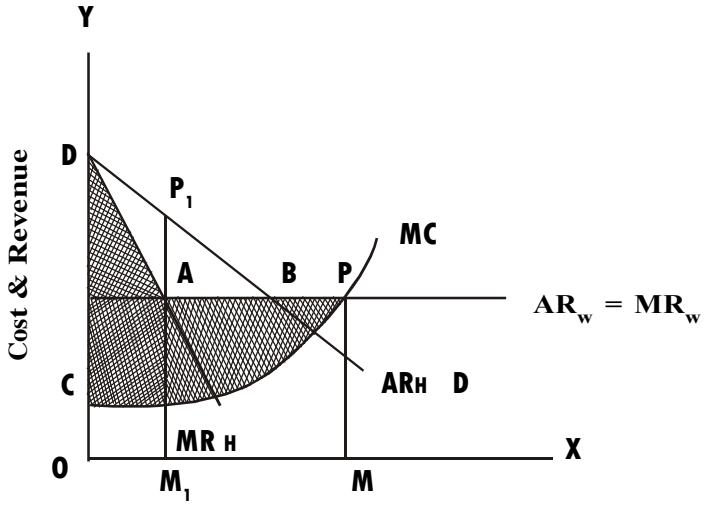


Figure 10 M.A. (Econ.) Paper I, Lesson No. 10

Since the producer faces perfect competition in the world market the average revenues and marginal revenues curve are indicated by horizontal straight line $AR_w = MR_w$, AR_H and MR_H represent average and marginal revenues curve in the home market. MC is the marginal cost curve of output. Equilibrium level of output is determined when MC intersects aggregate marginal revenue curve DABP at P. OM is the total output sold. This output is to be distributed in both markets in such a way that marginal revenues in both markets are equal to each other and also to marginal cost i.e. PM. Thus OM is sold in home market at $P_1 M_1$ price, and MM amount of output is sold in world market at PM price. As is evident from the figure 10, the producer is charging a lower price in the perfectly competitive world market and a higher price in the home market. He is said to be dumping in the world market. He is earning profit equal to the area CPAD.

It can be concluded that existence of price discrimination depends on the

difference between the elasticities in the market in which it is possible to sell. Total market is divided into sub-markets by monopolist on the basis of elasticity of demand. Highest price is charged in the least of elastic market and least price is charged in the most elastice market. Price charged thus depends upon the degree of elastic of demand.

Is Price Discrimination Desirable?

(a) We are now to examine how far price discrimination is harmful or advantageous to the customers of monopolist and to the society as a whole. According to G. J. Stigler, "Economists are considerably divided on the best pricing policy for an industry whose existence required discrimination." It is generally agreed that consumers are better off if a single price is charged for the same commodity. But various cases can be cited in which price discrimination is beneficial. Some commodities and services might not be produced at all if sellers were not able or were not allowed to practise price discrimination. The important question then is : is discrimination objectionable when the supply of the commodity will be smaller without discrimination or the supply completely stops. Let us take the cases of railway. Railways cannot operate if 'Cost of Service' principle is adopted. They can work only if permitted to charge according to the principle of what the traffic will bear, i. e. charge higher freight rates for materials that have a higher value per tonne than for materials for lower value. Similarly, higher rates are charged from customers of electricity for lighting purpose than from customers of electricity for industrial production. Such situations may be found in which costs of production cannot be covered by receipts unless price discrimination is permitted. Commenting on the desirability of price discrimination, Mrs. Joan Robinson has observed, "Since average revenue is greater under price discrimination than under simple monopoly, that there may be cases in which no output would be produced at all if price discrimination were not possible. If the average cost curve of a certain product lay above the demand curve for it throughout its length no profit could be made by producing it under any one price system. But if the average cost curve though above the demand curve, lay at some point below the average revenues curve under price discrimination, a profit could be made and some output would be produced provided that discrimination was possible. It may happen, for instance, that a railway would not be built or a country doctor would not set up in practice, if discrimination were forbidden. It is clearly desirable that price discrimination should be permitted in such cases, for the average revenues of the monopolist cannot be greater than average utility to the consumers. If average revenue is greater than average cost, average utility will also be greater, and the investment will lead to a gain to society."

The following illustration makes it clear, in the figure, the line D is the ordinary demand

(AR), showing amount that can be sold at prices that are same to all buyers. A curve shows the average price received when different prices are charged. Thus the price P_3 is the average of prices P_1 and P_2 . AP curve is thus constructed in the same way. As is shown in the figure, AC curve lies above D but below AP. No output would be produced if the firm

has to sell at single price on D curve. But there is output with the price discrimination since the firm is able to cover its cost. Thus the output or a commodity or service becomes profitable under discriminating monopoly which could not be produced at all under simple monopoly.

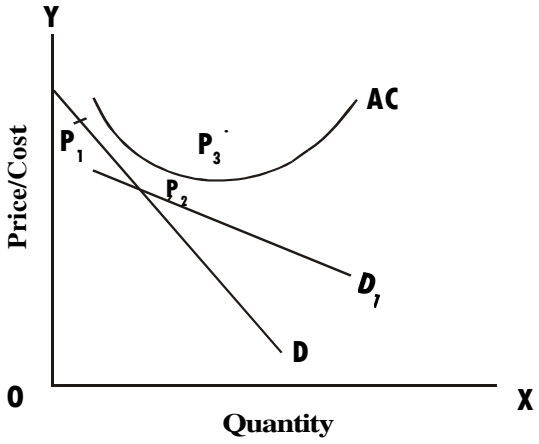


Figure11: M.A. 1 (Economics) Paper-I (L.No.10)

(b) Price discrimination is socially desirable when it makes possible for weaker sections of a society to have access to certain essential good or service (like that of a doctor) which would otherwise not be possible under simple monopoly. Again in many situations, weaker markets (with lower prices) are served and offered output at lower prices, while the same is refused in stronger markets (with higher demand prices).

(c) When price discrimination is practised, some sections of the society are affected adversely who are required to pay higher prices (this is especially true in the case of a good which is produced under law of increasing cost). But if higher price is charged from richer sections and lower price from poorer sections, it helps to mitigate inequalities in the society. It benefits the society as a whole. But if the good is produced under the law of decreasing costs or constant costs, it is quite likely that no body is affected adversely. The enlarged output of the product under discriminating monopoly becomes available to all customers at comparatively lower prices as compared to the price under simple monopoly.

(d) Will the output of a product under price discrimination be larger or smaller than or equal to the output produced under simple monopoly? The answer depends upon many factors, "if the monopolist were ever able to practise perfect discrimination, he would have an output as large as the purely competitive output with same demand and cost conditions. This is because the monopolist treats, the consumer's demand curve as his own marginal revenue curve. He equates this marginal revenue with his marginal cost. Thus he equates demand with marginal cost as in the equilibrium of pure competition." If we consider simple monopoly versus discriminating monopoly then output of the product under first degree price discrimination will be twice as large as output under simple monopoly, if the production is subject to constant cost (fig. 12a). But if the output of the product is subject to law of decreas-

ing cost or increasing returns, then output of the product under first degree price discrimination would be more than twice as large as output under simple monopoly (fig, 12b in the first case) and less than twice as large as output under simple monopoly (Fig. 12c in the second case). The figure given above illustrate this. Discrimination of the second degree tends to approach first degree price discrimination with an increase in the number of prices. Outcome thus tends to be as under first degree price discrimination. It follows that output is larger under second degree price discrimination than if the monopolist had a single price.

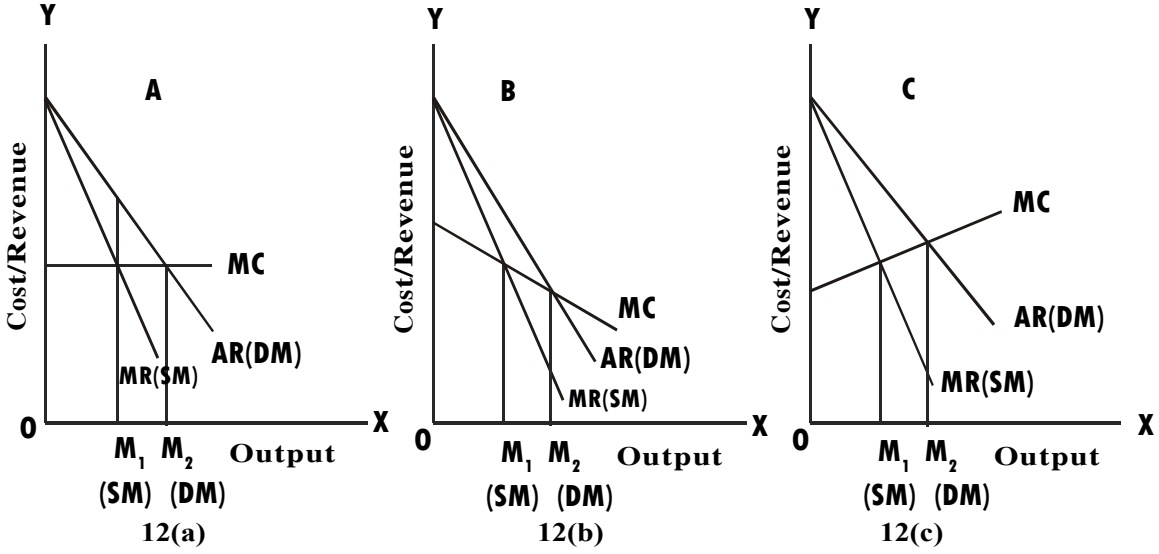


Figure 12

When a monopolist practises third degree price discrimination, his output can be equal to, or less than, or greater than his output at a simple monopoly price. It depends on the shape of demand curves in the monopolist's two markets. If the demand curves in the seperate markets are linear, then total output is the same as with a single price. With two markets, this means that the reduction in the market with the less elastic demand is equal to the expansion in the market with the more elastic demand. But when demand curves are concave and more elastic demand curve is more concave than the less elastic demand, the output will be increased by the introduction of price discrimination. On the other hand, when more elastic demand curve is less concave than the less elastic demand curve, the total output will be less under price discrimination than under simple monopoly.

However, Mrs. Joan Robinson in her authoritative discussion on price discrimination, comes to the conclusion that, in all likelihood, output is large with price discrimination than without it. Sometimes, output is possible only with price discrimination.

Long Question

Q. Explain about Equlibrium under Monopoly?

Q. Explain about First Degree Price Discriminition?

Q. Explain about Third Degree Price Discrimination?

Short Question

Q. Conditions for Discrimination in Monopoly?

Q. Conditions for Equlibrium under Perfect Discriminating Monopoly.