PREFACE

In the curricular structure introduced by this University for students of Post-Graduate degree programme, the opportunity to pursue Post-Graduate course in a subject introduced by this University is equally available to all learners. Instead of being guided by any presumption about ability level, it would perhaps stand to reason if receptivity of a learner is judged in the course of the learning process. That would be entirely in keeping with the objectives of open education which does not believe in artificial differentiation. I am happy to note that university has been recently accredited by National Assessment and Accreditation Council of India (NAAC) with grade 'A'.

Keeping this in view, study materials of the Post-Graduate level in different subjects are being prepared on the basis of a well laid-out syllabus. The course structure combines the best elements in the approved syllabi of Central and State Universities in respective subjects. It has been so designed as to be upgradable with the addition of new information as well as results of fresh thinking and analysis.

The accepted methodology of distance education has been followed in the preparation of these study materials. Co-operation in every form of experienced scholars is indispensable for a work of this kind. We, therefore, owe an enormous debt of gratitude to everyone whose tireless efforts went into the writing, editing, and devising of a proper layout of the materials. Practically speaking, their role amounts to an involvement in 'invisible teaching'. For, whoever makes use of these study materials would virtually derive the benefit of learning under their collective care without each being seen by the other.

The more a learner would seriously pursue these study materials the easier it will be for him or her to reach out to larger horizons of a subject. Care has also been taken to make the language lucid and presentation attractive so that they may be rated as quality self-learning materials. If anything remains still obscure or difficult to follow, arrangements are there to come to terms with them through the counselling sessions regularly available at the network of study centres set up by the University.

Needless to add, a great deal of these efforts are still experimental— in fact, pioneering in certain areas. Naturally, there is every possibility of some lapse or deficiency here and there. However, these do admit of rectification and further improvement in due course. On the whole, therefore, these study materials are expected to evoke wider appreciation the more they receive serious attention of all concerned.

Professor (Dr.) Subha Sankar Sarkar
Vice-Chancellor

Netaji Subhas Open University

Post Graduate Degree Programme

Subject : Commerce (M. Com)

Course : Managerial Economics

Course Code : PGCO - II

First Print : March, 2022

Printed in accordance with the regulations of the Distance Education Bureau of the University Grants Commission.

Netaji Subhas Open University

Post Graduate Degree Programme

Subject : Commerce (M. Com)

Course : Managerial Economics

Course Code : PGCO - II

: Board of Studies :

Members

: Course Writer :

Professor Dhirendra Nath Konar
Retd. Professor of Economics
Kalyani University

: Course Editor :

Professor Biswajit Chatterjee
Professor of Economics
Netaji Subhas Open University

: Format Editing :

Professor Anirban Ghosh
Netaji Subhas Open University

Netaji Subhas Open University

PG : Commerce
(M. Com)
New Syllabus

Course : Managerial Economics
Code : PGCO - II

Module - I

Module - II

❑

Unit 1 ❐ The Nature and Scope of Managerial Economics

Structure

1.0 Objectives

The objectives of this unit are—

- ❏ To know about Managerial economics, its scope and nature;
- ❏ To highlight the importance of microeconomic theory of the firm and business units;
- ❏ To study demand analysis, production and cost analysis, market behavior and pricing of products;
- ❏ To examine theory of capital and investment decisions.

1.1. Introduction

It is known to most of the students of Economics that the study of Economics as a subject dates back to 1776 when Adam Smith, the Father of Economics, published the famous title “The Wealth of Nations” in two volumes. Before this, the study of the subject had been carried out in ancient Greece where Economics was taken as the science of household management. However, the study of Managerial Economics had been first initiated in America immediately after the publication of the pioneering title “Managerial Economics” in 1951 by Joel Dean. During the last seven decades or so this subject has, increasingly, been popular and at the same time the book market all over the world has been flooded with a huge number of titles on this very subject.

The emergence of Managerial Economics as a separate discipline of management studies can be contributed to at least three factors : (i) Growing complexity of business decision-making process because of changing market conditions and environment of business, (ii) consequently, the increasing use of economic logic, theories, concepts and tools of economic analysis in the process of business decision-making and (iii) Finally, vast increase in demand for professionally trained managerial manpower.

The growing complexity of the business world has made the business decision-making process highly complex. About one hundred years ago most business units had been set up, owned and managed by individuals or business firms. Industries of large size had been few and scale of business operation was relatively small. The managerial capacity acquired through traditional family training and experience had, largely, been sufficient to manage small and at most medium scale business. It is a fact that a large part of private business is, still, run on small scale and managed in the traditional style of business management. However, there has been dramatic change in size, nature and content of the industrial business world. The highly growing complexity of the business world can, duly, be attributed to the growth of large scale industries, diversification of industrial products, growth of multinational corporations and mergers and take overs, specially after the Second World War. All these factors have contributed greatly to an increase in inter-firm, inter-industry and international rivalry, competition, risk and also uncertainty. Business decision-making in this kind of business environment is a highly complex matter. Naturally, the family training and experience are not at all sufficient to face such managerial challenges.

The ever-growing complexity of business decision making has, obviously, enhanced the application of economic concepts, theories and tools of economic analysis in this area. The basic reason is that making an appropriate business decision needs a clear understanding of the business environment, market conditions and market fundamentals. For this an intensive and extensive analysis of the market conditions in the product, input and financial markets is highly needed. At the same time, economic theories, logic and tools of analysis have been developed both to analyse and predict market behaviour. The application of economic concepts, theories, logic and analytical tools both for the assessment and prediction of market conditions and business environment has proved to be of vital help in business decision-making. In effect, economic theories and analytical tools that are highly used in business decision-making have crystallized into a different growth of management studies, very popularly called “Business Economics” or “Managerial Economics”.

1.2 Managerial Economics

From the above analysis we may say that Managerial Economis is constituted of economic theories and analytical tools that are widely applied to business decision-making. Naturally, we should know something about Economics. Economics is a social science whose basic function is to study how people maximize their gains from their limited resources and opportunities. In economic terminology this is known as optimizing behaviour. This behaviour is selecting the best out of available options with the basic objective of maximizing gains from the given resources. According to Lionel Robbins, "Economics is the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses". For example, Economics studies how households allocate their limited resources among various goods and services they seek to consume so that they are able to maximize their total satisfaction. Naturally, Economics is a study of the choice-making behaviour of the people. However, choice-making is not so simple as it appears to be as the economic world is highly complex and most economic decisions have to be taken under the condition of risk, uncertainty and imperfect knowledge. Now economic laws and tools of economic analysis are applied greatly in the process of business decision-making. This has led to the emergence of a separate branch of study known as "Managerial Economics".

Managerial Economics can, therefore, be defined as that branch of economic theory which analyses and discusses the problems of the business matters.

From the above we may say that Managerial Economics aims at increasing the manager's ability to understand and also to evaluate business problems and, therefore, help him in formulating right decisions. That means, Managerial Economics deals with the application of economic theory to business management, it is all economic aspects of managerial decisions. It is because of this reason that Managerial Economics is often called "Economics of Business Decisions". Accoring to G.P.E. Clarkson, "It is concerned with all decisions that have an economic implication irrespective of the level within the enterprise at which they are made. For management is a decision-making activity and to manage well is to control the decision-process so that the resulting behaviour is in keeping with stated objectives". The emphasis is highly placed on intelligent application of quantitative techniques to business.

1.3 Nature of Managerial Economics

After the definition of Managerial Economics let us concentrate our attention on the nature of this Economics.

1. Managerial Economics is primarily, micro economics in nature. Basically, there are two parts of Economics : (a) Micro Economics and (b) Macro Economics. The economics which is concerned with the whole economy is known as macro economics. Thus aggregate demand, general price level, business cycles etc. are within the purview of Macroeconomics. On the other hand, Microeconomics deals with smaller parts of an economy. For example, demand for a particular commodity, price of a commodity, profit of an individual firm are within the jurisdiction of Microeconomics. Managerial Economics, naturally, falls within the purview of Microeconomics since its basic concern is with the problems of the business firms.

2. Managerial Economics is not positive but normative. Economic theory has two divisions : (a) Positive Economics and (b) Normative Economics. In the former we concentrate on two things : there is a straight forward description, either of a part of the economy or of the whole economy. Besides, positive economic theory tries to explain the behaviour of either an economic unit or of economic aggregates. In Normative Economics, however, the economist says what he thinks should happen rather than what actually does happen. Managerial Economics is normative as it deals with what the management should do. Managerial Economics describes the goals of the entrepreneur and then prescribes how to attain these goals. That means, Managerial Economics is not descriptive but prescriptive. This Economics can better be called normative Microeconomic theory of the firms.

3. Students of higher Economics know that in pure Microeconomic theory they are to study theory on the basis of some assumptions most of which are abstract. Naturally, the analysis becomes far from reality. However, Managerial Economics is free from difficult abstract issues of economic theory and it concentrates, basically, on practical problems faced by the firms.

4. Together with Microeconomics, Macroeconomics is also highly useful in the study of Managerial Economics. This is so because this subject highlights an understanding of the environment where the firms have to operate. Some important topics of Macroeconomics, namely economic policies of the Government with regard to taxation, anti-monopoly measures, labour relations and the like affect the decisions of the firms.

Expectedly, the firm has no control over them and these are assumed to be externally given. However, the subject "Managerial Economics" deals with how the firm should adjust its various activities to cope with these important external Macroeconomic situations.

5. Last but not the least, though Managerial Economics is basically concerned with the theory of the firm, the entire demand analysis, specially, the concept of elasticity of demand is a very vital and practical topic of Managerial Economics.

1.4 The Scope of Managerial Economics

It is known to us that Economics has two major branches : (a) Microeconomics and (b) Macroeconomics. Both Microeconomics and Macroeconomics have been applied to business analysis and decision making, directly or indirectly. In general, the scope of Managerial Economics comprehends all those economic concepts, theories and tools of analysis which can be used to analyse the business environment and to find solutions to practical business problems. That means, Managerial Economics is Economics applied to the business problems and decision-making. Broadly speaking, we may say that Managerial Economics is Applied Economics.

Managerial Economics, though a part of economic theory, includes, basically, Microeconomics. What more, within Microeconomics, Managerial Economics is, basically, concerned only with those topics connected with the firm. Whatmore, the subject matter and scope of this subject has been increasing day-by-day. In spite of that we can say that the scope of Managerial Economics covers the following topics :

(a) Demand Analysis, specially elasticity of demand.

(b) Production Analysis

(c) Cost Analysis

(d) Market (Product) Analysis and Pricing Theory.

(e) Profit Analysis and Profit Managemet.

(f) Theory of Capital and Investment Decisions.

Of late, however, there has been a trend towards the integration of Managerial Economics and Operations Research. Naturally, techniques like Linear Programming,

Decision Theory, Inventory Models, Theory of Games have also come to be regarded as a very vital part of Managerial Economics. Let us now explain the above points, in short, one by one.

Theory of Deamand : Demand theory explains the consumers' behaviour. It answers the question : How do the consumers decide whether or not to purchase a commodity. How do they decide on the quantity of a commodity to be purchased? At what level of demand, does changing price become inconsequential in terms of total revenue? Therefore, the knowledge of demand theory can be helpful in the choice of commodities for production.

In the demand theory there is a topic called demand forecasting which is highly important for the management as it will decide the production pattern keeping an eye to the demand for its products. Besides, the management should have some knowledge of the factors affecting demand for a commodity. This is needed because the management by changing the variables at its control can affect the level of demand. Again, it is known that the advertisement expenditure made by firms for a product is a very important determinant of its demand. Therefore, the level of demand can be affected by making a change in the advertisement expenditure.

For the efficient organisation of the production process some knowledge of production and cost analysis is vital. For project planning, idea of production and cost analysis is vital. To obtain some level of profit some amount of output has to be produced and at the same time some amount of cost has to be incurred. The management is to decide how to obtain a given level of output at the minimum cost or how to obtain maximum output with a given cost.

Profit management is an important area of Managerial Economics. The difference between total revenue and total cost is known as profit. One basic aim of firms is to make maximum profits and the amount of profit enjoyed by the firm is a clear indicator of the degree of efficiency of a firm. Generally, the more efficient firms get more profit while less efficient firms get less profit.

A very important area of Managerial Economics is pricing because the price policy followed by a firm influences the demand conditions and the revenue earned by the firm. Prices are determined under different markets in different ways. However, pricing is only an important aspect of market strategy. In a modern economy, however, the consumers want new and satisfying experiences. In keeping with the expectation of the consumers

the firms introduce new products and services and try to bring superior variety of older products and thereby a business manager widens the range of choices open to the consumers in the expectation of expanding sales and revenues. Obviously, every successful business must have a product policy and keep its product line under constant review in accordance with the change in demand. Besides, every business should have a clear policy on promoting its own products. All these aspects indicate that product policy, market strategy and sales promotion are some of the vital aspects with which managers must be familiar and all these should be included in Managerial Economics.

1.5 Sample Questions

A. Objective-type Questions :

Choose the Correct alternative :

1. The Father of Economics is : (a) J. M. Keynes, (b) Adam Smith, (c) Alfred Marshall, (d) P. A. Samuelson.

Ans. (b).

2. The study of Managerial Economics had been first initiated by : (a) J. R. Hicks, (b) Amartya Sen, (c) Robert Solow, (d) Joel Dean.

Ans. (d).

3. A part of the subject matter of Managerial Economics is: (a) Determination of employment, (b) Determination of factor prices, (c) Demand analysis, (d) Interest rate determination.

Ans. (c).

4. How prices of factors of production are determined is studied in : (a) Microeconomic Theory, (b) Managerial Economics, (c) Macroeconomic Theory, (d) Economics of Public Finance.

Ans. (a)

5. Demand forecasting is studied in: (a) Only Microeconomics, (b) Only Managerial Economics, (c) Only Macroeconomics and (d) in both Microeconomics and Managerial economics.

Ans. (d).

6. Mathematical tools are conveniently applied in : (a) Only Managerial Economics, (b) Only Macroeconomics, (c) Microeconomics and Managerial Economics, (d) Only Microeconomics.

Ans. (c).

7. Theory of employment is a subject matter of : (a) Microeconomics, (b) Macroeconomics, (c) Managerial Economics, (d) Managerial Economics and Microeconomics.

Ans. (b).

B. Short-type Questions

(a) Give a definition of Managerial Economics.

(b) "Managerial Ecoomics has emerged from Microeconomics.'' Do you agree with this view? Give reasons in favour of your answer.

(c) Explain the role of Mathematices in Managerial Economics.

(d) Discuss, in brief, the subject matter of Managerial Economics.

(e) Write a brief note on the nature of Managerial Economics.

C. Essay-type Questions

1. Define Managerial Economics. Explain the nature and scope of Managerial Economics.
2. "Managerial Economics bridges the gap between economic theory and business practice." Explain with examples.
3. "Managerial Economics is applied Microeconomics" Elucidate.
4. Explain the role and responsibilities of a Managerial Economist.
5. "Managerial Economics is the integration of economic theory in the business practice for the purpose of facilitating decision making and forward planning by management." Discuss.
6. How does the study of Managerial Economics help a business manager in decision-making? Illustrate you answer with examples from production and pricing issues.
7. "Managerial Economics is essentially the application of Microeconomic theory of business decision making." Discuss the statement.

Unit 2 ❐ Demand Analysis

Structure

2.0 Objectives

The objectives of this unit are—

- To know about Law of Demand, determinants of demand, and the market demand schedule;
- To know the exceptions to the law of demand- Giffen goods, Bandwagon, Snob and Veblen effects;
- To understand various concepts of elasticity of demand and their roles in demand analysis, and the factors on whom these elasticities of demand depend;
- To understand demand forecasting, the different methods and estimation of demand functions.

2.1 Introduction

A very important topic of Managerial Economics is “Demand Analysis”. We are going to discuss it, in brief.

2.2 Meaning of Demand

The concept “demand” in Economics refers to the quantity of a good or service that consumers are willing and able to purchase at various prices during a period of time given his/her income. It is to be remembered that demand in Economics is something more than the desire to purchase though desire is one element of it. A beggar, for instance, may desire food, but due to lack of means to purchase it, his demand is not effective. Thus effective demand for a thing depends on : (a) desire, (b) means to purchase and (c) willingness to use those means for that purchase. Unless demand is backed by purchasing power or

ability to pay, it does not constitute demand.

Two things are to be noted about quantity demanded. One is that quantity demanded is always expressed at a given price. At different prices different quantites of a commodity are generally demanded. Secondly, quantity demanded is a flow. We are concerned not with a single or isolated purchase, but with a continuous flow of purchases and we must, therefore, express demand as so much per period of time : one hundred kgs. of apples per day, one thousand kgs. of apples per week and so on.

2.3 Determinants of Demand

There are a number of factors which influence household demand for a commodity. Important among these are : (a) price of the commodity itself, (b) income of the consumers, (c) prices of related commodities, (d) tastes and preferences of consumers, (e) consumers' expectation, (f) size and composition of population, (g) distribution of income and (h) other factors.

The above-listed factors can, easily, be presented in the form of a demand function as follows :

$$Q_{dc} = f\ (P_c,\ Y,\ P_R,\ T,\ E,\ D)$$

where Q_{dc} is the quantity demanded for a commodity c, P_c is the price of the commodity itself, Y is the money income, P_R is the price of related commodities, T is the taste of the household, E is the expectation and D representes the size of the population and other remaining factors.

The important determinants of demand are now being discussed :

(a) Price of the commodity : The quantity demanded for a commodity is, generally, inversely related to the price of that commodity. From the day-to-day general behaviour of consumers, it is observed that the price of a commodity is the most important factor for determining its demand. If price of a commodity increases, the demand for that commodity will decrease and vice-versa. That means, demand for a commodity is negativety related to the price of that commodity. However, in reality, exceptions to this negative relationship are also observed. We may cite here that in case of Giffen goods the law of demand is positively related to the price of the good.

(b) Income of the Consumer : Income of a consumer is another important factor in determining the demand for a commodity. The demand for a commodity is, in general, directly related to the income of the consumers. For normal goods, demand is directly related to income, that is, as incomes of the consumers go on increasing, the demand will

also go on increasing. However, for inferior goods (say, some coarse grains), demand is negatively related to income of the consumers.

(c) Prices of Related Goods : The related goods are of two types : Complementary and substitutable. Two goods are substitutable if when the price of one good increases (decreases), the demand for another good also increases (decreases). Tea and coffee are two substitute goods. On the other hand, in case of complementary goods (say, tea and sugar), if the price of one good increases (decreases), the demand for another good also decreases (increases).

(d) Tastes and preferences of the Consumers : Tastes and preferences of the consumers also affect the consumers' demand for a commodity. If the consumers' tastes and preference move in favour of the commodity, the consumer will demand more of the commodity.

(e) Consumer's expectations also affect the demand for the commodity. That is, if consumers expect a rise (fall) in future price of a commodity, its demand will tend to increase (decrease).

Apart from the factors discussed above, the market demand for a commodity will also depend on a number of other factors.

2.4 Law of Demand

Under the above background we are in a position to explain the Law of Demand in the following way :

Under the ceteris paribus clause the demand for a commodity is inversely related to the price of the commodity itself. In symbols, this law may be stated as

$$D = f(p)$$

$$\text{with } \frac{dD}{dp} < 0 .$$

The law of demand is one of the most important laws of economic theory. According to this law, other things being equal, if the price of a commodity falls (increases), the quantity demanded for it will rise (fall). That means, there is an inverse relationship between price and quantity demanded for a commodity, other things being the same.

The law of demand may by illustrated with the help of a demand schedule and a demand curve. A demand schedule is a tabular presentation of different prices of a commodity and its corresponding quantity demanded per unit of time.

Table 1 : Demand Schedule of an individual consumer

Commodity	Price (Rs.)	Quantity Demanded (Units)
A	5	10
B	4	15
C	3	20
D	2	35
E	1	60

From the above table it is seen that when price of the commodity concerned is Rs. 5/- per unit, a consumer purchases 10 units of the commodity. However, when price falls to Rs. 4, the consumer purchases 15 units of it. Again, when the price further falls, the quantity demanded by the consumer goes on rising until at price of Re. 1, the quantity demanded by him rises to 60 units. From the table we thus see an inverse relationship between price and quantity demaded : as the price of the commodity goes on rising, the demand for it goes on falling.

2.5 The Demand Curve

From the law of demand we have already seen a negative relationship between price of a commodity and quantity demanded for that commodity. This negative relationship between price and quantity demanded can, graphically, be presented in the following diagram, that is, Fig. 2.1 measuring price vertically and quantity demanded horizontally.

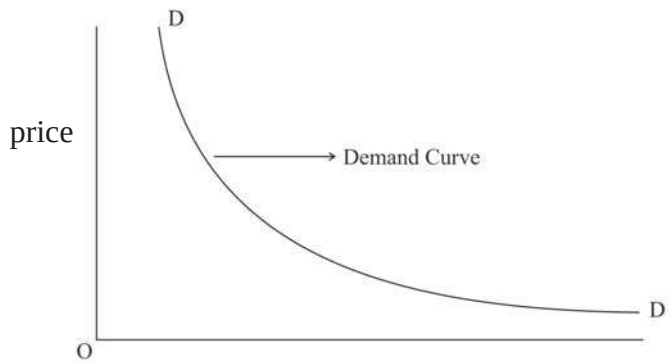


Fig 2.1 : Quantity Demanded

The hypothetical figures presented in Table-1 constitute the demand schedule. The graphical presentation of the demand schedule is known as a demand curve.

The above-drawn downward sloping demand curve DD, connecting price-quantity combinations is known as the demand curve for a commodity. The curve shows the quantity of X that a consumer would like to purchase at each price and its downward slope indicates that the quantity of 'X' demanded increases as its price falls given other factors. Hence the downward sloping demand curve is in accordance with the law of demand which as already stated, describes an inverse price–demand relationship.

2.6 Market Demand Schedule

When we add up the various quantities demanded by the number of consumers in the market we can obtain the market demand schedule. The market demand refers to the sum of the quantities demanded by all households at various prices. For different households we can get different individual demand curves. Naturally, each individual demand curve expresses the quantity demanded by a household at a particular price. Obviously, by summing all the quantities demanded by all the households, at a particular price we obtain the market demand curve at that price. In this manner if we find the total quantities demanded at all prices, we can get a market demand curve. This market demand curve shows total quantities demanded by all households at various prices.

Geometrically, by the horizontal summation of the individual demand curves we obtain the market demand curve. If the individual demand curves have normal properties, the market demand curve will also have normal properties. That means, if each individual household increases his quantity demanded when price decreases, the total quantity demanded will also increase as price decreases. However, even if some individuals purchase less of the commodity, when its price falls, the total quantity demanded may increase as price decreases, simply because most of the individuals will increase the quantity demanded as price decreases. Besides, with the fall in price of the commodity some new consumers who were not in a position to purchase the commodity previously, will now come to the market to purchase it. This reason also explains that total quantity demanded increases as price starts falling. Therefore, it can be said that under the 'ceteris paribus clause' there exists an inverse relationship between price of a commodity and the total quantity demanded for that commodity. Obviously, the market demand curve will be downward sloping.

We draw the market demand curve on the basis of the assumption that all other prices, total household income and its distribution among households, tastes and preferences of the households are kept constant. In case any of these factors changes, the market demand curve will shift its position. For instance, if we assume that there is a rise in income for all households, each household will go into increasing quantity demanded at each price.

Naturally, each household's individual demand curve will go on shifting to the right. Consequently, the market demand curve will also shift to the right. In that manner, if prices of other commodities change or, if tastes of households go on changing, then also the market demand curve will make a shift in its position.

The procedure for drawing the market demand curve from the individual demand curves has been demonstrated in the following diagram (Fig. 2.2) where along the horizontal axis we measure quantity demanded while price has been done vertically. For the sake of simplicity we have considered only two individuals : individual 1 and individual 2.

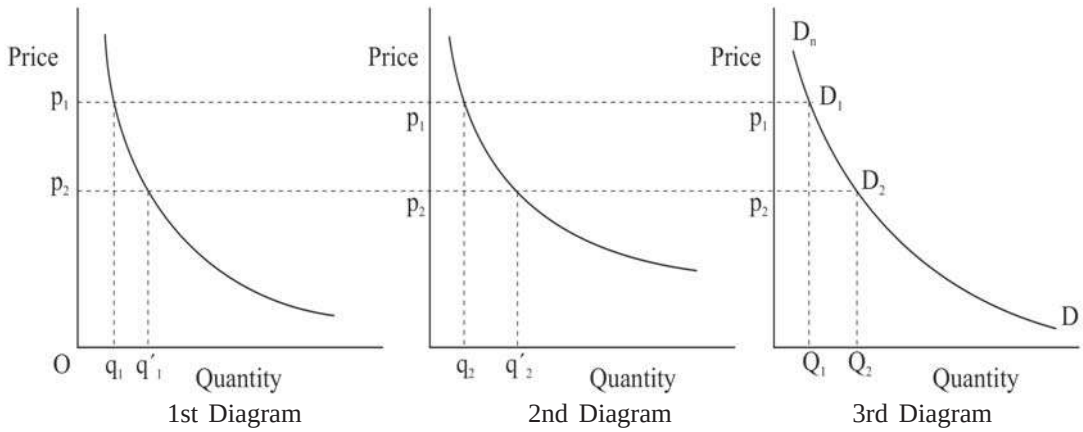


Fig 2.2 : Horizontal summation of individual demand curves to get the Market Demand Curve

At price op_1, the demand for the first individual is oq_1, while that of the second one is oq_2 and hence the market demand will be $oq_1 + oq_2 = OQ_1$, say. Similarly, when price gets reduced from op_1 to op_2, the demand for both the individuals increases, for the first individual the demand has become oq^1_1 while for the second one it has become oq^1_2, and hence the market demand will be $oq^1_1 + oq^1_2 = OQ_2$ in the third diagram. In this way we get another point (D_2) on the market demand curve. Naturally, the locus of these points will give the market demand curve.

We have drawn the market demand curve assuming that household's total income and its distribution among households, all other prices and tastes of households will remain constant. However, if any of these factors changes the market demand curve will shift its position. If we assume that there takes place a rise in income for all households, then naturally, each household will go on increasing quantity demanded at each price. Obviously,

each household's individual demand curve will shift to the right. Naturally, the market demand curve will shift to the right.

2.7 Exceptions to the Law of Demand

According to the law of demand, more of a commodity will be demanded at lower prices than at higher ones, other things being equal. This law is valid in most of the cases. In spite of that there are certain cases where this law does not hold good. The followings are the important exceptions to the law of demand.

(a) Conspicuous goods : Some consumers are habituated in measuring the uitility of a commodity by its price i.e., if the commodity is expensive, they think that it has got more utility. Accordingly, they buy less of this commodity at low price and more of it at a higher price. Diamonds are often given as an example of this case. Higher the price of diamond, higher is the prestige value attached to them and hence higher is the demand for them.

(b) Giffen goods : Sir Robert Giffen, an economist, was surprised to find out that as the price of bread increased, the workers of Britain started purchasing more bread and not less of it. This is something against the law of demand. The reason for this is that when the price of bread went up, it caused such a significant decline in the purchasing power of the poor people that they were forced to cut down the consumption of meat and also other expensive food items. Since bread even when its price was higher than before was still the cheapest food item, people started consuming more of it and not less when its price shot up.

(c) Conspicuous necessities : The demand for some goods gets affected by the demonstration effect of the consumption pattern of a social group to which an individual belongs. These goods, because of their constant use, have become necessities of life. For instance, in spite of the fact that the prices of refrigerators, coolers, cooking gas etc. have continuously been rising, their demand does not show any tendency to fall.

(d) Future expectations about prices : It has been observed that when prices are rising, households expecting that the prices in the future will be still higher, tend to buy a large quantity of the commodites. For instance, when there is wide-spread drought, people expect that prices of foodgrains would rise in future. They start demanding larger quantities of foodgrains as their price rises. But it is to be noted that here it is not the law of demand which is invalidated but there is a change in one of the factors which was held constant while deriving the law of demand, namely change in the expectations of price of the people.

(e) Behaviour in the share market : In the share market it is found that as the price

of any share increases, its demand also increases. In the same way when price of any share falls, the demand for that share also decreases. Therefore, the law of demand does not become applicable in the case of share market.

(f) The law has been derived assuming consumers to be rational and knowledgeable about market conditions. However, at times consumers tend to be irrational and make impulsive purchases without any cool calculation about prices and usefulness of the product. In such contexts also the law of demand does not hold good.

(g) Similarly, in practice, a household may demand larger quantity of a commodity even at a higher price because he may be ignorant of the ruling price of the commodity. Under such circumstances, the law of demand will remain invalid.

2.7.1 Giffen Good

The name Giffen Good owes its origin to Sir Robert Giffen. He found that a rise in the price of bread had been followed by an increase in the demand for bread. As the price of bread goes up, the real income of the consumers decreases very much. Consequently, the consumers are forced to curtail the consumption of meat and other expensive food items.

In the opinion of Nobel laureate Professor J. R. Hicks for a good to be a Giffen good, three conditions are of utmost importance : (i) the good must be an inferior one with large negative income effect, (ii) the substitution effect must be small and less strong than the income effect and (iii) the proportion of income spent upon the good must be very large.

The concept “Giffen good” is, theoretically, possible. But, in practice, it can hardly occur. The basic reason behind this is that consumption of the people is mostly diversified so that people spend a small proportion of their incomes on a single commodity. Consequently, the strength of the income effect induced by the change in price is likely to be very small and thus the negative income effect can not, generally, outweigh the substitution effect in practice.

Besides, a commodity may be a Giffen commodity to a single consumer. However, it is unlikely that the same commodity will be a Giffen one to all cansumers at the same time. Thus for a heterogeneous group of consumers the income effect will be positive for some consumers and negative for some other consumers. Thus the market income effect will be weak and the market substitution effect will always be stronger than the market income effect.

2.8 Bandwagon, Snob and Veblen Effects

We want to close our discussion on Demand Analysis by briefly making a brief discussion on the above-mentioned three effects on demand for a commodity.

We are to recall that the great developmental economist, Harvey Leibenstein in his article entitled "Bandwagon, Snob and Veblen Effect in the Theory of Consumers' Demand" published in the Quarterly Journal of Economics (1950) had explained these effects. He had clearly shown that in the presence of these three effects the true market demand curve will diverge from the horizontal summation of the individual demand curves. We shall discuss these effects, in brief, in the following paragraphs.

BANDWAGON EFFECT : A bandwagon effect is said to exist if any consumer tries to purchase the goods in order to behave like any member of his social group. He thinks that if other's demand for a good goes on increasing, so should his, as he wishes to identify him with them. Consequently, this effect makes the market demand curve more elastic. This effect has been demonstrated in figure 2.3(A).

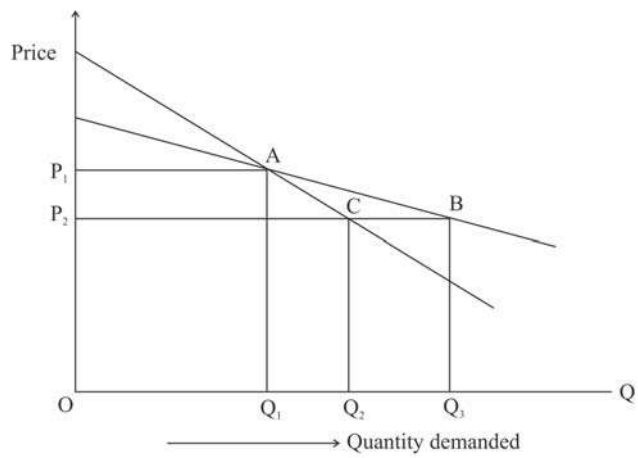


Fig 2.3 (A) : Bandwagon Effect

In the above figure AC is the market demand curve when the selfishness axiom is considered. When price falls from OP_1 to OP_2, the quantity demanded increases from OQ_1 to OQ_2, assuming that the bandwagon effect is absent. However, if this effect is present, many more consumers will enter the market for having this good. Consequently, the demand has extended to OQ_3. Thus the true market demand curve becomes AB instead of AC.

SNOB EFFECT : A ***snob effect*** exists when a consumer wants to differentiate himself from his social group by purchasing commodities which they do not purchase and conversely, reducing his purchase of commodtities which they purchase. Accordingly, when their demand increases, his demand will come down. Obviously, snob effect is opposite to that of the bandwagon effect and has the effect of making the demand curve more inelastic. Fig. 2.3 (B) illustrates this case. From Fig. 2.3(B) we see that when price falls from OP_1 to OP_2, the quantity demanded increases from OQ_1 to OQ_2 when the snob effect gets ignored. However, if the snob effect is operative, some consumers will reduce their purchase and the quantity demanded will increase to OQ_3 only. Therefore, the true market demand curve becomes AB which is relatively inelastic than AC.

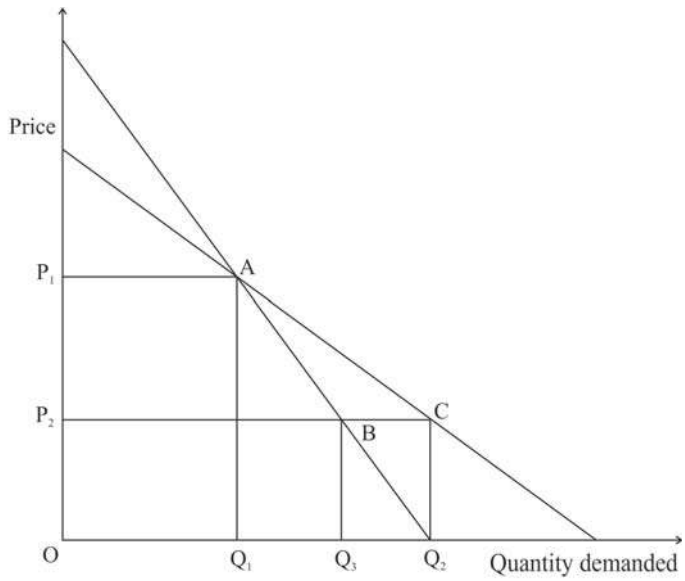


Fig 2.3 (B) : SNOB EFFECT

VEBLEN EFFECT : In any society there are some individuals who judge the quality of a commodity by merely looking at the price of the commodity itself. Therefore, whenever price of a commodity falls, some consumers will not be willing to purchase it, thinking that the fall in price is an indicative of a fall in quality of the commodity itself. In such a case the quantity demanded may even fall as price falls, resulting the market demand curve positively sloped. The Veblen effect may also make the market demand curve more inelastic. Figure 2.3 (C) has demonstrated the impact of this effect. In the following diagram AC is the demand

curve when the Veblen effect becomes absent and when the price falls from OP_1 to OP_2, the quantity demanded has increased from OQ_1 to OQ_2. However, when this effect is operative, some consumers will disappear from the market whenever price starts falling and the quantity demanded will increase upto OQ_3, not upto OQ_2. But if the veblen effect is strong enough the quantity demanded may even be OQ_4. That means, the quantity demanded may even decrease as price falls. However, the true demand curve becomes AB or even AD which is upward rising.

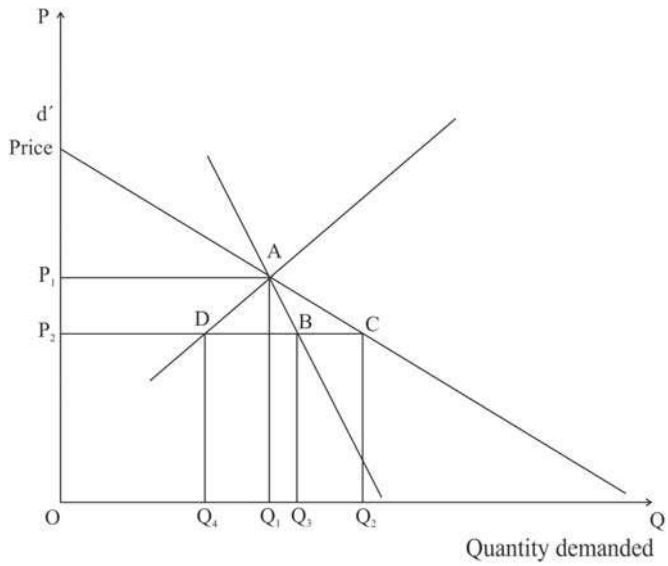


Fig 2.3 (C) : VEBLEN EFFECT

2.8.1 Summary of These Three Effects

The above-mentioned three effects actually imply interdependence among the purchase plans of the consumers. However, such interdependencies imply divergences between the true demand curve and the ordinary demand curves which have been obtained when the selfishness axiom has become operative. However, the extent of divergence will depend on the strength of these effects for any consumer and also on the number of consumers subject to such effects.

Of course, there is no reason to suppose that the market is composed of people subject to only one of the effects. Most of the times, however, these effects will be eliminated as they play their role in opposite directions. Besides, these effects are relevant only when market demand, not individual demand, is considered.

2.9 Elasticity And Its Applications In Demand Analysis

A very important concept in Managerial Economics is the concept of elastcity of demand which has got wide applications in both theory and practice.

We know that the demand for a commodity depends, basically, upon three factors : (A) price of the commodity itself (p), (B) income of the consumers (y) and (C) price of the related commodity (p^1). That means, the demand can, symbolically, be expressed as

$$D = f(p, y, p^1).$$

Elasticity is a very general concept of how responsive a dependent variable is to variation in an independent variable.

Since demand for a commodity depends primarily upon three variables, there are three types of elasticity of demand. These are : (i) Own price elasticity of demand, (ii) Cross price elasticity of demand and (iii) Income elasticity of demand. However, by elasticity of demand we normally understand own price elasticity of demand or simply price elasticity of demand as it has wide applications in different fields in theory and practice. So we shall, first of all, concentrate our attention on price elasticity of demand.

2.10 Price Elasticity of Demand

By elasticity of demand we normally understand price elasticity of demand. It measures the degree of reponsiveness of quantity demanded for a commodity following a change in own price of the commodity itself, assuming money income of the consumer and prices of related goods remain constant. This elasticiy is the relative difference in the dependent variable (that is, quantity demanded) divided by the relative difference in the independent variable (that is, price). In symbols, this is expressed as

$$ep = \frac{\Delta q}{\Delta p} ...(i)$$

This is the measurement used by Professor Boulding and it is called the absolute measure of elasticity of demand. The basic drawback of this measure is that it is not unit-free and since it is not unit-free, it cannot be compared among different price elasticities for different commodities. Hence it has later been discarded and replaced by the new measure called the “relative measure of elasticity of demand”. This measure is expressed as

$$ep = \frac{d \log q}{d \log p} \ldots (ii)$$

The basic difference between measurement (i) and (ii) is that in place of the absolute value of q and that of p the change in log value of q and that of p has been taken into consideration. Since log value of any number, irrespective of the unit of measurement, is free from unit of measurement, the measurement of price elasticity of demand given by (ii) is unit-free and can, therefore, be used to compare price elasticities of demand among a number of commodities expressed in various units. Besides, from (ii) we get

$$ep = \frac{dq}{q} \Big/ \frac{dp}{p}$$

$$= \frac{\text{Rate of change in quantity demanded}}{\text{Rate of change in price}}$$

$$= \frac{\text{Percentage rate of change in quantity demanded}}{\text{Percentage rate of change in price}}.$$

Since price and quantity demanded move in opposite directions, ep must always be a negative number. However, we want to make this expression positive and we write it as

$$ep = (-)\frac{dq}{q} \Big/ \frac{dp}{p} = (-) \frac{dq}{q} \times \frac{p}{dp} = (-) \frac{dq}{dp} \times \frac{p}{q}$$

$$= (-) \frac{p}{q} \times \frac{dq}{dp} = (-) \frac{p}{q} \times \frac{1}{\frac{dp}{dq}} = (-) \frac{p}{q} \times \frac{1}{\text{Slope of the demand curve}}.$$

$$= (-) \frac{p}{q} \times \text{Reciprocal of the slope of the demand curve.}$$

2.10.1 Types of Own Price Elasticty of Demand

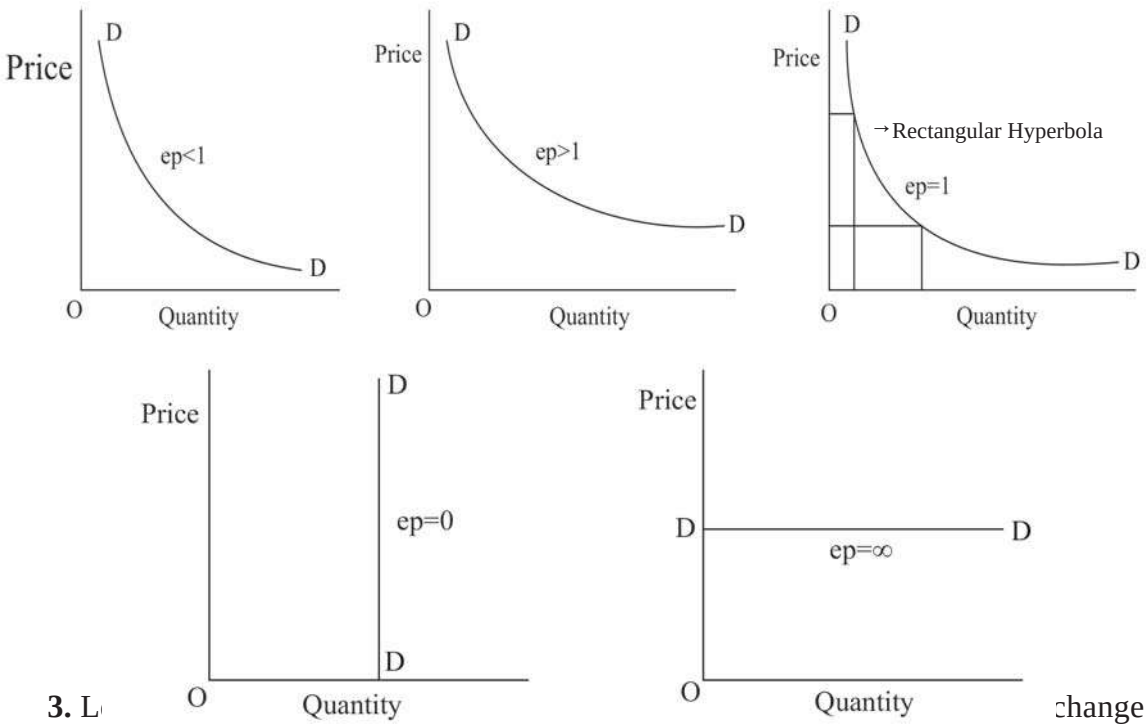
Since own price elasticity of demand has been expressed as a ratio which can assume any of the following values, own price elasticity of demand can also be of the following values :

$$ep \gtreqless 1 \text{ and } ep = 0, \text{ and also } ep = \infty.$$

Let us consider these cases one by one :

1. Let ep be less than one (that is, ep < 1). In this case $\frac{dq}{q} \Big/ \frac{dp}{p} < 1$, that is, rate of change in demand is less than the rate of change in price. If this happens for a commodity, the commodity concerned will be said to have inelastic demand. Normally, commodities which are very necessary for our life are said to have inelastic demand. Some such commodities are salt, text book, green vegetables, rice, clothes etc. In this case the percentage change in quantity demanded for a commodity is less than the percentage change in price of the commodity.

2. Let ep be greater than one (that is, ep > 1). Naturally, in this case $\frac{dq}{q} \Big/ \frac{dp}{p} > 1$. That means, the rate of change in quantity demanded is greater than the rate of change in price. If this happens for a commodity, the commodity concerned will be said to have elastic demand. Normally, demand is said to be elastic for luxurious goods. Some such goods are gold, diamond, very costly car etc. For such commodities the percentage change in quantity demanded is greater than the percentage change in price.



3. L change

Fig 2.3 (D) : Different demand curves according to the different values of price elasticity of demand

in quantity demanded is exactly equal to the percentage change in price. If this happens for a commodity, the commodity concerned will be said to have unitary elasticity of demand. It is not possible to name a commodity having unitary elasticity of demand. However, this elasticity has some applications in Ecomomic Theory. We shall later see that a firm, in order to maximise revenue, will be able to do so if ep equals one. According to Alfred Marshall, a commodity will be said to have unitary elasticity of demand if total expenditure to be incurred by the purchaser remains invariant whatever be the price of the commodity. That is why, the shape of the demand curve having unitary elasticity will be a rectangular hyperbola.

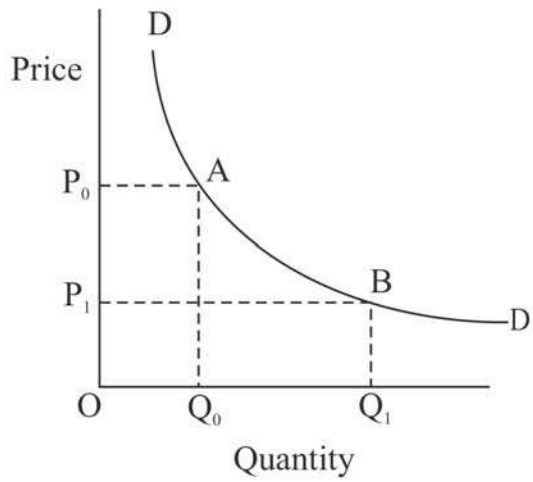


Fig 2.3 (E) :

4. Let ep be zero. In this case $\frac{dq}{q} \Big/ \frac{dp}{p}$ = 0 and this is possible when dq = 0. That means there is no change in quantity demanded whatever be the price of the commodity. In this situation the commodity concerned will be said to have perfectly inelastic demand. The demand curve will be a vertical straight line parallel to the vertical axis. Such a commodity is more relevant to supply.

5. Let ep be $\propto$. In this case $\frac{dq}{q} \Big/ \frac{dp}{p} = \propto$. This is possible when dp = 0. That means there is no change in price, that is, at the same price any amount may be purchased. This happens for the demand curve faced by a perfectly competitive firm where the demand curve is a horizontal straight line parallel to the horizontal axis.

The shapes of these five types of demand curves have been presented in the above diagram (Fig. 2.3 D.) at each of which the quantity demanded has been measured along the horizontal axis while price has been measured along the vertical axis.

2.10.2 Measurement of Price Elasticity of Demand

Price elasticity of demand is measured in two different ways : (a) Either over a portion of the demand curve or (b) on a particular point of the demand curve. When price elasticity is meausred over a portion of the demand curve, it is known as arc price elasticity of

demand. On the other hand, if price elasticity is measured on a point of the demand curve, we get the point elasticity of demand. These two methods of measuring price elasticity have been discussed below.

In the above diagram (Fig. 2.3 E) DD is the demand curve where over the portion AB of the curve we are interested in measuring the price elasticity of demand.

Let A be taken as the initial point at which price = OP_0 and quantity is OQ_0. Now change in price, $\Delta p = P_0P_1$ while change in quantity demanded, $\Delta q = Q_0Q_1$. Hence corresponding to point A, price elasticity is, $\left|e\frac{A}{p}\right| = \frac{p}{q}\left|\frac{\Delta q}{\Delta p}\right| = \frac{OP_0}{OQ_0}\cdot\frac{Q_0Q_1}{P_0P_1}$.

However, if B is taken as the initial point, the same price elasticity of demand in the zone AB of the demand curve DD will be

$$\left|e\frac{B}{p}\right| = \frac{p}{q}\left|\frac{\Delta q}{\Delta p}\right| = P\left|\frac{\Delta q}{\Delta p}\right| = \frac{OP_1}{OQ_1}\cdot\frac{Q_0Q_1}{P_0P_1}.$$

However, $\left|e\frac{A}{p}\right| \neq \left|e\frac{B}{p}\right|$ since initial price and initial quantity demanded change in these two situations. To get a unique value of price elasticity of demand over the arc AB, the initial price and initial quantity of demand are taken as the average of the prices of P_0 and P_1, that is, $\frac{P_0+P_1}{2}$ and the average quantities, that is, $\frac{Q_0+Q_1}{2}$. Hence the arc price elasticity of demand also known as the average elasticity of demand is obtained as

$$|ep| = \frac{(P_0+P_1)/2}{(Q_0+Q_1)/2}\times\frac{Q_0Q_1}{P_0P_1}.$$

$$= \frac{P_0+P_1}{Q_0+Q_1}\times\frac{Q_0Q_1}{P_0P_1}.$$

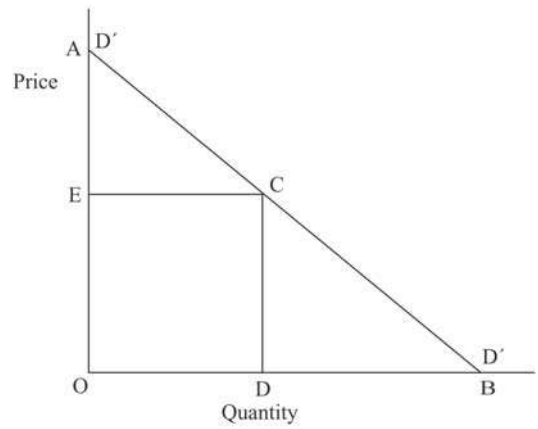


Fig 2.3 (F) :

2.10.3 Point Price Elasticity of Demand

The point price elasticity of demand is the limiting value of the arc price elasticity of demand when the change in price, that is, Δp tends to zero. Thus, in symbols, the point price elasticity of demand will be

$$ep = \underset{\Delta p \to 0}{Lt}\left(\frac{p}{q}\cdot\frac{\Delta q}{\Delta p}\right) = \frac{p}{q}\cdot\frac{dq}{dp}$$

$$\text{and} \quad |ep| = -\left(\frac{p}{q}\cdot\frac{dq}{dp}\right).$$

Here $\frac{dp}{dq}$ denotes the slope of the demand curve.

Let us now measure the point price elasticity of demand diagrammatically. The process of measuring the point price elasticity of demand has been demonstrated in the above diagram [Fig. 2.3 (F)].

Suppose that we want to measure price elasticity of demand on the point C of the straight line demand curve D^1D^1. From the diagram we notice that at point C, p = OE = DC, q = OD and the absolute slope of the demand curve D^1D^1 is

$$\frac{dp}{dq} = \left|\frac{CD}{BD}\right| \text{ so that } \left|\frac{dq}{dp}\right| = \frac{BD}{CD}$$

$$\text{That is, } |ep| = \frac{p}{q}\cdot\left|\frac{dq}{dp}\right| = \frac{CD}{OD}\cdot\frac{BD}{CD} = \frac{BD}{OD} = \frac{BD}{EC}.$$

Let us now consider the triangles AEC and CDB where $\angle AEC = \angle CDB$ (right angles), $\angle EAC = \angle DCB$ (since OA||DC) and $\angle ACE = \angle CBD$ (since EC||OB).

Therefore, ΔEAC and ΔCDB are similar and so their corresponding sides are proportional. That means, $\frac{DB}{EC} = \frac{CD}{AE} = \frac{CB}{CA}$.

$$\text{That is, } \frac{DB}{EC} = \frac{EO}{AE} = \frac{CB}{CA}.$$

$$\text{That means, } |ep| = \frac{DB}{OD} = \frac{EO}{AE} = \frac{CB}{CA} = \frac{\text{Lower segment of the demand curve}}{\text{Upper segment of the demand curve}}.$$

Thus at a particular point on a straight line demand curve the point elasticity of demand is obtained as the ratio between the lower segment of the demand curve below the point and the upper segment of the demand curve above the point.

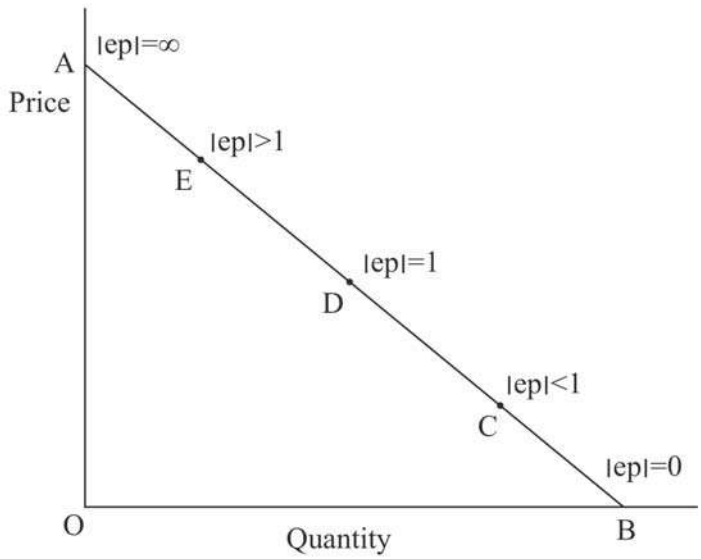


Fig 2.3 (G) :

By applying this formula, we can show that on a straight line demand curve the absolute value of the price elasticity of demand varies between O and ∝.

In the annexed diagram at point B, the point elasticity of demand is $\frac{\text{Lower segment}}{AB} = \frac{O}{AB} = 0$.

At point C, the point elasticity of demand is $\frac{CB}{AC} < 1$ as CB < AC. At point D, this elasticity is $\frac{DB}{AD} = 1$ as D is the mid-point between A and B. At point E the point elasticity of demand is $\frac{EB}{AE} > 1$ as EB > AE. Finally, at point A the point elasticity of demand is $\frac{AB}{\text{upper segment}} = \frac{AB}{O} = \infty$.

Thus the absolute value of the point elasticity of demand varies between O and ∝.

2.10.4 Income Elasticity of Demand

The income elasticity of demand is defined as the responsiveness of the quantity demanded for a commodity resulting from a change in income of the consumer alone, assuming that price of the commodity concerned and price of the related commodity remain unchanged. Thus, it is defined as

$ey = \frac{d \log q}{d \log y}$, y being income of the consumer.

$$= \frac{dq}{q} \Big/ \frac{dy}{y} = \frac{\text{Rate of change in quantity demanded for a commodity}}{\text{Rate of change in income of the consumer}}.$$

That is, $ey = \frac{\text{Proportionate change in quantity demanded}}{\text{Proportionate change in income of the consumer}}$.

2.10.5 Uses of Income Elasticity of Demand

Income elasticity of demand is very useful in classifying a number of commoditites around us. Its uses will be understood in the following manner :

1. Let ey be negative, that is, ey < 0. Normally ey is supposed to be positive. That is, with the increase/decrease in income more (less) of a commodity is demanded. But when ey < 0, it will mean that with the increase (decrease) in income, demand for a commodity decreases (increases). If this happens for a commodity, the commodity concerned is called an inferior commodity. Cheap washing powder might be an example. Those living on a tight budget may be unable to afford reasonable quality of washing powder. But as their incomes increase, they give up the cheap washing powder and switch to a better quality of washing powder. Another example is with the increase (decrease) in income the demand for coarse variety of rice decreases (increases). Hence for an inferior commodity the income elasticity of demand is negative.

2. Let ey be positive but greater than one. In such a case $\frac{dq}{q} > \frac{dy}{y}$. That is, the rate of change in quantity demanded is greater than the rate of change in income. This happens for a luxurious or prestigious commodity. For example, with the increase (decrease) in

income the demand for gold, jewellery, luxury car etc. generally, increases (decreases). Thus for a luxurious or prestigious good the income elasticity of demand is positive and its numerical value is greater than one.

3. Let ey be positive but less than one, that is o < ey < 1. In this case the rate of change in quantity demanded is less than the rate of change in income of the consumer. This happens for a commodity which is a normal necessary, that is, commodities which are useful for our daily life.

The uses of income elasticity of demand can also be understood from the following presentation :

Nature of a commodity[a]	Inferior	Normal necessary	Normal luxurious
Value of ey →	0	→ 1	→

The income elasticity of demand for various categories of goods may, however, vary from time to time and also from household to household, depending on the choice and preference of the consumers, levels of consumption and income and their susceptibility to "demonstration effect". The other factor which may cause deviation from the general pattern of income elasticities is the frequency of increase in income. If frequency of rise in income is high, income elasticities will conform to the general pattern.

2.10.6 Cross Price Elasticity of Demand

Demand for a commodity is also influenced by the price of other related goods and services. The responsiveness for a change in quantity demanded for a commodity resulting from a change in price of its related commodity is measured by cross price elasticity of demand. Thus,

$ep^1 = \frac{d \log q}{d \log p^1}$ where q is the quantity demanded for a commodity (say x) and p^1 is the price of a related commodity (say y).

$$= \frac{dq}{q} \Big/ \frac{dp^1}{p^1} = \frac{dq}{q} \times \frac{p^1}{dp^1} = \frac{dq}{dp^1} \times \frac{p^1}{q}.$$

$$= \frac{\text{Rate of change in demand for a commodity , say x}}{\text{Rate of change in price of a related commodity , say y}}.$$

2.10.7 Uses of Cross Price Elasticity of Demand

Cross price elasticities are used to classify the relationship among a number of goods. It has been demonstrated in the following manner :

Case 1 : Let $ep^1 > 0$. In this case an increase in the price of y causes an increase in the quantity demanded for x and the two products are said to be substitutes. That means, one product can be used in place of (substituted for) the other. Suppose that the price of y increases. This means that the opportunity cost of y in terms of x has increased. The result is that the consumers purchase less of y and more of the relatively cheaper good, x. Chicken and Mutton (tea and coffee) are examples of substitutes. An increase in the price of chicken (tea) usually leads to an increase in the demand for mutton (coffee) and vice versa.

Case 2 : Let $ep^1 < 0$. In this case an increase in the price of y causes a decrease in the quantity demanded for x and the two products are said to be complements (goods that are used jointly or together). That means an increase in the price of y leads to a reduction in the quantity demanded for that product. The diminished demand for y causes a reduced demand for x. Butter and bread, cars and tyres, computers and computer software are examples of pairs of goods that are complements.

Cross elasticities are not always symmetrical. This means that a change in demand for good x caused by a change in the price of y may not equal the change in demand for y generated by a change in the price of x. Although the two elasticities are different, they are of the same sign.

Case 3 : Let ep^1 be zero. In this case an increase or decrease in the price of y does not have any impact on the demand for x so that ep^1 becomes zero. The two goods will, obviously, be independent or unrelated.

Therefore, the sign of cross price elasticity of demand helps us in understanding the relationship between two commodities : whether they are substitutes or complements or they are unrelated or independent of each other.

2.11 Relation between Elasticity of Demand and Total Expenditure

Alfred Marshall had suggested to measure elasticity of demand in terms of total expenditure to be incurred by a consumer.

We know that under the ceteris paribus clause, demand for a commodity is a function of the price of the commodity itself. Thus as price changes along a demand curve, the quantity demanded will change. With such changes, an individual consumer's expenditure on the commodity would also change. By observing how this expenditure changes in response to price change, we can predict whether demand is elastic, inelasic or unitary elastic.

We should remember that total outlay or expenditure or revenue is price multiplied by quantity demanded, that is,

R = pq, R being total expenditure, p is the price of the commodity and q is the quantity demanded by a consumer. We consider the following two cases :

(a) We suppose that price of the commodity falls. Then according to this method,

(i) the total outlay or expenditure on the good could increase and demand for the good would be price elastic, that is, elasticity > 1,

(ii) the total outlay or expenditure on the good could fall and demand for the good would be price inelastic, that is, elasticity < 1

and (iii) the total outlay or expenditure on the good could remain the same and demand would be unitary elastic, that is, e = 1.

(b) We suppose that price of the commodity rises. Then according to this method,

(i) the total outlay or expenditure on the good falls and demand is elastic, that is, e > 1,

(ii) the total outlay or expenditure on the good also increases and demand is inelastic, that is, e < 1

and (iii) the total outlay or expenditure on the good remains the same and demand is unitary elastic, that is, e = 1.

It is to be noted that with the help of the total expenditure or outlay method, we fail to derive the exact value of elasticity of demand. The calculated value can, however, be used to indicate only whether elasticity is greater than or less than or equal to 1 (one).

Exercise :

1. From the following demand schedule find the direction of elasticity when price falls from Rs. 10/ to Rs. 9/; from Rs. 9/ to Rs. 8/; from Rs. 8/ to Rs. 7/ and from Rs. 7/ to Rs. 6/, using the outlay method.

Price (Rs.)	10	9	8	7	6
Quantity demanded (Kg)	60	75	80	84	88

Solution : When price falls from Rs. 10/- to Rs. 9/-, total expenditure increases from Rs. 600/- to Rs. 675/. So in this case demand is elastic. Again, when price falls from Rs. 9/ to Rs. 8/, total expenditure decreases from Rs. 675/ to Rs. 640/. So in this case the demand is inelastic. Moreover, when price falls from Rs. 8/ to Rs. 7/, expenditure decreases from Rs. 640/ to Rs. 588/. So in this case the demand is inelastic. Finally, when price falls from Rs. 7/ to Rs. 6/, total expenditure decreases from Rs. 588/ to Rs. 528. So in this case also the demand is inelastic.

2. The monthly market demand curve for calculators among management students is given by p = 400 – Q, p being price per calculator in rupees and Q is the number of calculators purchased per month. Obtain the price elasticity of demand when price is fixed at Rs. 300/ per calculator.

Solution : We have Q = 400 – p. That is, p = 400 – Q

$$\therefore \frac{dp}{dQ} = -1, \text{ or, } \frac{dQ}{dp} = -1.$$

Thus the price elasticity of demand at Rs. 300/ will be

$$\frac{p}{Q}\cdot\frac{dQ}{dp} = \frac{300}{400}\times(-1) = -0.75. \text{ Ans.}$$

That is, R = pq, R being total expenditure, p is the price of the commodity and q is the quantity demanded by a consumer. We consider the following two cases :

(a) We suppose that price of the commodity falls. Then according to this method,

(i) the total outlay or expenditure on the good could increase and demand for the good would be price elastic, that is, elasticity > 1,

(ii) the total outlay or expenditure on the good could fall and demand for the good would be price inelastic, that is, elasticity < 1

and (iii) the total outlay or expenditure on the good could remain the same and demand would be unitary elastic, that is, e = 1.

(b) We suppose that price of the commodity rises. Then according to this method,

(i) the total outlay or expinditure on the good falls and demand is elastic, that is, e > 1,

(ii) the total outlay or expenditure on the good also increases and demand is inelastic, that is, e < 1

and (iii) the total outlay or expenditure on the good remains the same and demand is unitary elastic, that is, e = 1.

It is to be noted that with the help of the total expenditure or outlay method, we fail to derive the exact value of elasticity of demand. The calculated value can be used to indicate only whether elasticity is greater than or less than or equal to 1 (one).

2.12 Advertisement or Promotional Elasticity of Demand (or Sales)

The expansion of demand by means of advertisement and other promotional methods may be measured by advertising elasticity of demand, also called promotional elasticity. The promotional elasticity measures the responsiveness of demand to changes in advertising or other promotional expenses. The formula for its measurement is given by

$$ea = \frac{\text{Proportional change in sales}}{\text{Proportional change in advertisement expenditure}}$$

$$= \frac{\Delta Q}{Q} \Big/ \frac{\Delta A}{A} = \frac{\Delta Q}{Q} \times \frac{A}{\Delta A} = \frac{\Delta Q}{\Delta A} \times \frac{A}{Q}.$$

$$= \frac{Q_2 - Q_1}{A_2 - A_1} \times \frac{A_1 + A_2}{Q_1 + Q_2}.$$

Sales of different goods react differently to the doses of advertisement expenditure.

Even the same commodity may not respond the same way at different levels of advertising expenditure. This implies that advertising elasticity of demand differs between products and also between different levels of sales of the same product.

The advertising elasticity of demand is affected by a number of factors, namely (1) stage of the product market, (2) effect of advertising in terms of time and (3) influence of advertising by rivals.

2.12.1 Factors on Which Elasticity of Demand Depends

There are a number of factors which influence the price elasticity of demand for a good. The followings are some important determinants of price elasticity of demand.

(i) Number of substitutes available : Larger the number of substitutes available for a given commodity, the higher is the price elasticity of demand for it. The substitution effect is felt very strongly in such cases. For example, demand for electric goods is very elastic. On the other hand, fewer the number of substitutes available, the lower will be the elasticity of demand. The substitution effect is felt very weakly or not felt at all. The demand for salt is a nice exmaple of this case.

(ii) Nature of the good : A good can be basic or non-basic, a necessity or a luxury. For necessity and basic goods demand is less elastic. They have to be consumeed in certain quantity, irrespective of price prevailing. For example, food items (cereals, cooking oil, sugar, potatoes, salt, onions etc.) coarse clothing, transport to and from the place of work etc. have inelastic demand. The substitution effect is very weak for such goods. For non-basic and luxury goods demand would be elastic and both substitution and income effects are felt very strongly. For example, demand is elastic in cases like entertainment, electrical gadgets, finer clothes etc.

(iii) Proportion of income spent on a good, that is, importance of the commodity in the consumer's budget also affects elasticity of demand for a commodity. Higher the budget proportion more strongly will the income effect be felt. As a result, demand for such goods will be highly responsible for price change. Therefore, demand will be elastic. For example, all consumer durables like refrigerator, geyser, motor bikes and washing machines tend to be more demand elastic. On the other hand, smaller the budget proportion, the more weak

will be the income effect. Demand for such goods will be insensitive to price change. Naturally, demand will be less elastic.

(iv) Number of uses to which a commodity can be put : The more the possible uses of a commodity, the greater will be its price elasticity and vice versa. For example, milk has a number of uses. If its price falls, it can be used for a variety of purposes, like preparation of cream, curd, ghee and sweets. However, if its price increases, its uses will be restricted only to essential purposes, like feeding the children and sick persons.

(v) Consumer habits : If a consumer is a habitual one of a commodity, no matter how much and how its price changes, the demand for the commodity will be inelastic. That is why the demand for liquor, cigarettes is inelastic.

(vi) Height of price and range of price change : There are some goods like costly luxury items or bulky goods such as refrigerators, T.V. sets, which are highly priced in general. In such a case, a small change in price will have an insignificant effect on their demand. Their demand will, therefore, be inelastic. However, if the price change is large enough, then their demand will be elastic. In the same way, there are perishable goods like potatoes and onions etc. which are relatively low-priced and bought in bulk, so a small variation in their prices will not have much effect on their demand and hence their demand tends to be inelastic.

(vii) Durability of the commodity : In the case of durable goods, the demand generally tends to be inelastic in the short run, e.g., furniture, motor bike, television etc. In the case of perishable commodities, on the other hand, demand is relatively elastic, e.g., vegetables, milk etc.

(viii) Tied demand : The demand for those goods which are tied to others is normally inelastic as against those whose demand is of autonomous nature.

(ix) Possibility of postponement : When the demand for a product is postponable, it will tend to be price elastic. In the case of consumption goods which are urgently and immediately required, their demand will be inelastic.

2.12.2 Uses of The Concept of Elasticity of Demand

The concept of elasticity of demand has a good number of uses :

(a) Elasticity of demand is considered very useful in price determination. Whereas in case of perfect competition it determines how a change in supply would affect its price, in case of monopoly a supplier has to consider it in both normal price determination and in price discrimination.

(b) This concept is considered useful in budget formulation, specially, the incidence of taxation on consumers depends upon elasticity of demand. A tax on a commodity having inelastic demand may be shifted to the consumers.

(c) The terms of trade depend upon elasticity of demand and supply. Normally, the terms of trade are unfavourable to a country whose demand for imported foreign articles is inelastic, the reverse is also true when the import demand is elastic.

(d) Elasticity of demand is helpful in forecasting demand. Given the elasticity of demand and the state of independent variable, it is possible to forecast the demand for a good.

(e) The effectiveness of price controls also depends upon elasticity of demand. With an inelastic demand for a controlled product, it would, generally, sell in the black market.

(f) The effect of devaluation of currency also depends upon elasticity of demand and elasticity of supply of goods and services entering into international trade. In case of inelastic demand for imports as well as inelastic supply for exports, devaluation may not be beneficial in terms of reducing the adverse balance of payments.

2.13 Demand Forecasting : Meaning

Forecasting of demand is the art of predicting demand for a product or a service at some future date on the basis of certain present and past behavioural patterns of some related events. However, forecasting is no simple guessing, but it refers to estimating scientifically and objectively on the basis of some facts and events relevant to the art of forecasting.

Every business enterprise should have distinct knowledge about the demand for its

product since business decisions regarding the amount of capital to be raised, labour needed, requirement of working capital, sources of raw materials, sales promotional activities etc. are highly dependent on the perception of the demand for its product. However, if this perception is erroneous, the decisions of the enterprise will not be correct and that will lead to huge losses. Hence, for a successful venture, correct estimation of future demand is highly significant. Demand forecasting is merely an attempt to estimate this future demand.

Forecasting of demand plays a vital role in the process of planning and decision making whether at the national level or at the level of a firm. The importance of demand forecasting has increased overtime on account of mass promotion and production in response to demand. No useful business planning can be done without proper estimates of demand forecast. It can be said that all business planning starts with forecasting of demand because capital investment, production scheduling etc. have to be related with expected demand. Although it is said no forecast is completely fool-proof and correct but the very process of forecasting helps in evaluating various forces which affect demand and is, in itself, a reward because it enables the forecasting authority to know about various forces relevant to the study of demand behaviour.

2.13.1 Importance of Demand Forecasting

Demand forecasting is of high use both to the firms and to the Government. Some reasons can be advanced in support of this :

(a) For production planning demand forecasting is a must. If a firm can, more or less, accurately forecast the future demand for its product, it can, accordingly, arrange its production schedule. Besides, it can also hire various factors of production in advance just to meet the necessities of production.

(b) For inventory planning also demand forecasting is highly useful. If future demand is within the knowledge of the entrepreneurs, optimum level of inventory can be determined and maintained.

(c) Apart from proper scheduling of production, demand forecasting enables a firm to avoid overproduction as well as underproduction.

(d) For estimating the financial requirements a firm should be in the knowledge of demand forecasting.

(e) Demand forecasting for a particular product gives an appropriate idea about demand forecast for future product. Forecast for bi-cycles gives an appropriate idea about future demand for tyres and tubes in bi-cycles.

(f) Finally, demand forecasting is helpful to the Govt. of a country in a number of ways. To distribute any commodity among the public, to determine target of exports and imports, to give subsidy to any commodity, to estimate various components of Government budget, to formulate the overall economic planning in the country and the like more or less an accurate demand forecasting is a pre-condition.

2.13.2 Methods of Demand Forecasting

There is no easy method or simple formula which enables an individual or a business person to predict the future demand with certainty or to escape the hard process of thinking. The firm has to apply a proper mix of judgement and scientific formula in order to correctly predict the future demand for a product.

Broadly speaking, demand forecasting may be categorised under two broad methods :

(a) Survey Methods and (b) Statistical Methods.

In method (a) the demand for a product is forecast by enquiring the intentions and views of the participants going to the market. In method (b), on the other hand, past data are used to forecast future demand.

Survey methods can, again, be divided into three categories :

(i) Survey of Buyers' intentions

(ii) Collective Opinion Method

and (iii) Expert Opinion Method.

Let us now discuss these methods, in brief.

(i) Survey of Buyers' Intentions : The most direct method of estimating demand, in the

short run, is to ask customers what they are planning to buy for the forthcoming period, usually, a year. So in this method the burden of forecasting is put on the customers. However, it would not be wise to depend wholly on the buyers' estimates and they should be used very cautiously in the light of a seller's own judgement. A number of biases may creep into the surveys. Besides, the customers may themselves misjudge their requirements or may mislead the surveyors. However, this method is useful when bulk of sale is made to industrial producers who, generally, have firm future plan. In the case of household customers this method may not prove very helpful.

(ii) Collective Opinion Method : In this method salesmen are required to estimate expected sales in their respective territories. The rationale of this method is that salesmen being closest to the customers are likely to be the most intimate feel of the market. The estimates of salesmen are consolidated to find out the total estimated sales. These estimates are reviewed to estimate the bias of optimism on the part of some salesmen and pessimism on the part of others. These revised estimates are further examined in the light of factors like proposed change in selling prices, product designs and advertisement, expected changes in competition and changes in secular prices like purchasing power, income distribution, employment, population etc. The final sales forecast would emerge after these factors have been taken into account.

Although this method is simple and is based on first hand information of those who are directly, connected with sales, this method is subjective as personal opinions can, possibly, influence the forecast. Besides, salesmen may be unware of the broader economic changes having the impact of future demand. Therefore, forecasting could be useful in the short run, for the long run analysis, however, a better technique is to be applied.

(iii) Expert Opinion Method : This method is also known as delfi method of investigation. In this method instead of depending on the opinions of buyers and salesmen, firms can obtain views of the specialists or experts in their respective fields. Opinions of different experts are sought and analysed. The process goes on until some sort of unanimity is arrived at among all the experts. This method is best suited in circumstances where intractable changes are occurring. Besides, it has the advantage of speed and cheapness.

2.14 Statistical Methods

Statistical methods have proved to be very useful in forecasting demand. The important statisical methods of demand forecasting are :

(i) Trend Projection Method

(ii) Graphical Method

(iii) Least Squares Method

and (iv) Regression Method.

Let us now discuss the above methods one by one.

(i) Trend Projection Method : A firm which has been in existence for some time, will have accumulated considerable data on sales pertaining to different time periods. Such data when arranged chronologically yield 'time series'. This can be done either through graph or through least squares method.

(ii) Graphical Method : A trend line can be fitted through a series graphically. Old values of sales for different areas are plotted on a graph paper and a free hand curve is drawn passing through as many points as possible. The direction of this free hand curve shows the necessary trend. The main drawback of this method is that it may show the trend but it fails to measure it.

(iii) The Least Squares Method : This method is based on the assumption that the past rate of change of the variable under study will continue in the future. It is a mathematical procedure for fitting a line to a set of observed data points in such a manner that the sum of squared differences between the calculated and observed values is minimised. This technique is used to find a trend line which best fits the available data. This trend is then used to project the dependent variable in the future. This method is very popular because it is simple and inexpensive.

(iv) Regression Method : This is a very common method of forecasting demand. Under this method a relationship is established between quantity demanded (the dependent variable) and independent variables such as income, price of the good, prices of related goods and so on. Once the relationship is established, we derive a regression equation

assuming the relationship to be linear. The eqauation will be of the form

$$Y = a + bx.$$

There could also be a curvy linear relationship between the dependent and independent variables. Once the regression equation is derived, the value of y, i.e. quantity demanded can be estimated, for any given value of x. The values of a and b get estimated from given data on the basis of the following two normal equations :

$$\sum_{i=1}^{n} yi = na + b\sum_{c=1}^{n} xi$$

and $$\sum_{i=1}^{n} yi\, xi = a\sum_{i=1}^{n} xi + b\sum_{i=1}^{n} xi^2 .$$

Once the values of a and b get estimated, we can forecast y for a particular value of x from the equation

$$Y = a + bx.$$

This method can be estimated in two ways. Instead of linear regression equation the non-linear regression equation can also be fitted. Moreover, the multiple regression equation can be estimated in place of the simple regression equation incorporating more than one independent variable.

The regression method is objective in the sense that we will get a unique forecast value of the demand for the unique sets of historical data on x and y, given the form of the regression equation. Besides, this method is less expensive and is simple to understand and, therefore, it is very popular.

2.15 Estimation of Demand Functions

A demand function can be estimated given the information on price (p) and quantity demanded (q). This has been demonstrated in the following manner :

Let the general form of the demand function with one independent variable be

$$Y = a + bx$$

where 'a' and 'b' are two parameters.

Given a set of n pairs of observations (x_1, y_1), (x_2, y_2), (x_3, y_3),..., (x_n, y_n) we are to estimate the values of the parameters 'a' and 'b'. We have to choose the values of a and b in such a way that the sum of squares of errors gets minimised. Hence the method is known as the least squares method.

Now, error, ei = observed value – estimated value.

$$= (y_i - Y_i) = (y_i - a - bx_i).$$

Thus, the sum of squares of errors = S

Now, $\sum e_i^2 = \sum_{i=1} (y_i - a - bx_i)^2$.

Since (x_i, y_i) are given, we see that S is a function of a and b only.

For the minimization of S, the first order conditions are :

$$\frac{\partial S}{\partial a} = \sum(y_i - a - bx_i)\,(-2) = 0 \quad \text{.....} \quad \text{(A)}$$

$$\frac{\partial S}{\partial b} = \sum(y_i - a - bx_i)\,(-2x_i) = 0 \quad \text{.....} \quad \text{(B)}$$

Equations (A) and (B) can be written in the following forms :

$$\sum y_i = na + b\sum x_i \quad \text{.....} \quad (A^1)$$

$$\sum y_i x_i = a\sum x_i + b\sum x_i^2 \quad \text{.....} \quad (B^1)$$

Equations (A^1) and (B^1) are known as normal equations.

From the observed data we get the values of $\sum y_i$, $\sum x_i$, $\sum y_i x_i$ and $\sum x_i^2$. Thus there are two equations needed for solving two unknowns, a and b. Therefore, solving equations (A^1) and (B^1) we get the estimated values of the parameters a and b.

These estimated values when put in Y = a + bx give the necessary demand function. This can be explained with the help of the following example :

Example 1 : Fit a linear demand function to the following data :

Price (p) :	4	5	6	8	9
Quantity demanded (q) :	90	85	75	77	80

Solution : Suppose the linear demand function is $q = a + bp$. The values of the parameters 'a' and 'b' can be obtained by solving the two normal equations :

$$\sum q = na + b\sum p$$

$$\text{and } \sum pq = a\sum p + b\sum p^2.$$

p	q	p^2	pq
4	90	16	360
5	85	25	425
6	75	36	450
8	77	64	616
9	80	81	720
$\sum p_i = 32$	$\sum q_i = 407$	$\sum p^2_i = 222$	$\sum p_i q_i = 2571$

Here n = 5, so that the normal equations are : $5a + 3b = 407$

and $32a + 222b = 2571$.

Solving the above two equations we get a = 93.99 and b = –1.96

Hence the linear demand curve will be $q = 93.99 – 1.96p$.

Verification

At p = 4, q = 86.15
At p = 5, q = 84.19
At p = 6, q = 82.23
At p = 8, q = 78.31
At p = 9, q = 76.35
Thus $\sum q = 407.23$.

2.16 Sample Questions

A. Objective-type Questions:

1. Giffen goods are those goods for which the slope of the demand curve is :

   (a) Negative, (b) Positive, (c) Zero, (d) undefined.

Ans. (b)

2. Marginal revenue will be zero when :

   (a) ep < 1, (b) ep > 1, (c) ep = 1, (d) ep = 0.

Ans. (c)

3. If cross price elasticity of demand for good x and y is negative, the goods will be :

   (a) Substitutes, (b) Complements, (c) Independent and (d) Giffen.

Ans. (b)

4. If regardless of change in its price, the quantity demanded for a good remains unchanged, then the demand curve for the good will be :

   (a) Horizontal, (b) Vertical, (c) Positively sloped, (d) Negatively sloped.

Ans. (b)

5. If price elasticity of demand is zero, (ed = 0), the demand curve will be :

   (a) A horizontal straight line, (b) A vertical straight line, (c) A downward falling straight line, (d) An upward rising straight line.

Ans. (b)

6. In case of an inferior good, the income elasticity of demand is :

   (a) Positive, (b) Zero, (c) Negative, (d) Infinity.

Ans. (c)

7. In case of a straight line demand curve meeting the two axes, the price elasticity of demand at the mid-point of the line would be :

   (a) 0, (b) 1, (c) 1.5, (d) 2

Ans. (b)

B. Short-type Questions :

1. What is the difference between individual demand and market demand? How are these two types of demand related?
2. What is cross price elasticity of demand? Is it possible to express or to measure complementariy or substitutability between two goods by means of cross price elasticity of demand?
3. Establish the relation among MR, AR and price elasticity of deamnd. From this relation find out the value of total revenue when price elasticity of demand is unity.
4. What is Giffen's paradox? Explain it clearly with an example.
5. What is the difference between individual demand and market demand? How are these two types of demand related?
6. Suppose the price of a commodity increases from Rs. 200/- to Rs. 240/- and the quantity demanded for it decreases from 100 units to 75 units. Evaluate the price elasticity of demand for the commodity and interpret the result you arrive at.
7. "Elasticity of demand is ''unit-free''. Justify the statement.
8. What is income elasticity of demand? Classify commodities on the basis of the sign and magnitude of income elasticity of demand.
9. Show that all Giffen goods are inferior goods but all inferior goods are not Giffen goods.
10. Show that the absolute value of price elasticity of demand is different at different points on the same straight line demand curve.
11. Make a distinction betwen (i) an individual demand curve and the market demand curve, (ii) movement along the demand curve and any movement off the demand curve.

C. Essay-type Questions :

1. Explain, clearly, your idea about bandowagon effect, snob effect and veblen effect.
2. State and explain the law of demand. What are the non-price determinants of demand?

3. Explain, clearly, the factors affecting elasticity of demand for a commodity.
4. What is advertising elasticity of demand? Explain the importance of studying it.
5. Explain, clearly, the distinction between “Change in demand” and “Change in quantity demanded”. What are the exceptions to the law of demand?
6. What is demand forecasting? What is the utility for studying it? Explain the statistical method of forecasting demand.
7. Make a clear distinction between elastic demand and inelastic demand. Explain the factors affecting price elasticity of demand. What is the utility for studying cross price elasticity of demand?
8. Distinguish, clearly, between the own price elasticity of demand and cross price elasticity of demand. Explain the necessity for studying these two types of elasticity of demand.
9. Explain with illustrations, the various types of elasticity of demand. Draw the necessary demand curves according to the numerical values of elasticity of demand.
10. Explain the law of demand with illustrations. Give circumstances under which it is not operative.

Unit 3 ❐ The Theory of Production

Structure

3.0 Objectives

The objectives of this unit are—

- To explain production functions, laws of production and their presentation, alternative forms of production functions;
- To explain equilibrium of a firm- least cost combination of inputs. Returns to scale and expansion path
- Cobb Douglas and CES Production function, Elasticity of substitution
- Problem solutions

3.1. Production Function : An Introduction

In Economics, the term production means transformation of inputs into outputs. We assume that production takes place only in firms. By a firm we mean a technical unit in which inputs are employed and outputs get produced. An input is a good or service that goes into the process of production. According to Professor Baumol, "An input is simply anything which the firm buys for use in its production or other processes". In other words, an input is any good or service which contributes to the production of an output. Economists have classified inputs as : (i) Land (N), (ii) Labour (L), (iii) Capital (K) and (iv) Organizatoin (O). Of these four basic inputs the two crucial ones are : L and K. Labour is a crucial factor as labour can speak, it can construct and destruct also. On the other hand, capital is another crucial factor as lack of capital is behind the backwardness of a country. Some inputs such as labour and land are not produced and these are known as primary inputs. An output may be tangible or intangible.

Nature of inputs : Inputs may be either fixed or variable. Fixed and variable inputs are defined in economic sense and in technical sense also. In economic sense, a fixed input is one whose supply is inelastic in the short run, and is used in a fixed quantity in the short run. On the other hand, in technical sense, a fixed factor is one that remains fixed (or constant) for all levels of output.

A variable input, on the other hand, is defined as one whose supply in the short-run is elastic. Thus labour and raw materials are variable inputs. All users of such factors can employ more quantity in the short run. In technical terminology, a variable input is one that changes with change in output. In the long run, however, all inputs are variable.

3.2. Production Function

The key concept in the theory of production is the production function. The word 'function' as used in Mathematics expresses the relationship between the dependent and the independent variables. Alternatively, a function shows a cause and effect relationship or an engineering relationship between inputs (independent variables or cause) and output (dependent variable or effect). Using one-output and two-input function, the relationship between output and inputs has been summarised in the following form :

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

Q being the quantity of output produced; L and K denote the respective quantities of labour and capital. By using L and K in a production process, we obtain Q or output. The engineer or the technician of an establishment describes the production process in terms of three or more variables (like density, pressure and horsepower). The economist's production function incorporates the engineering technology. Production engineers tell us how many units of L and K are needed to yield output by using a particular process of production. However, which particular production process (out of a number of alternatives) will be chosen by the firm depends on the prices of inputs and output that are bought and sold in the market. An isoquant gives a consont level of output for atternative combinalism or L and K. Therefore, a production function, being a technical one, presents physical combinations only. To an economist, however, production function is something more. A production activity shows costs for using inputs and revenues for output sold. Since the production function is given in the form of a table showing physical combinations of various inputs to obtain some units of output, it is not strictly within the domain of Economics. Economists are concerned not with the physical combinations but with the costs, revenue, output behaviour in respect to changes in inputs used, etc. Whenever there is a change in input in a production process, output changes. As a consequence various laws governing output have been obtained.

The form of a production function of an establishment gets determined by the nature of technology whose size, naturally, depends on the time span. This implies that a short run production function is not the same thing as the long run production function. We shall, therefore, consider the short run and the long run production functions separately.

3.2.1 Fixed Proportion Production Function

Fixed proportion production function can be expressed as $Q = \text{Min}\left[\frac{x_1}{a_1}, \frac{x_2}{a_2}\right]$.

This implies that the amount of output that can be produced equals the minimum of the ratios, namely $\frac{x_1}{a_1}$ and $\frac{x_2}{a_2}$. However, if $\frac{x_1}{a_1} = \frac{x_2}{a_2}$, both inputs, namely x_1 and x_2 will be fully utilised. But if $\frac{x_1}{a_1} \neq \frac{x_2}{a_2}$ one input will be fully utilised while some amount of the other will be redundant. We assume that the two inputs are always employed in a fixed proportion. In such a case there is only one ratio in which the two factors, namely x_1 and x_2 can be employed. We assume that $x_1 : x_2 = a_1 : a_2$, where a_1 and a_2 are constants. That means $\frac{x_1}{x_2} = \frac{a_1}{a_2}$ whence $x_2 = \left(\frac{a_2}{a_1}\right)x_1$. In the following **Fig.** 3.1 all points on the line OA satisfy the relation $x_1 : x_2 = a_1 : a_2$. Naturally, in this case there is only one process of production available to the producer; there is no substitution between the two factors in the production process. By a process of production is meant a ratio in which the two inputs get used.

Let us consider a production function $Q = f(x_1, x_2)$, x_1 and x_2 are two inputs used for production. Herein comes the concept of an isoquant or a production indifference curve or an equal product curve which is the locus of various combinations of x_1 and x_2 that produce the same level of output Q. Assuming strict complementarity between the inputs, i.e., no substitutability of factors of production, that is, factors of production are to be employed in a fixed proportion, the isoquants will be L-shaped as shown in figure 3.1

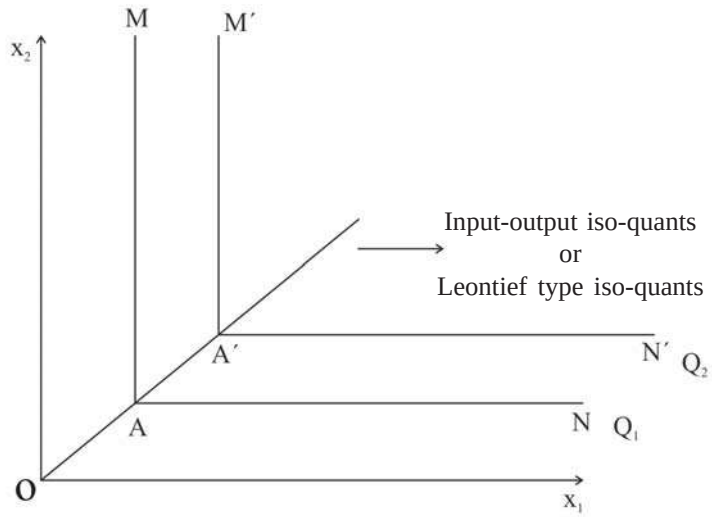


Fig : 3.1 : L-shaped iso-quants

In figure 3.1 "MAN" is the isoquant for the output level Q_1. For higher level of output we shall get a higher isoquant that lies to the right of the initial isoquant. $M^1A^1N^1$ is another such isoquant. Obviously, only the points A or A^1 are efficient points of production in the sense that at these points both factors of production are fully utilised. Any rational entrepreneur will always use the level of inputs given by the points A, A^1 etc. Besides, only on these points total cost of producing a given level of output will be minimum. That means, in the long run all inputs are variable.

In spite of such a difference between short run and long run, the actual length of time, that is, whether short run or long run, does not refer to any calender time as it varies from a firm to another firm or from one industry to another industry.

In consequence of the above differences between short run and long run we get two types of production function : (a) Short run production function and (b) long run production function. A short run production function assumes capital to be fixed while labour to be variable. So the short run production function takes the form

$$P = f(L, \bar{K}).$$

Here labour or L is a variable input while capital or K is assumed to be fixed.

In the short run production is subject to the law of variable proportions whereas production in the long run is subject to the law of returns to scale.

3.3 The Law of Variable Proportions or the Law of Diminishing Returns to Factors

As told above, a firm, in the short run, increases output only by varying variable inputs. So the firm's production function becomes

$$P = f(L, \bar{K}), \ \bar{K} \text{ being fixed.}$$

For various values of L, we get different values of output, Q. The total product curve of labour expresses the relation between L and K. From this we also get the idea of average product (AP) and marginal product (MP).

A firm, in the short run, goes on increasing its level of production by making changes in input-mix. Here we are interested in learning the returns to a variable input. Changes made in the input-mix and their impact on output have been studied under an empirical and celebrated law in Economics, that is, the law of variable proportions or the law of non-proportional returns or the law of diminishing returns to factors.

3.4.1 Statement of The Law

The law states that under a given state of technology if the quantity of one factor of production, say, labour is increased by equal increments while the quantity of the other factor of production, say, capital, remains fixed, the consequent increment of total product will, eventually, decline (that is, total product will first increase but decrease after some point). This law has also been known as the law of eventually diminishing returns to factors. More explicitly, the law states that if we go on increasing more and more of a variable input (say, labour), the amount of other inputs, say, capital being held fixed, the returns to the variable input will become non-proportional : initially, it will show increasing returns, then constant returns for a while and eventually diminishing returns. In brief, as additional units of a variable input are added to a given amount of fixed input, the average and marginal products of the variable input will, eventually, come to decline.

Let us clarify the meanings of TP, AP and MP.

TP : Various values of a variable input, given a specified amount of the fixed input, will give us various values of output. Total output is usually called the total product (TP). A typical TP curve shows that initially it increases slowly, then more rapidly and thereafter more slowly again until it finally reaches a maximum level and then begins to decline.

AP : The average product (AP) of an input is the total product per unit of the variable input which is needed to produce this output. It is obtained by dividing total product by total units of the input. In symbols, the average product of labour (L) input will be :

$$AP_L = \frac{TP}{L}.$$

The AP curve rises initially, reaches a maximum and thereafter declines.

MP : The marginal product (MP) is the change in total product resulting from a change in the use of the variable inputs. In symbols,

$MP_L = \frac{\Delta TP}{\Delta L}$ = change in total product resulting from the change in labour.

This may also be defined as

$MP_L = \Delta TP_n - \Delta TP_{(n-1)}$.

That means, MP is the difference in total product of n units and that of (n – 1) units.

Like AP, MP initially rises, reaches a peak and thereafter, declines. However, MP can be zero and negative if the MP curve touches and cuts the horizontal axis, respectively.

Therefore, the returns to the variable input are non-proportional. Initially, we obtain increasing stage where TP, AP and MP are on the rise. Then we have the diminishing stage where TP increases at a slower rate and both AP and MP decline. Finally, we have a negative stage when TP decreases and MP becomes negative.

3.4.2 Graphical Presentation of the Law

The above-stated law can, neatly, be presented in the following diagram (Fig. 3.2), measuring the variable input, labour, along the horizontal axis (assuming the amount of capital used is fixed) while total product (TP), average product (AP) and marginal product (MP) have been measured along the vertical axis.

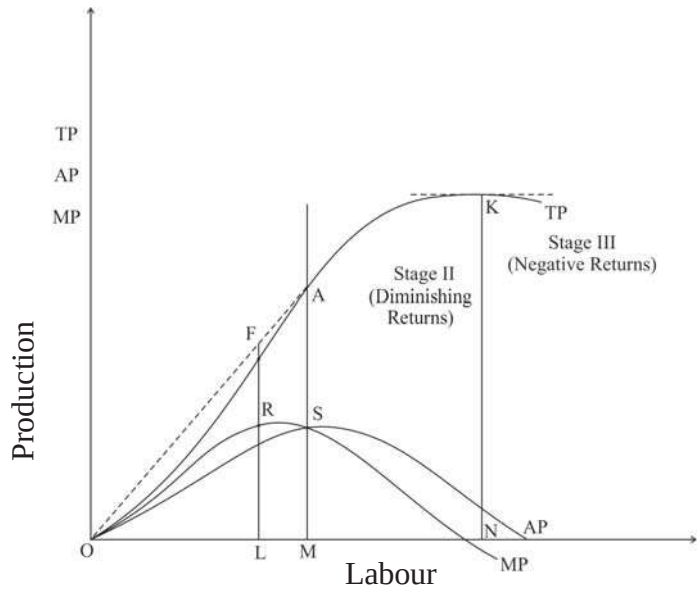


Fig. 3.2: Law of Variable Proportions

Let us concentrate our attention first on the TP curve. Since labour is increased relative to the fixed inputs, the TP curve rises very rapidly, reaching its maximum slope at point F, the point of inflection. Upto the point of inflection, TP rises at an increasing rate. TP starts declining after the point of inflection has been reached. MP is the slope

of TP. Therefore, the maximum slope of TP (i.e., point F) must correspond to the maximum on MP (that is, point R). Once the employment of labour exceeds OL amount, TP increases at a diminishing rate, that is, its slope diminishes. When OM unit of labour is employed, AP becomes maximum. Since tangency of the dotted ray to the TP curve at point A defines the condition for highest AP, point S, thus, lies directly below point A.

Beyond OM, TP continues to rise, though, at a diminishing rate and reaches maximum at point K. Here the slope of TP becomes zero and hence, MP becomes zero (corresponding to ON volume of labour employed).

Beyond ON, if labour is employed, TP will decline and MP will become negative.

Let us now explain the relations among TP, AP and MP in the above-mentioned three stages of production.

Stage I is known as the stage of increasing returns where TP increases at an increasing rate and hence, AP and MP rise. This is the stage of increasing returns as in this stage AP of labour rises throughout. However, in this stage MP increases though declines after OL amount of labour gets employed. Throughout this stage, however, MP exceeds AP.

Stage II is called diminishing stage since here both AP and MP decline, though positive. However, this is the most crucial stage as far as decision to produce is considered.

Stage III is known as the stage of negative returns where TP declines and MP becomes negative.

Let us now find out the relationship between AP and MP in the three stages of production :

If MP > AP, AP will rise as labour increases.

If MP < AP, AP will decline as labour increases.

If MP = AP, then AP will be maximum.

In stage I, MP > 0 and AP is rising. Thus MP > AP. This is the increasing stage.

In stage II, MP > 0 but AP is falling. Thus MP < AP but TP is increasing because MP > 0, this is the diminishing stage.

In stage III, MP < 0 and TP is falling, this is the negative stage.

From the above analysis we are now interested in answering a question : Out of the three stages, where will the firm produce? Evidently, no rational profit-making firm would produce either in stage I or in stage III. A rational producer will always like to produce only in the second stage where both AP and MP of the variable input (that is, labour) as well as AP and MP of fixed inputs are positive.

3.5 Long Run Production Function Returns to Scale

Now we are interested in studying the effect of changes in output when all factors of production in a particular production function are changed together. That means we intend to study the behaviour of output in response to a change in the scale. A change in the scale means that all factors of production are increased or decreased in the same propotion. Changes in scale are different from changes in factor proportions.

RELATION BETWEEN AVERAGE PRODUCT (AP) AND MARGINAL PRODUCT (MP)

There is a unique relation between AP and MP and it is true for all "average" and "marginal" productivity conditions. This relation is as follows :

(a) When AP is rising, MP > AP.

(b) When AP is maximum, MP = AP.

and (c) When AP is falling, MP < AP.

The above relation may be established in the following manner :

We have $AP_L = \frac{Q}{L}$.

Differentiating AP_L partially with respect to L, we get

$$\frac{\partial}{\partial_L}(AP_L) = \frac{\partial}{\partial_L}\left(\frac{Q}{L}\right) = \frac{\frac{L\partial Q}{\partial_L} - \frac{Q\partial_L}{\partial_L}}{L^2} = \frac{\frac{L\partial Q}{\partial_L} - Q}{L^2}.$$

$$\text{or, } \frac{\partial}{\partial_L}(AP_L) = \frac{\partial Q}{\partial_L}.\frac{1}{L} - \frac{Q}{L^2}.$$

$$\text{or, } \frac{\partial}{\partial_L}(AP_L) = \frac{1}{L}\left[\frac{\partial Q}{\partial_L} - \frac{Q}{L}\right] = \frac{1}{L}\ [MP_L - AP_L] \ldots\ldots(i)$$

From equation (i) $\frac{\partial}{\partial_L}(AP_L) > 0$ when $MP_L > AP_L \because L > 0$;

$\frac{\partial}{\partial_L}(AP_L) = o$ when $MP_L = AP_L$;

Finally, $\frac{\partial}{\partial_L}(AP_L) < 0$ when $MP_L < AP_L \because L > 0$.

It is an empirically observed feature that all inputs have positive but diminishing marginal products. If L and K are the only factors of production $MP_L > 0$ and $MP_K > 0$.

Thus for any factor, initially, MP is positive. Then a situation arises when MP is zero and further on it falls as more of the factor is employed in production. Changes in output as a result of the variation in factor proportions form the subject matter of the law of variable proportions, whereas the study of changes in output as a consequence of changes in scale forms the subject matter of returns to scale. The study of changes in output following a change in the scale of production results in a "law of returns to scale", "which has, obviously, to be distinguished from the law of returns to a variable input." A firm can, obviously, change its scale of operation only if a sufficiently long time is allowed. Therefore, the "law of returns to scale" forms the subject matter of long run production function.

If all inputs are increased by the same proportion, thereby keeping factor proportions unaltered, there occurs an increasse in scale. The percentage increase in output due to one percent increase in all inputs determines the returns to scale. Let the one-output-two-variable input production function be given by

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

If, now, these inputs are exactly doubled, output may be doubled. In this case the returns to scale are said to be constant. Thus

$$2Q = f(2L, 2K)...(b)$$

However, if doubling or trebling of inputs causes output to increase by more than that, then there will be increasing returns to scale. Finally, returns to scale will be said to be decreasing, when doubling of inputs results in turning out an output less than that.

As we increase the scale of production initially the increase in output becomes subject to increasing returns, then constant returns and ultimately diminishing returns to scale. At the beginning, if we increase all inputs by one percentage point, output will increase by more than one percent. Here we have the stage of increasing returns to scale (IRS). Next, if the

rate of increase in output remains constant following a one percent increase in input then we have the stage of constant returns to scale (CRS). Ultimately, the firm experiences diminishing returns to scale (DRS) as the scale of production rises.

It is to be noted that what is called CRS in economic theory, in Mathematies we call it linear homogeneous production function or homogeneous production function of degree one. This at once suggests that "linearly homogeneous" and "constnat returns to scale" are two interchangeable terms.

For an easy understanding of the relation between stages of returns to scale and change in output we have used the following figure

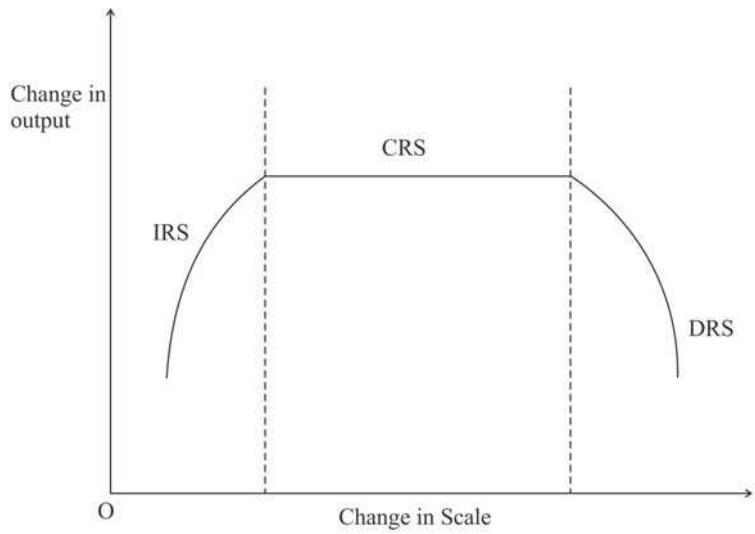


Fig. 3.3 Stages of Returns to Scale

3.6 Isoquant or Isoproduct Curve : Two-Variable Inputs

We have so far concentrated our attention on only one variable input, namely labour. We say, now, that to produce a commodity we are in need of two inputs (that is, two- variable inputs). Moreover, we assume that these two inputs are substitutable throughout the production process. We can represent these two substitutable input production function in a diagram and get a curve called isoproduct curve or equal product curve or production indifference curve or an isoquant, which is similar to an indifference curve in the theory of consumer behaviour (made quite popular by Nobel Laureate Economist Professor J. R. Hicks

and R. G. D. Allen). We may mention here that modern economists analyse the theory of production in terms of an isoquant.

Let our production function be : P = f(L, K).

We assume that this production function is continuous and differentiable. Besides, the two factors are substitutable : we can employ more of one factor and less of the other to get the same amount of output. The various combinations of L and K that can produce a given level of output, when plotted on a graph paper, give us an equal product curve or an isoquant [We are to note that the Greek word 'iso' means 'equal' or 'same']. Therefore, an isoquant is a curve along which output is same. Supposing that the output level is constant at Q^0, the equation of the isoquant will be given by $Q^0 = f(L, K)$, Q^0 being the fixed amount of output to be produced.

For one possible level of output we get one isoquant. A higher level of output is indicated by a higher isoquant while a lower output level is indicated by a lower isoquant. This is a smooth isoquant which assumes continuous substitutability between labour and capital.

In figure 3.4 capital has been measured along the vertical axis while labour is measured along the horizontal axis. We have drawn an isoproduct curve represented by a curve IQ100. Point A, for example, on IQ100 represents just one possible combination of labour and capital that can be utilized for producing 100 units of output. Point B is another possible combination of labour and captial to produce the same level of output. However, at point B, more labour and less capital have been used. Points C, D and E have similar meanings. By joining these points by a continuous curve we get an isoquant. It is to be remembered that each and every point on an isoquant represents the same level of output.

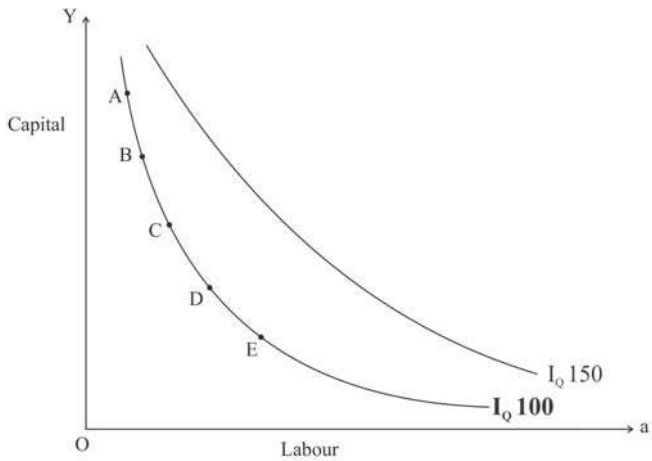


Fig. 3.4 : Isoquant

Therefore, an isoquant represents various combinations of two inputs, namely labour and

capital that yield the same level of output. An isoquant is a contour line showing different combinations of inputs that are physically capable of producing a given amount of output. A family of isoquant is called an isoquant map. Higher-order isoquant represents higher level of output. By utilising more labour and more capital, a firm can produce larger output. For example, I_Q150 lies above I_Q100.

An isoquant is analogous to an "indifference curve" having two points of distinction : (a) An indifference curve is made of two consumer goods whereas an isoquant is constructed by using two producer goods, namely labour (L) and capital (K) and (b) an indifference curve measures "Utility" not measurable, while an isoquant measures "output" which is always measurable.

3.6.1 Properties of an Isoquant

Like an indifference curve in the theory of consumer behaviour, an isoquant has the following properties :

(i) An isoquant normally slopes downward from left to right. Or, an isoquant has a negative slope.

(ii) Isoquants are convex to the origin.

(iii) No two isoquants have a common point. That means, isoquants cannot intersect or be tangent to each other.

(iv) Higher isoquants represent higher level of output.

Let us prove these properties in the simple possible manner :

An isoquant is negatively sloped : If we move along an isoquant the amount of one input should increase while that of another must decrease. If both labour and capital yield positive marginal products, then increasing the number of labour while holding the number of capital constant, will enhance output. Therefore, if we want to maintain a constant amount of output when the quantity of labour (or capital) gets increased, the amount of capital (or labour) must decrease. This naturally implies a negatively sloped isoquant.

An isoquant is convex to the origin : In course of our analysis of isoquant we assume that the two inputs which are needed for producing a given level of output are substitutes, of course, not perfect substitutes. Therefore, by virtue of the assumption of the diminishing "marginal rate of technical substitution between labour and capital" ($MRTS_K$ for L/$MRTS_L$ for K) an isoquant is convex to the origin.

No two isoquants have a common point :

Proof : If two isoquants intersect or touch each other, we will encounter an absurd result. In Fig 3.5 we have drawn two intersecting isoquants that have cut each other at point C. In such a case we shall show that this will lead to an absurd result.

In the diagram (Fig. 3.5), points C and A lie on IQ_1. As points C and A lie on IQ_1, we can produce the same amount of output at both these points. In the same way we can produce the same amount of output at both points C and B. Thus, along points C and A we have Ok_1 amount of capital + OL_1 amount of labour = OK_2 amount of capital + OL_2 amount of labour. Similarly, OK_1 amount of capital + OL_1 amount of labour = OK_2 amount of capital + OL_3 amount of labour.

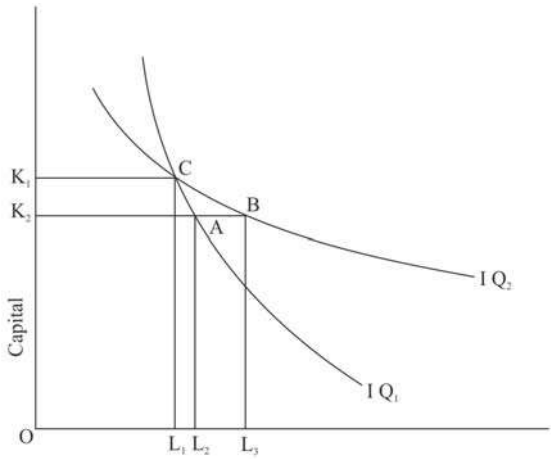


Fig. 3.5 : Labour

Equating the R.H.S. of the above two expressions,

we get OL_2 amount of labour = OL_3 amount of labour. However, this is impossible.

Hence we can say that two isoquants can not intersect each other.

3.6.2 Isocost Line

Another important concept in the theory of production is Isocost line or equal cost line which represents the prices of two factors of production, namely labour and capital. Naturally, the cost of production is the sum of all costs on purchasing inputs (here labour and capital) required to produce some amount of output. Such cost information can be obtained from the isocost line similar to the budget line discussed in the context of demand analysis. An isocost line expresses different combinations of two inputs, namely labour and capital that can be purchased from a given amount of money.

We assume that a typical firm uses two inputs, namely labour (L) and capital (K) whose respective prices are wage (W) and rate of interest (r). If 'C' representes the total cost of production for employing various units of L and K, then the equation of the isocost line will be

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

We suppose that the firm decides to purchase input K only. Then

C – W.L. = r.K so that $K = \frac{C}{r} - \frac{W}{r}.L.$

From the above expression we note that with given amount of money at the disposal of the firm, $\frac{C}{r}$ will be the maximum amount of capital that can be bought and utilised by the firm. This means that $\frac{C}{r}$ is the ordinate intercept of the isocost line. Also, $\frac{W}{r}$ is the slope of the line or price-ratio of the two inputs and this slope is negative. In the same way if the entire amount is spent on purchasing labour only, then the firm can purchase maximum amount of labour, i.e., $\frac{C}{W}$. Solving for L we get $L = \frac{C}{W} - \frac{r}{W}.K$. Measuring labour on the horizontal axis we note that $\frac{C}{W}$ will indicate the horizontal intercept.

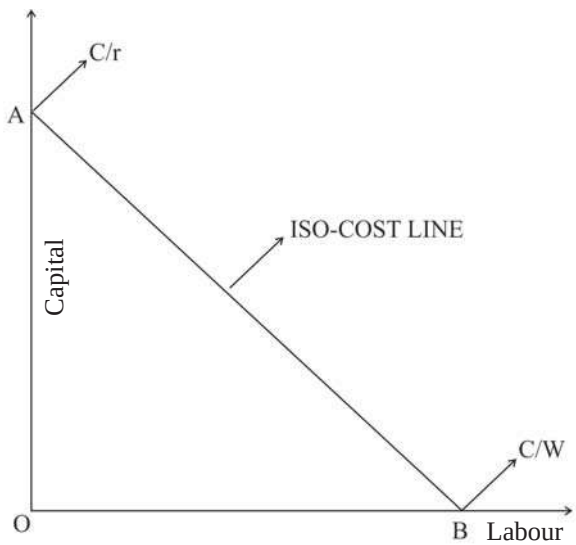


Fig. 3.6 : ISO Cost Line

Measuring all the necessary information on a graph paper we get the isocost line AB shown in the annexed figure 3.6. In terms of the isocost line AB, point B is nothing but $\frac{C}{W}$, that is, the horizontal intercept. In the same way, point A is nothing but $\frac{C}{r}$, that is, the vertical intercept. A firm, however, goes on purchasing both the inputs, namely labour and capital. However, between points A and B, there are many purchasable combinations of capital and labour. If all these purchasable combinations of inputs are joined together, we will get the isocost line, AB. Therefore, an isocost line represents alternative combinations of capital and labour that may be collected with a given money expenditure.

An increase in monetary expenditure will cause the isocost line to shift to the upward direction. However, there will be an inward shift of the isocost line if there is a decline in monetary outlay.

The slope of the isocost line is –OA/OB = $-\frac{C}{r}\Big/\frac{C}{W} = -\frac{W}{r}$.

In fine, we may say that if there is a parallel shift of the isocost line following a change in total expenditure, the slope of the equal cost line will remain the same. But, a change in the relative input prices will lead to a change in the slope of this line.

3.7 Equilibrium of a firm : Least-Cost Combination of Inputs

Now, we are interested in studying the optimal combination of two inputs to be used by the firm. To put it in another way : how should a firm choose the correct or optimal combination of the two inputs?

We know that the primary objective of a firm is to maximise profit which is the difference between total revenue (TR) and total cost (TC). Since revenue is reaped through markets, it is better not to talk about that at this moment. So we are to think of how through total cost, profit can be maximized. One possible way is to make the total cost as minimum as possible. There are two alternative ways to perform this job : (a) The firm can maximise output subject to a given cost or (b) The firm can minimise cost subject to a given output. It is to be remembered here that a firm cannot achieve the situation of maximum output and minimum cost simultaneously. Maximum output can be attained only at a given cost or, minimisation of cost situation will emerge at a given output level. So we are to consider both these two situations separately.

Situation 1 : Output maximisation subject to a cost constraint.

Let us suppose that the firm decides on a given cost level C^0. With this cost, the firm can purchase various combinations of the said two factors of production. As told earlier, all these combinations will lie on the iso-cost line, C^0. The objective of the firm is to maximise the level of output while remaining on this given iso-cost line. Fig 3.7. has demonstrated this picture.

We assume that the firm remains on the iso-cost line AB. Naturally, the firm can purchase any combination of the two inputs labour and capital lying on the line AB, because all of them are equally expensive. However, by moving from P_2 to P_1 the firm can increase the level of output as P_1 is on a higher isoquant compared to P_2. In the same way, by moving from P_1 to P, the firm will be able to enhance the level of output further. We see that P is the point of tangency between an isoquant and the iso-cost line. Further, at P the highest possible level of output can be achieved. We are to note that the movement from

P to P_3 is not, at all economic as we move from a higher isoquant to a lower one. As a rational producer the basic objective of the firm is to maximise output subject to the cost constraint. Geometrically, this means that the firm tries to attain the highest possible isoquant while remaining on the given iso-cost line. This can be achieved when one of the isoquants (here q^2) becomes tangent to the given iso-cost line.

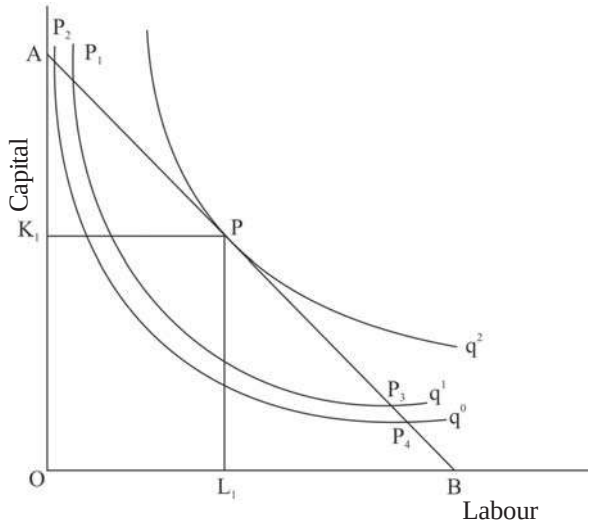


Fig. 3.7

The necessary condition for the maximisation of output is that the iso-cost line must be tangent to one of the isoquants. That is, the slope of the isoquant must equal the slope of the iso-cost line. We have already seen that the slope of the isocost line is $-\frac{W}{r}$. The slope of the isoquant equals the negative of the ratio of marginal productivities, i.e., $-\frac{\partial q/\partial x_1}{\partial q/\partial x_2}$.

Hence, the necessary condition for maximisation of output subject to a cost constraint is :

$-\frac{\partial q/\partial x_1}{\partial q/\partial x_2} = \frac{W}{r}$. That, is, $MRTS_{KL} = \frac{MP_L}{MP_K} = \frac{W}{r}$.

That is, the ratio of marginal productivities of two factors of production must equal the ratio of factor prices.

This condition is known as the necessary condition or, mathematically, the first order condition.

The above condition can also be written as $\frac{\partial q/\partial x_1}{W} = \frac{\partial q/\partial x_2}{r}$.

Here, $\frac{\partial q / \partial x_1}{W}$ is the amount of output that can be obtained by spending one unit of money in purchasing the factor L. In the same way, $\frac{\partial q / \partial x_2}{r}$ is the amount of output that can be obtained by spending one unit of money in purchasing the factor K. As they are equal, we may say that the firm gets the same amount of output by spending one unit of money either on L or on K.

Besides, there is another condition known as the sufficient condition or, mathematically, the second order condition which, in plain language, means that the isoquant must be convex to the origin at the point of equilibrium. This condition is, again, fulfilled at point P in diagram 3.7.

Situation 2 : Cost minimisation subject to an output constraint :

An alternative to maximising output for a given cost constraint is minimisation of cost subject to a stipulated level of output, say q^0. Under this least cost combination condition, a firm faces more than one isocost line and only one isoquant to represent the stipulated volume of output, q^0.

We know that the same level of output can be produced by various combinations of the said two factors of production. The locus of all such combinations is an isoquant representing the output level q^0. As per the objective of the firm, the firm will, always, remain on this isoquant. The problem of the firm is to select a point on this isoquant which is least costly. This problem can be explained and solved with the help of the figure 3.8.

Under this least cost combination condition, the firm faces more than one iso-cost line and only one isoquant q^0 to represent the stipulated volume of output. From the diagram it is clear that the firm can produce q^0 amount of output staying at any point such as P_2, P_1, P_3 and P_4 each lying on the isoquant. However, the cost level at P_1 is less than that at P_2. Here P is the point on the lowest possible iso-cost line. When we move from P_2 to P_1 we substitute labour for capital, such a substitution is profitable as total cost gets reduced as a result of this substitution. However, once we reach the point P no further substitution is profitable since total cost increases as we move to the right of P. Therefore, only at point P, the cost of producing the output level q^0 is minimum.

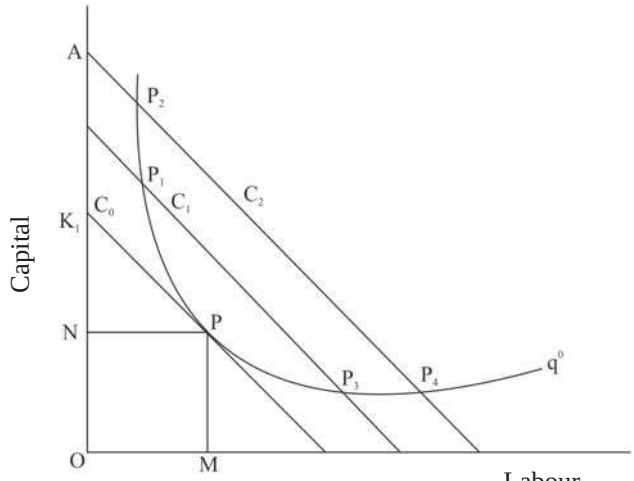


Fig. 3.8

Geometrically, the principle of cost minimisation requires that the firm under consideration tries to reach the lowest possible iso-cost line while remaining on the given isoquant. For this the necessary condition is that the isoquant will be tangent to one of the iso-cost lines. That means, when the slope of the iso-quant and that of the iso-cost line are identical, that is, $\frac{\partial q / \partial L}{\partial q / \partial K} = \frac{W}{r}$.

That is, the ratio of marginal productivities = Ratio of factor prices.

In addition to the fulfilment of the necessary or first order condition for equilibrium, we need another condition known as the sufficient condition or the second order condition which states that the isoquant must be convex to the origin. Thus point P is the equilibrium point.

3.8 Expansion Path

A highly relevant concept in the theory of production is the "expansion path" which is defined to be the locus of all input combinations for which the marginal rate of technical substitution equals the factor price ratio. If there are two factors of production and if their prices are constants, we will get a number of parallel iso-cost lines.

Each of these iso-cost lines will be tangent to one of the isoquants. The locus of all these points of tangency between isoquants and parallel iso-cost lines is the expansion path of the firm concerned. Needless to say, the points on the expansion path are the most efficient combinations of the two factors of production needed for producing a commodity. The

expansion path is also known as the scale-line because it shows how an entrepreneur will change the quantities of the two factors when it goes on increasing the level of output. Since an expansion path represents the minimum cost combinations for various levels of output, it shows the cheapest way of producing each level of output, given the relative prices of the two factors.

Along an expansion path, the cost of production increases, though the input prices remain the same. Thus there will be an output effect along an expansion path. The expansion path will be upward rising if both factors of production are non-inferior. A firm desirous of producing more output will employ more of both the factors. The first order condition of cost minimisation gives the equation of the expansion path which is

$$\frac{MP_L}{MP_K} = \frac{P_L}{P_K} \quad \text{or,} \quad \frac{f_1}{f_2} = \frac{W}{r}.$$

For expanding its level of output, the firm will select only those combinations of the two inputs which lie only on the expansion path. The expansion path will be upward rising if both the factors are non-inferior. However, it may be backward bending or downward sloping after some point if one of the factors becomes inferior. It is worth mentioning that the expansion path of a linear homogeneous production function is a straight line through the origin. The Cobb-Douglas production function is a linear homogeneous production function. We shall see later that for this production function the expansion path is a straight line through the origin.

3.9 Elasticity of Substitution

A very important concept in connection with production function is ***"Elasticity of Substitution"***.

We know that an isoquant represents a fixed amount of product for various levels of factors used in the production and its slope represents the marginal rate of technical substitution between the factors. The slope of the isoquant after a particular point (when the marginal product of the factor on the relevant axis becomes zero) becomes positive which, in turn, means that the substitution between the factors is no more possible. In other words, we may say that the slope or the curvature of the isoquant reflects the complementarity and substitutability between the two factors of production.

Every firm is interested in knowing the extent of substitutability between the factors of production in order to maximise output or minimise cost looking at the prices which the factors demand in the market. The measure of substitution between the factors is known as the

elasticity of technical substitution. This is a pure number and it measures the extent to which the substitution between the factors can take place. Since the substitution depends mainly on the slope of the isoquant (or RTS), elasticity of substitution is defined as the proportionate change in the ratio between the factors due to the propotional change in RTS.

Thus elasticity of substitution, σ, is defined as

Proportionate Change in the ratio between the factors
Proportionate change in the rate of technical substitution (RTS) .

Symbolically, elasticity of substitution between the two factors, say, labour (L) and capital (K) will be

$$\sigma = \frac{d\log\left(\frac{K}{L}\right)}{d\log\left(\frac{MP_L}{MP_K}\right)} = \frac{d\left(\frac{K}{L}\right)}{\frac{K}{L}} \Big/ \frac{d\left(\frac{MP_L}{MP_K}\right)}{\frac{MP_L}{MP_K}} .$$

The followings are the uses of the elasticity of substitution :

1. The magnitude of σ is an indication of the extent with which the same amount of product can be maintained by substituting one factor for another.

2. The higher the value of σ, the higher is the degree of substitutability between the factors.

3. The value of σ can vary between 0 and ∞.

4. As $\sigma \to \infty$, the shape of the isoquant tends to be flatter.

5. As $\sigma \to 0$, the shape of the isoquant tends to be such as makes an angle of 90^o.

6. When $\sigma = \infty$, the two factors are perfect substitutes and if $\sigma = 0$, the two factors are incapable of substitution.

3.10 Linearly Homegeneous Production Function

Production functions homogeneous of the first degree are commonly referred to as linearly (or linear) homogeneous production functions. In such a case the production function itself need not be linear at all.

This may be illustrated with an example.

Let $f(x, y) = 3\frac{x^2}{y} + \frac{y^3}{x^2}$ and it is definitely not a linear function.

But $f(Kx, Ky) = \frac{3(Kx)^2}{Ky} + \frac{(Ky)^3}{(Kx)^2} = 3\frac{K^2x^2}{Ky} + \frac{K^3y^3}{K^2x^2} = K\left[\frac{3x^2}{y} + \frac{y^3}{x^2}\right]$.

This shows that the above production function is linearly homogeneous.

Linearly homogeneous production functions are widely used in economic literature. They are not only mathematically simple but also possess some important properties :

1. The function obeys constant returns to scale (CRS). Thus, for any value of C, f(CL, CK) = C f(L, K) by the definition of linear homogeneity. This implies that a proportionate change in all inputs results in an equal proportionate change in output and this is what is meant by CRS.

2. The average product (of either input) depends upon the capital labour (K/L) ratio alone and is independent of the absolute amounts of inputs used.

Proof : From Q = f(L, K) we may write

$$\frac{Q}{L} = f\left(\frac{L}{L}, \frac{K}{L}\right) = f\left(1, \frac{K}{L}\right).$$

Now, $AP_L = \frac{Q}{L} = f\left(1, \frac{K}{L}\right) = \phi\left(\frac{K}{L}\right)$.

This means that the AP_L depends only on K/L ratio.

Similarly, $APK = \frac{Q}{K} = f\left(\frac{L}{K}, \frac{K}{K}\right) = f\left(\frac{L}{K}, 1\right) = \psi\left(\frac{L}{K}\right)$,

revealing that the APK will also depend on the $\frac{K}{L}$ ratio.

3. The marginal product (of either input) is the function of capital/labour ratio only.

Proof : We have seen that

$\frac{Q}{L} = \phi\left(\frac{K}{L}\right)$ so that $Q = L\,\phi\left(\frac{K}{L}\right)$.

Differentiating partially $Q = L\,\phi\left(\frac{K}{L}\right)$ with respect to L we get

$$\frac{\partial Q}{\partial L} = \frac{\partial}{\partial L}\left[L\phi\left(\frac{K}{L}\right)\right]$$

$$= L\,\phi^1\left(\frac{K}{L}\right)\frac{d}{dL}\left(\frac{K}{L}\right) + Q\left(\frac{K}{L}\right)$$

$$= L\,\phi^1\left(\frac{K}{L}\right)\left(-\frac{K}{L^2}\right)+\phi\left(\frac{K}{L}\right)$$

$$= \phi\left(\frac{K}{L}\right) - \frac{K}{L}\,\phi^1\left(\frac{K}{L}\right).$$

Thus $MP_L = \frac{\partial Q}{\partial L}$ will depend upon the $\frac{K}{L}$ ratio.

In the same way, $MP_K = \frac{\partial Q}{\partial K} = \frac{\partial}{\partial K}\,L\left[\phi\left(\frac{K}{L}\right)\right]$.

$$= L\,\phi^1\left(\frac{K}{L}\right)\left(\frac{1}{L}\right)+\phi\left(\frac{K}{L}\right).0$$

$$= \phi^1\left(\frac{K}{L}\right).$$

Therefore, MP_K also depends upon the input ratio alone.

4. The Euler's theorem for homogeneous function holds good. This means that for this production function $L\frac{\partial Q}{\partial L} + K\frac{\partial Q}{\partial K} = Q$.

Proof : $L\frac{\partial Q}{\partial L} + K\frac{\partial Q}{\partial K}$

$$= L\left[\phi\left(\frac{K}{L}\right) - \frac{K}{L}\phi^1\left(\frac{K}{L}\right)\right] + K\phi^1\left(\frac{K}{L}\right).$$

$$= L\,\phi\left(\frac{K}{L}\right) - K\phi^1\left(\frac{K}{L}\right) + K\phi^1\left(\frac{K}{L}\right) = L\phi\left(\frac{K}{L}\right) = Q.$$

Mathematically, Euler's theorem states that the value of a linearly homogeneous function can be shown as the sum of the terms, each of which shows the product of the value of a variable and the partial derivative of the function with respect to the same variable.

Economically, the theorem states that total product is the sum of the quantities of each input multiplied by their respective marginal products. That means, if every factor is paid according to its marginal product, total product will be fully exhausted. This is also known as the product exhaustion theorem.

3.11 The Cobb-Douglas Production Function

Probably the best known production function in Economics is the Cobb-Douglas production function which had been so named after the pioneering study in which Douglas fitted a function suggested by Cobb for the U. S. data.

This function takes the form :

$$Q = A\ L^{\alpha}K^{\beta}$$

where Q is the amount of output produced, L and K are labour and capital inputs respectively, A is a positive constant, α and β are positive fractions.

It is a very popular function and it has the following properties :

1. The function is non-linear. However, it is log linear because with the help of logarithmic transformation it can be written as

$\log Q = \log A + \alpha \log L + \beta \log K.$

Since it is log linear, it is very simple to handle.

2. The function is very often written in the form

$$Q = A\ L^{\alpha}\ K^{1-\alpha}.$$

In this special case where $\alpha + \beta = 1$, the function shows CRS.

Proof : From $Q = A\ L^{\alpha}\ K^{1-\alpha}$, we have

$$= A\ (CL)^{\alpha}\ (CK)^{1-\alpha}$$

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

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

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

$$= CQ.$$

It is not necessary that $\alpha + \beta = 1$ in the Cobb-Douglas production function. If the function is homogeneous of degree one, the CRS will prevail. If it is of degree less than one, decreasing returns to scale prevails. Similarly, if it is of degree greater than one, there are increasing returns to scale.

3. The function yields diminishing returns to each input.

Proof : From $Q = A\ L^{\alpha}\ K^{\beta}$ we get

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

$$= A\alpha\ L^{\alpha-1} K^{\beta}.$$

Also, $\frac{\partial^2 Q}{\partial L^2} = A\ \alpha\ (\alpha - 1) L^{\alpha-2}\ K^{\beta} < 0.$

Since α is a positive fraction, $(\alpha - 1)$ is negative. This means that the rate of change in the marginal product of input L is negative, implying that MP_L declines.

Similarly, $MP_K = \frac{\partial Q}{\partial K} = \beta\ A\ L^{\alpha} K^{\beta-1}$

$$= A\ \beta\ L^{\alpha} K^{\beta-1}.$$

Also, $\frac{\partial^2 Q}{\partial K^2} = A\ \beta\ L^{\alpha}(\beta - 1) K^{\beta-2} < 0.$

This suggests that MP_K also declines.

4. α and β show the output elasticity of inputs L and K respectively.

Proof : The output elasticity of an input is defined as the ratio of the proportionate change in the output resulting from the proportionate change in the input.

Thus, $e_{Q,L} = \frac{\partial Q}{Q} \Big/ \frac{\partial L}{L} = \frac{\partial Q}{Q} \times \frac{L}{\partial_L} = \frac{\partial Q}{\partial_L} \cdot \frac{L}{Q}$

$= \alpha A L^{\alpha-1} K^{\beta} \times \frac{L}{AL^{\alpha}K^{\beta}} = \alpha.$

In the same way,

$e_{Q,K} = \frac{\partial Q}{Q} \Big/ \frac{\partial K}{K}.$

$= \frac{\partial Q}{\partial K} \cdot \frac{K}{Q}.$

$= A\beta L^{\alpha} K^{\beta-1} \times \frac{K}{AL^{\alpha}K^{\beta}} = \beta.$

5. α and β show the relative distributive shares of inputs L and K respectively.

Proof : The relative distributive share of input L is given by

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

Therefore, the relative distributive share of input L is α.

Similarly, the relative distributive share of input K is β.

Since $\frac{K \times MPK}{Q}$

$= \frac{K.A\beta L^{\alpha}K^{\beta-1}}{AL^{\alpha}K^{\beta}} = \beta.$

6. Labour and capital are essential factors of production, that is, if either input is zero, output will also be zero.

7. The expansion path of the Cobb-Douglas production function is linear and it passes through the origin.

Proof : The Cobb-Douglas production function is $Q = AL^{\alpha} K^{\beta}$.

Thus, $\log Q = \log A + \alpha \log L + \beta \log K$.

Differentiating, partially, with respect to L and K we get

$\frac{1}{P}\cdot\frac{\partial P}{\partial L} = \frac{\alpha}{L}$ and $\frac{1}{P}\frac{\partial P}{\partial K} = \frac{\beta}{K}$.

Now, $MP_L = \frac{\partial P}{\partial L} = \alpha\frac{P}{L}$ and $MP_K = \frac{\partial P}{\partial K} = \beta\frac{P}{K}$.

The first order condition for the optimization requiers that $\frac{MP_L}{MP_K} = \frac{P_L}{P_K}$.

That is, $\frac{\alpha\frac{P}{L}}{\beta\frac{P}{K}} = \frac{P_L}{P_K}$

That is, $\frac{\alpha}{\beta}\cdot\frac{K}{L} = \frac{P_L}{P_K}$.

That is, $\alpha K P_K = \beta_L P_L$.

That is, $\alpha K P_K - L\beta P_L = 0$.

That means, the expansion path generated by the Cobb-Douglas production function is linear and it passes through the origin.

8. For this production function the elasticity of substitution is unity.

Proof : We know that the elasticity of substitution is

$$\sigma = \frac{\text{Percentage change in factor quantity ratio}}{\text{Percentage change in marginal productivity ratio}}.$$

$$= \frac{d\log\left(\frac{K}{L}\right)}{d\log\left(\frac{MP_L}{MP_K}\right)}$$

$$= \frac{d\frac{K}{L}}{\frac{K}{L}} \Bigg/ \frac{d\left(\frac{MP_L}{MP_K}\right)}{\left(\frac{MP_L}{MP_K}\right)}$$

For this production function $\frac{MP_L}{MP_K} = \frac{A\alpha L^{\alpha-1}K^{\beta}}{AL^{\alpha}\beta K^{\beta-1}} = \frac{\alpha}{\beta}\cdot\frac{K}{L}$.

Thus, $d\left(\frac{MP_L}{MP_K}\right) = \frac{\alpha}{\beta}d\left(\frac{K}{L}\right)$.

Hence, $\sigma = \frac{d\left(\frac{K}{L}\right)}{\frac{K}{L}} \Big/ \frac{\frac{\alpha}{\beta}d\left(\frac{K}{L}\right)}{\frac{\alpha}{\beta}\cdot\frac{K}{L}}$

$= \frac{d\left(\frac{K}{L}\right)}{\frac{K}{L}} \cdot \frac{\frac{\alpha}{\beta}\left(\frac{K}{L}\right)}{\frac{\alpha}{\beta}d\left(\frac{K}{L}\right)}$

$= 1.$

This also means that if the production function is linear and homogeneous then the elasticity of substitution between the factors (say, labour and capital) will be one.

9. The Cobb-Douglas production function gives an example which exhibits constant returns to scale but at the same time there are diminishing returns to each factor of production.

3.11.1 Limitations of The Cobb-Douglas Production Function

Although the Cobb-Douglas production function has a wide use and scope in the field of Economics and its use is constantly increasing in diffrent sectors of agriculture and manufacturing industry, it has been criticized by different economists. The main critics of this function are K. J. Arrow, H. B. Chenery, B. S. Minhas and R. M. Solow. The followigs are the main criticisms :

1. The main drawback of this production function is that it contains only two inputs i.e., Labour and capital. But, in reality, other factors are equally important in the production techniques. So this function is inapplicable for multiple inputs.

2. The production function operates under constant returns to scale. The law of increasing

and decreasing returns to scale are also operating in the production process. However, this function is not applicable in these situations.

3. The function assumes that technological knowledge remains constant or there is no technological change in the production process. But the production scale can also change with the change in technology. Thus stagnant technology assumption of the function is unrealistic.

4. This function assumes that all inputs are homogeneous. In fact, all units of a factor are not equally efficient. For example, in the labour population some labourers are efficient while others may not be.

5. In the function, the parameters α and β represent the shares of labour and capital in total production respectively. It is only true when there is perfect competition in the market. But if there is imperfect competition or monopoly in the economy, the above relationship does not hold good.

6. The function does not give any information regarding inter-connection or relations between the production factors, thus it ignores the possibility of the negative marginal productivity of a factor.

7. Since it has considered only the positive marginal productivity of factors, it ignores the possibility of the negative marginal productivity of a factor.

3.12 The Constant Elasticity of Substitution Production Function [CES Production Function]

A more general type of production function having constant elasticity of substitution (not necessarily equal to one) is the CES production function having the form

$$q = A\,[\alpha\, x_1^{-\rho} + (1 - \alpha)x_2^{-\rho}]^{\frac{-1}{\rho}},$$

where $A > 0$, $0 < \alpha < 1$, ρ is the parameter.

From this production function we can find out the marginal productivity of two inputs x_1 and x_2.

Marginal productivity of x_1,

$$MPx_1 = f_1 = \frac{\partial q}{\partial x_1} = \frac{-A}{\rho}\left[\alpha x_1^{-\rho} + (1-\alpha)x_2^{-\rho}\right]^{-\frac{1}{\rho}-1}\left(-\rho_{\alpha x_1}{}^{-\rho-1}\right)$$

$$= A\alpha\left[\alpha x_1^{-\rho} + (1-\alpha)x_2^{-\rho}\right] - \left(\frac{1}{\rho}+1\right)\cdot x_1^{-(\rho+1)}.$$

$$= A\alpha x_1^{-(\rho+1)}\left[\alpha x_1^{-\rho} + (1-\alpha)x_2^{-\rho}\right]^{-\left(1+\frac{1}{\rho}\right)} = A\alpha x_1^{-(\rho+1)}\left[\frac{q}{A}\right]^{\rho+1} = \frac{A^{\alpha}}{A^{\rho+1}}\left(\frac{q}{x_1}\right)^{\rho+1}.$$

$$\left[\because \frac{q}{A} = \left[\alpha x_1^{-\rho} + (1-\alpha)x_2^{-\rho}\right]^{-\frac{1}{\rho}}\right.$$

$$\because \left(\frac{q}{A}\right)^{\rho+1} = \left[\alpha x_1^{-\rho} + (1-\alpha)x_2^{-\rho}\right]^{-\frac{1}{\rho}\times(\rho+1)} = \left[\alpha x_1^{-\rho} + (1-\alpha)x_2^{-\rho}\right]^{-1-\frac{1}{\rho}}\left.\right]$$

$$= \frac{\alpha}{A^{\rho}}\left(\frac{q}{x_1}\right)^{\rho+1} \quad ...(1)$$

Again, the marginal product of input x_2 is

$$MPx_2 = f_2 = \frac{\partial q}{\partial x_2}.$$

$$= -\frac{A}{\rho}\left[\alpha x_1^{-\rho} + (1-\alpha)x_2^{-\rho}\right]^{-\frac{1}{\rho}-1}.(1-\alpha).(-\rho)x_2^{-\rho-1}.$$

$$= A(1-\alpha)\left[\alpha x_1^{-\rho} + (1-\alpha)x_2^{-\rho}\right]^{-\frac{1}{\rho}-1} x_2^{-(\rho+1)} = \frac{1-\alpha}{A^{\rho}}\left(\frac{q}{x_2}\right)^{\rho+1} \quad ....(2)$$

$$\left[\because A(1-\alpha)\left[\alpha x_1^{-\rho} + (1-\alpha)x_2^{-\rho}\right]^{-\left(1+\frac{1}{\rho}\right)-(\rho+1)}\right.$$

$$= \frac{A(1-\alpha)}{x_2^{\rho+1}}\left[\alpha x_1^{-\rho} + (1-\alpha)x_2^{-\rho}\right]^{-\left(1+\frac{1}{\rho}\right)} = \frac{1-\alpha}{x_2^{\rho+1}} q^{\rho+1}$$

$$\because\ q^{\rho+1} = A^{\rho+1}\left[\alpha x_1^{-\rho} + (1-\alpha)x_2^{-\rho}\right]^{-\frac{1}{\rho}\times(\rho+1)}$$

$$= A^{\rho+1}\left[\alpha x_1^{-\rho} + (1-\alpha)x_2^{-\rho}\right]^{-\left(1+\frac{1}{\rho}\right)} \quad = \frac{1-\alpha}{x_2^{\rho+1}} \cdot \frac{q^{\rho+1}}{A^{\rho}} = \frac{1-\alpha}{A^{\rho}}\left(\frac{q}{x_2}\right)^{\rho+1}]$$

Therefore, $\frac{f_1}{f_2} = \frac{\alpha}{A^{\rho}}\left(\frac{q}{x_1}\right)^{\rho+1} \Big/ \frac{1-\alpha}{A^{\rho}}\left(\frac{q}{x_2}\right)^{\rho+1}$

$$= \frac{\alpha}{1-\alpha}\left(\frac{x_2}{x_1}\right)^{\rho+1}$$

Thus, $\log\left(\frac{f_1}{f_2}\right) = \log\ \frac{\alpha}{1-\alpha} + (\rho+1)\log\frac{x_2}{x_1}$.

That is, $d\ \log\left(\frac{f_1}{f_2}\right) = d\log\ \left(\frac{\alpha}{1-\alpha}\right) + (\rho+1)d\log\frac{x_2}{x_1}$.

$= d\ \log K + (\rho + 1)\ d\ \log\ \frac{x_2}{x_1}$, where $K = \frac{\alpha}{1-\alpha}$.

$= 0 + (\rho + 1)\ d\ \log\frac{x_2}{x_1} = (\rho + 1)\ d\ \log\ \frac{x_2}{x_1}$.

Now, elasticity of substitution, $\sigma = -\frac{d\log\frac{x_2}{x_1}}{d\log\frac{f_1}{f_2}} = \frac{1}{1+\rho}$ = constant.

Hence, $\sigma = \frac{1}{1+\rho}$ so that $(1 + \rho)\ \sigma = 1$.

That is, $\sigma + \rho\sigma = 1$.

That is, $\rho\,\sigma = 1 - \sigma$ whence $\rho = \frac{1-\sigma}{\sigma}$.

This reveals that the parameter ρ of the CES production function is closely related with the elasticity of substitution.

For the CES production function RTS $= \frac{f_1}{f_2} = \frac{\alpha}{1-\alpha}\left(\frac{x_2}{x_1}\right)^{\rho+1}$ and RTS will be decreasing if $\rho > -1$. Again, when $\rho > -1$, $\sigma > 0$. The isoquants generated from the CES production function will be downward sloping and convex. However, the particular shape will depend upon the value of ρ.

3.12.1 Few Limiting Cases of this Production Function

Some limiting cases of the above production function can now easily be pointed out :

1. If $\sigma \to 1$ and $\rho \to 0$ the CES production function is reduced to the Cobb-Douglas production function.

2. However, as $\rho \to \alpha$, $\sigma \to 0$; the RTS approaches zero if $x_1 > x_2$ or approaches ∞ if $x_1 < x_2$. In such a case no substitution is possible and the isoquants become L-shaped. Such a case happens with a production function having fixed coefficients.

3. If $\sigma > 1$, that is if $-1 < \rho < 0$ the isoquants will be convex to the origin and will cut both the axes. On the other hand, if $\sigma < 1$, that is, if $\rho > 0$, the isoquants are convex to the origin. However, in this case instead of cutting the two axes the isoquants will be asymptotic to both the axes.

4. If $\sigma \to \infty$, that is, if $\rho \to -1$, the isoquants become straight lines and inputs are perfect substitues of each other.

5. All these results show that the CES production function is a completely general type of production function from which isoquants of different shapes can be derived as special cases.

3.12.2 Some Problems and their Solutions

1. Suppose the production function of an imaginary firm is given by $Q = x_1x_2 - 0.2x_1^2 - 0.8x_2^2$. Assume $x_2 = 10$, determine the equations of the AP curve and of the MP curve

of x_1. Find at what values of x_1 will AP and MP be zero.

Solution : Putting the value of $x_2 = 10$, the given production function becomes $Q = 10x_1 - 0.2x_1^2 - 80$.

Now, AP of x_1= $AP_1 = \frac{Q}{x_1} = 10 - 0.2x_1 - \frac{80}{x_1}$ and MP of x_1 will be

$MP_1 = \frac{\partial Q}{\partial x_1} = 10 - 0.4x_1$.

When $AP_1 = 0$, $10 - 0.2x_1 - \frac{80}{x_1} = 0$.

That is, $10x_1 - 0.2x_1^2 - 80 = 0$.

That is, $2x_1^2 - 10x_1 + 80 = 0$. That is, $2x_1^2 - 100x_1 + 800 = 0$.

That is, $x_1^2 - 50x_1 + 400 = 0$. That is, $(x_1 - 10)(x_1 - 40) = 0$.

Thus when $x_1 = 10$ and 40, AP1 will be zero.

When $MP_1 = 0$, $10 - 0.4x_1 = 0$. That is, $x_1 = \frac{10}{0.4} = 25$.

2. Given the production function $Q = \sqrt{x_1 x_2}$ and prices per unit of x_1 and x_2 Rs.2/- and Rs.4/- respectively and total cost of output Rs. 80, obtain the maximum output subject to the cost constraint.

Solution : The theoretical expression of the cost function is given by $r_1x_1 + r_2x_2 = c$. Hence, here the equation of the cost line is $2x_1 + 4x_2 = 80$.

That is, the slope of the isocost line is $\frac{r_1}{r_2} = \frac{2}{4} = \frac{1}{2}$.

From the given production function $Q = \sqrt{x_1 x_2}$,

We get $f_1 = \frac{\partial Q}{\partial x_1} = \frac{1}{2} x_1^{-\frac{1}{2}} x_2^{\frac{1}{2}}$

Also, $f_2 = \frac{\partial Q}{\partial x_2} = \frac{1}{2} x_1^{\frac{1}{2}} x_2^{-\frac{1}{2}}$.

That is, the slope of the isoquant is $\frac{f_1}{f_2} = \frac{\frac{1}{2} x_1^{-\frac{1}{2}} x_2^{\frac{1}{2}}}{\frac{1}{2} x_1^{\frac{1}{2}} x_2^{-\frac{1}{2}}} = \frac{x_2}{x_1}$.

The necessary condition for maximisation of output is $\frac{f_1}{f_2} = \frac{\frac{1}{2} x_1^{-\frac{1}{2}} x_2^{\frac{1}{2}}}{\frac{1}{2} x_1^{\frac{1}{2}} x_2^{-\frac{1}{2}}} = \frac{x_2}{x_1}$.

For the maximisation of output the necessary condition is $\frac{f_1}{f_2} = \frac{r_1}{r_2}$.

That is, $\frac{x_2}{x_1} = \frac{1}{2}$.

That is, $x_1 = 2x_2$.

Putting $x_1 = 2x_2$ in the cost equation we get $2(2x_2) + 4x_2 = 80$.

That is, $x_2 = 10$ and $x_1 = 20$.

Let us put the above values of x_1 and x_2 so that we get

$Q = \sqrt{20 \times 10} = 10\sqrt{2} = 10 \times 1.414 = 14$ units (approx).

Thus, the maximum producible output is 14 units.

3. Obtain the equation of the expansion path when the production function is of the form

$Q = 8x_1^{\frac{1}{2}} + 20x_2^{\frac{1}{2}}$.

Solution : From the given production function

$Q = 8x_1^{\frac{1}{2}} + 20x_2^{\frac{1}{2}}$ we get

$f_1 = \frac{\partial Q}{\partial x_1} = 4x_1^{-\frac{1}{2}}$ and $f_2 = \frac{\partial Q}{\partial x_2} = 10x_2^{-\frac{1}{2}}$.

Hence, $\frac{f_1}{f_2} = \frac{4x_1^{-\frac{1}{2}}}{10x_2^{-\frac{1}{2}}}$.

Moreover, $\frac{r_1}{r_2} = \frac{1}{5}$.

The equation of the expansion path is given by the relation

$\frac{f_1}{f_2} = \frac{r_1}{r_2}$. That means, $\frac{4x_1^{-\frac{1}{2}}}{10x_2^{-\frac{1}{2}}} = \frac{1}{5}$.

That is, $20x_1^{-\frac{1}{2}} = 10x_2^{-\frac{1}{2}}$. That is, $2x_1^{-\frac{1}{2}} = x_2^{-\frac{1}{2}}$.

That is, $4.\frac{1}{x_1} = \frac{1}{x_2}$. That is, $x_1 = 4x_2$.

That is, $x_1 - 4x_2 = 0$.

Therefore, the equation of the expansion path is given by $x_1 - 4x_2 = 0$.

4. Suppose the production function is given by $Q = Ax_1^2x_2^2 - Bx_1^3x_2^3$,

show that the expansion path is given by

$r_1x_1 + r_2x_2 = 0$, r_1 and r_2 are the unit prices of x_1 and x_2 respectivly.

Solution : From $Q = Ax_1^2x_2^2 - Bx_1^3x_2^3$, we get

$f_1 = \frac{\partial Q}{\partial x_1} = 2Ax_1x_2^2 - 2Bx_2^3x_1^2$.

Also, $f_2 = \frac{\partial Q}{\partial x_2} = 2Ax_1^2x_2 - 3Bx_1^3.x_2$. Hence,

$$\frac{f_1}{f_2} = \frac{2Ax_1x_2^2 - 3Bx_1^2x_2^3}{2Ax_1^2x_2 - 3Bx_1^3x_2^2} = \frac{x_1x_2^2(2A - 3Bx_1x_2)}{x_1^2x_2(2A - 3Bx_1x_2)} = \frac{x_2}{x_1}.$$

The equation of an expansion path is given by $\frac{f_1}{f_2} = \frac{r_1}{r_2}$. That is, $\frac{x_2}{x_1} = \frac{r_1}{r_2}$. That is, $r_1x_1 = r_2x_2$ or, $r_1x_1 - r_2x_2 = 0$ Ans.

5. Assume the production function of a commodity is given as $Q = K^2 - 3KL + 4L^2$. Obtain the maximum amount of capital that can be employed when 7 units of labour are employed.

Soultion : For the employment of 7 units of labour the prodcution function becomes

$Q = K^2 - 21K - 4 \times 49$.

Let us now find the marginal productivity of capital.

That is, $MP_K = \frac{dQ}{dK} = 2K - 21$.

The maximum amount of capital that is to be employed is equal to that unit where the marginal productivity of capital is zero That is, $\frac{dQ}{dK} = 0$. That is, $2K - 21 = 0$.

That is, $K = \frac{21}{2} = 10.5$ units.

That means, 10.5 units of capital are to be employed.

6. A short run production function is given as $Q = 10L + 15L^2 - L^3$, Q is output, L is labour employed per unit of time.

(i) Derive MP_L and AP_L schedules and (ii) Find the output at which $AP_L = MP_L$.

Solution : (i) $MP_L = \frac{dQ}{dL} = 10 + 30L - 3L^2$.

$AP_L = 10 + 15L - L^2$.

(ii) $AP_L = MP_L \Rightarrow 10 + 15L - L^2 = 10 + 30L - 3L^2$.

$\Rightarrow 3L^2 - L^2 - 30L + 15L = 0$.

$\Rightarrow 2L^2 - 15L = 0$.

$\Rightarrow 2L - 15 = 0$

$\Rightarrow L = \frac{15}{2}$.

Thus $Q = 10 \times \frac{15}{2} + 15\left(\frac{15}{2}\right)^2 - \left(\frac{15}{2}\right)^3$.

$= 75 + \frac{3375}{4} - \frac{3375}{8}$

$= 497$ (approx)

Thus the requried amount of output is 497 units (approx).

7. Let the production function of a firm be given by $Q = 8LK - L^2 - K^2$.

Find out the marginal productivity of L and K. Also show that $L\frac{\partial Q}{\partial L} + K\frac{\partial Q}{\partial K} = 2Q$.

Solution : From $Q = 8LK - L^2 - K^2$, we get

$MP_L = \frac{\partial Q}{\partial L} = 8K - 2L$ and $MP_K = \frac{\partial Q}{\partial K} = 8L - 2K$.

Now, $L\frac{\partial Q}{\partial L} + K\frac{\partial Q}{\partial K} = L(8K - 2L) + K(8L - 2K)$.

$= 8LK - 2L^2 + 8KL - 2K^2 = 16LK - 2L^2 - 2K^2 = 2(8KL - L^2 - K^2)$

$= 2Q$.

Hence proved.

3.13 Some Numerical Examples on Elasticity of Substitution

1. Find the elasticity of substitution for the following production functions :

(a) $q = \alpha x_1 + (1 - \alpha)x_2$ and (b) $q = Ax_1^{\alpha}x_2^{1-\alpha}$.

Solution : For (a) $q = \alpha x_1 + (1 - \alpha)x_2$. We have $f_1 = \frac{\partial q}{\partial x_1} = \alpha$.

and $f_2 = \frac{\partial q}{\partial x_2} = 1 - \alpha$.

Naturally, $\frac{f_1}{f_2} = \frac{\alpha}{1-\alpha}$ so that $\log\left(\frac{f_1}{f_2}\right) = \log\left(\frac{\alpha}{1-\alpha}\right) = K$.

That means, $d \log\left(\frac{f_1}{f_2}\right) = dk = 0$.

Thus, $\sigma = \frac{d\left[\log\left(\frac{x_2}{x_1}\right)\right]}{d\log\left(\frac{f_1}{f_2}\right)} = \frac{d\log\left(\frac{x_2}{x_1}\right)}{dk} = \frac{d\log\left(\frac{x_2}{x_1}\right)}{0} = \infty$.

For (b) $q = Ax_1^{\alpha}x_2^{1-\alpha}$, we have $f_1 = \frac{\partial q}{\partial x_1} = \alpha A x_1^{\alpha-1}x_2^{1-\alpha}$.

Also, $f_2 = \frac{\partial q}{\partial x_2} = A(1-\alpha)x_1^{\alpha}x_2^{-\alpha}$.

That means, $\frac{f_1}{f_2} = \frac{\alpha A x_1^{\alpha-1}x_2^{1-\alpha}}{A(1-\alpha)x_1^{\alpha}x_2^{-\alpha}} = \frac{\alpha}{1-\alpha}\cdot\frac{x_2}{x_1}$.

Hence, $\log\left(\frac{f_1}{f_2}\right) = \log\left(\frac{\alpha}{1-\alpha}\cdot\frac{x_2}{x_1}\right) = \log\frac{\alpha}{1-\alpha} + \log\frac{x_2}{x_1}$.

That is, $\log\left(\frac{x_2}{x_1}\right) = \log\left(\frac{f_1}{f_2}\right) - \log\frac{\alpha}{1-\alpha} = \log\left(\frac{f_1}{f_2}\right) - \log K$.

That is, $d\log\left(\frac{x_2}{x_1}\right) = d\log\left(\frac{f_1}{f_2}\right) - d\log K = d\log\frac{f_1}{f_2} - 0 = d\log\frac{f_1}{f_2}$.

Hence, $\sigma = \frac{d\left[\log\left(\frac{x_2}{x_1}\right)\right]}{d\log\left(\frac{f_1}{f_2}\right)} = \frac{d\log\left(\frac{f_1}{f_2}\right)}{d\log\left(\frac{f_1}{f_2}\right)} = 1.$

3.13.1 Some Problems and their Solutions

1. Examine the nature of returns to scale in case of the following production functions :

(a) $Q = 2K + 3L + KL$, (b) $Q = 20K^{0.6}L^{0.5}$ (c) $Q = 100 + 2L + 3K$, (d) $Q = 5K^aL^b$, $a + b = 1$ and (e) $Q = \frac{K}{L}$.

Solution : (a) For $Q = 2K + 3L + KL$, let us double the inputs K and L to get $2(2K) + 3(2L) + 2K.2L = 4K + 6L + 4KL$.

$= 4K + 6L + 2KL + 2KL = 2(2K + 3L + KL) + 2KL = 2Q + 2KL > 2Q$.

This indicates that there are increasing returns to scale.

(b) $Q = 20K^{0.6} L^{0.5}$. Doubling inputs K and L we get

$20(2K)^{0.6}(2L)^{0.5} = 20.2^{0.6}K^{0.6}2^{0.5}L^{0.5} = 20.2^{1.1}K^{0.6}L^{0.5}$

$= 20 \times 2.14 \times K^{0.6} L^{0.5} = 42.8\ K^{0.6}L^{0.5} > 2Q$.

This implies that returns to scale are increasing.

(c) $Q = 100 + 2L + 3K$. Doubling inputs we get,

$100 + 2(2L) + 3(2k) = 100 + 6k + 4L = 200 + 6k + 4L - 100 = 2Q - 100 < 2Q$.

This illustrates that returns to scale are decreasing.

(d) $Q = 5K^a L^b$, $a + b = 1$. Doubling the inputs we get,

$5(2K)^a (2L)^b = 5.(2K)^a (2L)^{1-a}$ $\because a + b = 1.$

$= 5.2^a.K^a.2^{1-a}.L^{1-a} = 5.2\, K^a L^{1-a} = 2 \times 5\, K^a L^b = 2Q.$

Therefore, returns to scale are constant.

(e) $Q = \frac{K}{L}$. Let us double the inputs to get $\frac{2K}{2L} = \frac{K}{L} = Q$. Thus as a result of doubling inputs, there is no change in output implying that returns to scale are decreasing.

3.14 Sample Questions

A. Objective-type Questions :

Choose the correct alternative

(a) One empirical production function is (i) $P = f(L, K)$, (ii) $P = AL^{\alpha}K^{\beta}$, (iii) $P = \frac{K}{L}$, (iv) $P = 2K + 3L + KL$.

Ans. (ii)

(b) The nature of returns to scale in case of the production function $P = 20K^{0.6}L^{0.5}$ is : (i) Increasing, (ii) Decreasing, (iii) Constant, (iv) Cannot be ascertained.

Ans. (i)

(c) Suppose the production function of a firm is given by $Q = 8LK - K^2 - L^2$. The marginal product of L is : (i) 8K – 2K, (ii) 8K – 2L, (iii) 8K – 2K – 2L, (iv) 8 – 2K – 2L.

Ans. (ii)

(d) One important property of an isoquant is : (i) It is concave to the origin, (ii) It is a vertical straight line, (iii) It is a horizontal straight line, (iv) It is convex to the origin.

Ans. (iv)

(e) One particular isoquant indicates : (i) A fixed amount of output, (ii) The amount of output cannot be ascertained, (iii) Zero level of output, (iv) Varying level of output.

Ans. (i)

(f) Which of the following statements is true? (i) All costs are fixed in the short run, (ii) All costs are variable in the long run, (iii) All costs are variable in the short run, (iv) All costs are fixed in the long run.

Ans. (ii)

B. Short-type Questions :

1. Give an empirical production function. Whose names are associated with this production function?
2. What is an isoquant? What are its properties?
3. What is an L-shaped isoquant? Why is it so called? In which other name is this isoquant associated?
4. Explain, in brief, the concept of linear homogeneous production function.
5. Give your idea about the concept of elasticity of substitution.
6. Suppose a firm's production function is $Q = 4L^{\frac{1}{2}} K^{\frac{1}{2}}$.

   Find out the degree of homogeneity of this production function.
7. State the law of increasing returns to scale.
8. Distinguish between the law of variable proportions and the law of returns to scale.

C. Essay-type Questions :

1. Explain, in brief, the law of variable proportions. Analyse why a profit-maximising firm will like to produce in stage II.
2. What is a Cobb-Douglas production function ? State and prove the important properties of this production function.
3. Give the form of the Constant Elasticity of Substitution production function. State and prove the basic properties of this production function.
4. What is a linearly homogeneous production function? State and prove the important properties of this production function.
5. State and illustrate with the help of a diagram the law of variable proportions.
6. Suppose the production function of a firm as $Q = \sqrt{x_1 x_2}$. Prices per unit of x_1

and x_2 are Rs. 4 and Rs. 8 respectively. Determine the maximum level of output if the firm spends Rs. 400/-.

7. Explain, clearly, the concept of elasticity of substitution. Discuss, in detail, the uses of this concept.

8. Explain, in brief, the following terms in connection with the theory of production : (a) Isoquant, (b) Elasticity of substitution, (c) Cobb-Douglas production function, (d) Linearly homogeneous production function, (e) CES production function.

9. "To reach the least cost combination of factors a firm must equalize the marginal productivity per rupee spent on every factor." Explain and Establish this statement.

Unit 4 ❑ The Theory of Costs

Structure

4.1 Objectives

The primary objective of this chapter is to analyse different aspects of the cost of production. This will facilitate the management of firms in taking appropriate decisions with regard to the level of production to be produced and at the same time the pricing of the product. In relation to this basic objective this unit has been arranged to make the readers aware of the following issues :

(a) Accounting Costs Vs. Economic Costs.

(b) Short Run Cost Analysis.

(c) Shapes of the Short Run Cost Curves.

(d) Relationship Between AC and AVC and also between AC and MC.

(e) Long Run Costs

(f) Some Numerical Problems And Their Solutions.

(g) Sample Questions of Various Types.

4.2 Introduction

It is a fact that the study of business behaviour concentrates on the production process, that is, the conversion of inputs into outputs and, obviously, the relationship between output and costs of production. In the previous chapter, that is, in chapter three, we have studied the firm's production analysis and we have dealt with how inputs are combined to produce output. In the present chapter we shall explain how cost is defined, determined and measured. In reality, however, economists use the term, 'cost' in different ways. We may say that economists are concerned with the opportunity costs whereas the accountants are interested in historical rather than opportunity cost.

4.3 Accounting costs and Economic costs

Cost considerations play a very important role for taking any business decision. However, very often cost means different things to different people. The basic fact is that economists and accountants measure cost of production according to their individual purpose and interests. This means that there is a big difference between the very concept of cost as understood by economists and as done by the accountants.

An accountant is primarily concerned with the financial statements. That is why he considers only those costs which involve direct actual cash payments of a firm. That is, accounting costs constitute prices of raw materials, wages and salaries for workers, prices of fuel and power, insurance tax etc. The accountants take into account actual expenses and costs on a historical basis. Various empirical studies of cost function go on utilising accounting data that record what has happened. Besides, it presents information which will protect various interests of people like tax collectors, creditors and stock holders.

However, an economist's thought of cost is, to some extent, different from that of an accountant's view. Accounting data that record acutal costs do not take into account economic data. This means that the notion of cost as understood by accountants has little relevance to decision-making. For the purpose of decision-making, economists besides considering accounting costs, also consider an additional cost, that is, opportunity cost.

The opportunity cost is the cost measured in terms of the next best alternative foregone. For example, when the owner-manager of a firm spends time for someone,

as a rule, he does not have to pay wage. Obviously, the owner manager sacrifices this wage by doing unpaid work for himself. This foregone wage is the opportunity cost, though this wage is not an actual cash outlay. This implicit cost or the so-called opportunity cost, shoutd be included in economic costs. Therefore, decision-making depends not only on accounting costs but also on opportunity cost.

So far as the decision-making is concerned, accounting costs or accounting data will provide only little information. Historical data provide an imperfect guide for the future. Any decision-making should be concerned with future costs and revenues. It is because of this reason that the managerial economists cannot ignore future costs and also revenues for decision-making, as they consider forward-looking view of the firm.

Besides, accountants and economists differ over the question of depreciation. Usually, accountants record depreciation on a time-related basis and depreciation is very often not related to actual use of various assets. So far as accounting standard is concerned, accountants use tax rules to determine depreciation which may not reflect the actual wear and tear of an asset. Economists calculate not only explicit cost for the acquisition of capital assets but also cost associated with wear and tear. Accountants' notion of depreciation is based on the original cost of plant and machinery whereas economists' notion should be based on replacement value.

Economists' notion of cost is very much useful for managerial decisions as accounting costs involve several limitations.

4.4 Short Run Cost Analysis

We know that the cost function is a derived function and it has been derived from the production function. As the nature of production in the short period is different from that of the long period, nature of cost function is different under different time periods.

The short run has been defined to be that time period when the firm cannot change its fixed inputs for increasing output. It is only by varying variable inputs that a firm can change its volume of output. In the short run, therefore, total cost (TC) may be divided into two broad components : (i) Fixed cost (F.C.) and (ii) Variable Cost (V.C.). It is to be remembered that total cost encompasses all necessary expenses required to do a business.

Fixed Costs : Fixed costs are those costs which do not vary with the level of output. These costs represent the payments made for the use of fixed factors of production.

Some common examples of fixed costs are expenditure on advertisement, fixed plant and equipment, rent charges etc. More precisely, we may say that costs that arise due to the use of fixed inputs are termed fixed costs. Whether a firm produces or not, it will have to incur fixed costs. Therefore, fixed cost is independent of the level of output. Whatever the level of output (even if it is zero), fixed costs remain unchanged. The fixed cost curve has been presented in Fig. 4 (a) where the TFC curve has been drawn parallel to the horizontal axis along which the level of output has been measured.

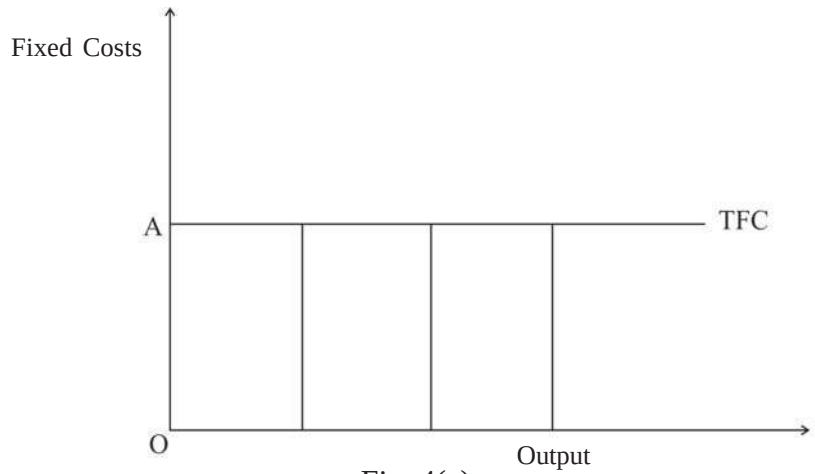


Fig. 4(a)

It is clear from the figure 4(a) that the firm incurs a fixed cost amounting to OA, whether it produces or not. That means even for zero amount of output total fixed cost is OA unit. Fixed costs are also known as supplementary costs or overhead costs. Some examples of such costs are : (i) salaries of top officials, (ii) interest on borrowed funds, (iii) expenses for municipal rent on land, (iv) depreciation of machinery etc.

Variable Cost (VC) : Variable costs are those costs which change directly with the change in output. If the firm does not produce, it will not have to use variable inputs. Therefore, when nothing is produced, variable cost will be zero. However, if the firm wants to produce, it will have to increase the use of variable inputs. Naturally, more output implies employment of more variable inputs and more variable cost. Therefore, variable cost rises in response to an increase in output. Fig. 4(b) has demonstrated the graph of the variable cost. We see from the diagram that the VC curve has started from the origin revealing that when no output (or zero ouput) has been produced, the variable cost has become zero. As soon as output goes on increasing, VC curve starts rising. Therefore.

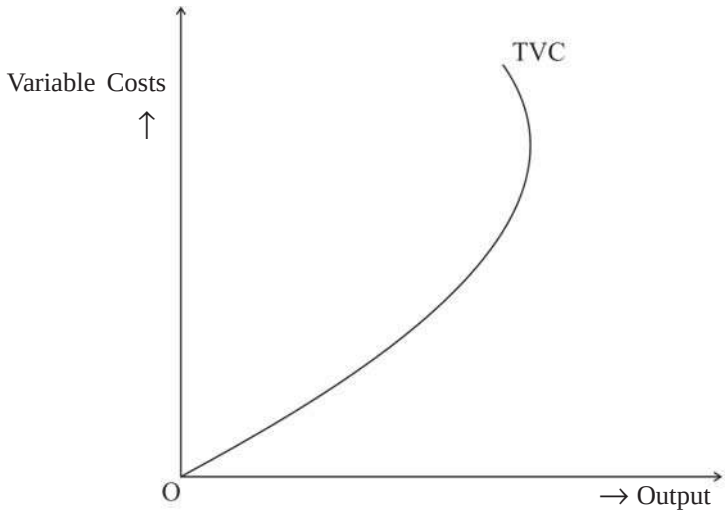


Fig. 4(b)

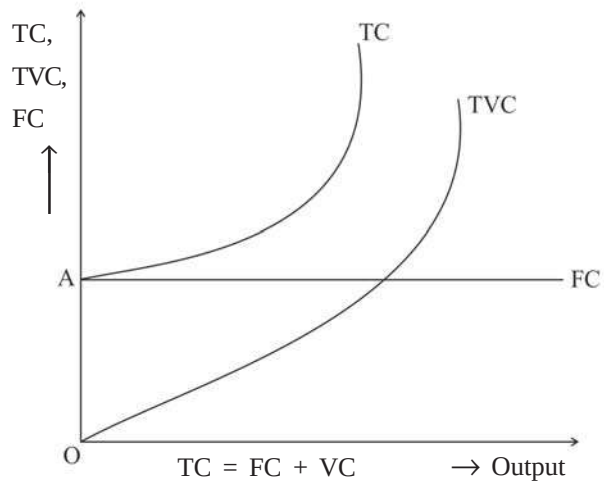


Fig. 4(c)

the VC curve has, broadly, been an inverse S-shaped because of the operation of the law of variable proportions. Initially, as more variable inputs get employed, the average variable cost starts falling. However, afterwards, use of more variable inputs leads to a rise in average variable cost. It is because of this reason that, initially, at a low level of output VC rises more than the rise in output. Variable costs include wages of daily wage earners, cost of packaging, transport cost, cost of raw materials etc.

Total Cost (TC) : The sum total of fixed cost and variable cost gives rise to total cost. The TC curve has been presented in Fig. 4(c). From this figure we notice that the TC curve, in the short run, has started from the point where from the FC curve has started. This implies that when the level of output is zero, total cost covers only fixed cost. However, with the increase in the level of output, TC rises since variable costs come into operation. That is why the TC curve takes the shape of the VC curve. At any level of output, the vertical distance between TC and FC measures the VC. In this way, by vertically adding FC and VC, we get TC.

4.5 Shapes of the Short Run Cost Curves

(i) Average Fixed Cost (AFC) : Fixed cost per unit of output produced is known as average fixed cost. So by dividing the total fixed cost by the level of output we get AFC. In symbols,

AFC = $\frac{FC}{Q}$ where Q = quantity of output. As the level of output goes on increasing and FC remains fixed, the AFC goes on declining continuously. This is so because as the volume of fixed cost is divided by the larger and larger volume of output, AFC must be declining. Besides, the AFC is a rectangular hyperbola in the sense that all rectangles formed by AFC are of equal sizes. [Here we see that AFC × Q = FC = constant and this is the equation of a rectangular hyperbola]. Thus the AFC curve is asymptotic to both the axes. This implies that the AFC curve touches neither the horizontal axis nor the vertical axis.

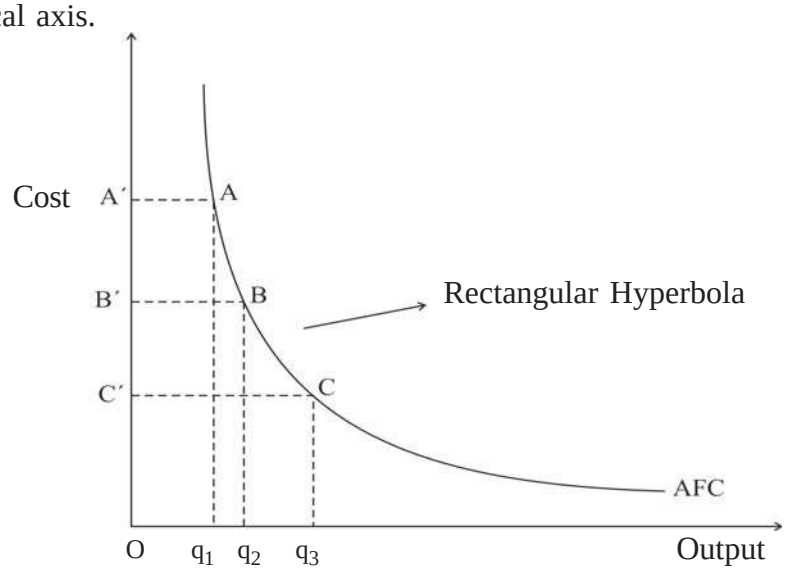


Fig. 4 (d) illustrates the shape of the AFC curve

In the diagram the [Fig. 4(d)], the AFC curve is a rectangular hyperbola where all the rectangles are of equal sizes since area of each rectangle say, $\square$ OA'Aq$_1$, $\square$ OB'Bq$_2$ and $\square$ OC'Cq$_3$ is equal as total fixed cost which is equal to the area of each rectangle is fixed.

(ii) Average variable cost (AVC) : AVC is the variable cost per unit of output produced. The AVC has been obtained by dividing TVC by the level of ouput (Q).

That is, AVC = $\frac{TVC}{Q}$, TVC is total variable cost and Q is the amount of output produced.

If we put AVC on a graph paper, we will find that, initially, AVC will come down, will reach a minimum, and, thereafter, it will start rising. This means that the AVC curve is U-shaped. Behind this U-shapedness nature of the AVC curve lies the operation of the law of variable proportions. In shot, there is a relationship between costs of production and input productivity. We assume that labour is the only variable input and, thus, labour cost is the only variable cost. Labour cost is nothing but quantity of labour employed, multiplied by the wage rate. That means, Variable cost = L.W, L being amount of labour employed and W is the wage rate.

Now, AVC = $\frac{VC}{Q} = W.\left(\frac{L}{Q}\right)$.

However, average product (AP) is $\frac{Q}{L}$.

Hence AVC = $W.\left[\frac{1}{\frac{Q}{L}}\right]$.

That is, AVC = $W\left(\frac{1}{AP_L}\right)$ where AP_L is the average productivity of labour.

Therefore, AVC is the wage rate multiplied by the reciprocal of average productivity of labour. Since, at the initial stage, AP rises, thereafter, it reaches a maximum and then declines, AVC must, therefore, initially fall, reach a minimum and thereafter it will rise. That means, the movement of the average product (AP) curve has just been the reverse

of the AVC curve. The relation between the AP_L curve and the AVC curve has been portrayed in Fig. 4(e).

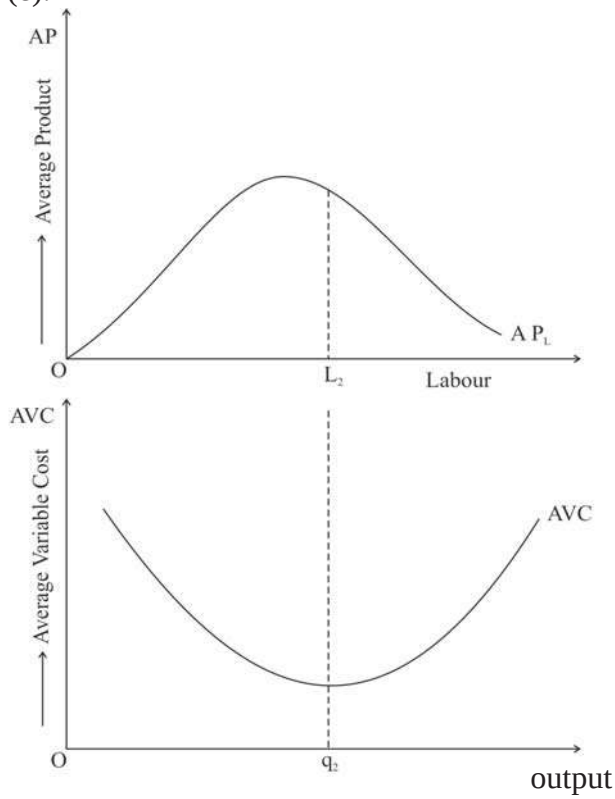


Fig 4(e) : AP and AVC Curves

The inverse relationship betwen AP and AVC curves has been demonstrateord in figure 4(e). The upper panel reveals that as labour starts increasing AP_L rises and consequently AVC declines. At OL_2 level of employment since AP_L is maximum AVC at the Oq_2 amount of output must be minimum. Beyond OL_2 amount of labour, as AP_L goes on declining, AVC rises. This means that the AVC is U-shaped.

AC or ATC : Average Cost (AC) is the total cost per unit of output produced. Thus AC is obtained by dividing total cost (TC) by output. That means,

$$AC = \frac{TC}{Q} = \frac{FC + VC}{Q} = \frac{FC}{Q} + \frac{VC}{Q} = AFC + AVC.$$

This means that the sum total of AFC and AVC gives AC. Naturally, the shape of the AC curve is governed by the shapes of AFC and AVC curves. Evidently, AVC

is U-shaped. We know that as output gets expanded both AFC and AVC decline, so AC must decline. Though AVC now rises, the significant decline in AFC causes AC to decline more. It is because of this reason that AC's minimum point comes later than AVC's minimum point. From Fig. 4(g) it is seen that upto OQ_2 amount of output, AC declines. However, beyond OQ_2 level of output, AC starts rising. This shows that increase in output beyond OQ_2 level will lead to an increase in AVC. That means, the increase in AVC more than offsets the fall in AFC. This suggests that in this region of output, AVC becomes stronger than AFC and AC continues to rise. Therefore, the AC curve is U-shaped. It is U-shaped because of the operation of the law of variable proportions.

Marginal Cost (MC) : Marginal cost is the change in total cost attributable to a change in output. That means, MC is an addition to the total cost when one more unit of output is produced. Symbolically, it is defined as

$$MC = \frac{\Delta TC}{\Delta Q}.$$

That is, $MC = TC_n - TC_{(n-1)}$, that is, MC is the additional cost for producing the n-th unit of a product when (n – 1) units of the product have already been produced. We should remember that in the short run, fixed costs do not change. In this period output changes due to change in the use of variable inputs. Therefore, by change in total cost we mean here change in total variable costs only. Naturally, fixed costs do not have any bearing on MC. That means, MC is independent of fixed cost. Therefore,

$$MC = \frac{\Delta TVC}{\Delta Q}.$$

MC refers to change in total variable cost which results from a change in output. This can be proved in the following manner :

$$MC = TC_n - TC_{(n-1)} = (FC + VC_n) - (FC + VC_{n-1}).$$
$$= FC + VC_n - FC - VC_{n-1}.$$
$$= VC_n - VC_{n-1}.$$

Thus, $MC = \frac{\Delta TVC}{\Delta Q}$.

The MC may also be defined as the reciprocal of marginal product (MP) multiplied

by the price of the variable input. To show that let us assume that labour is the variable input whose cost, i.e., wage cost, is the variable cost. Therefore, $\Delta TVC = W.\Delta L$.

Dividing both sides by ΔQ we get, $\frac{\Delta TVC}{\Delta Q} = W.\frac{\Delta L}{\Delta Q}$.

However, $\frac{\Delta Q}{\Delta L} = MP$. Therefore, $MC = W\left(\frac{1}{MP}\right) = \frac{W}{MP}$.

It is because of this reason that it is said that the shape of the MC curve is governed by the shape of the MP curve.

Since MP initially increases, reaches a maximum and thereafter declines, MC initially falls, reaches a minimum and thereafter rises. Therefore, the MC curve is U-shaped because of the operation of the law of variable proportions. This information has been portrayed in the following Fig. 4(f).

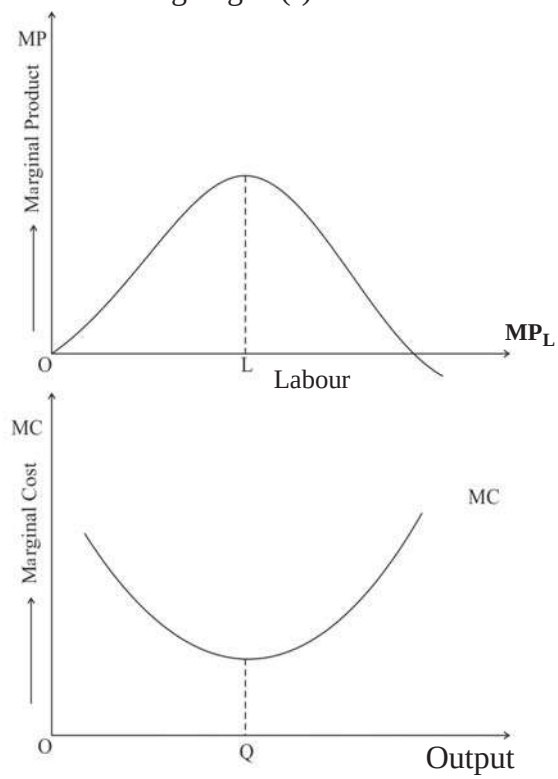


Fig. 4(f) : MP and MC curves

In the above figure, as labour employment increases, MP initially rises and consequently, MC falls with the rise in output. MP reaches maximum when OL units of labour are employed. The corresponding output level is OQ where MC has become minimum. Thereafter, additional employment of labour causes marginal product to fall even to zero or to become negative whereas MC goes on rising continuously. Therefore, the MC curve is the reciprocal of the MP curve.

All the short run cost curves can be represented in a single diagram shown below, Fig 4(g). The AFC curve has been drawn as a rectangular hyperbola. This curve has been declining steadily. The AVC curve is U-shaped, through its minimum point (N) passes the MC curve. The sum of AFC and AVC gives the AC curve which lies above both AFC and AVC curves. Also, at

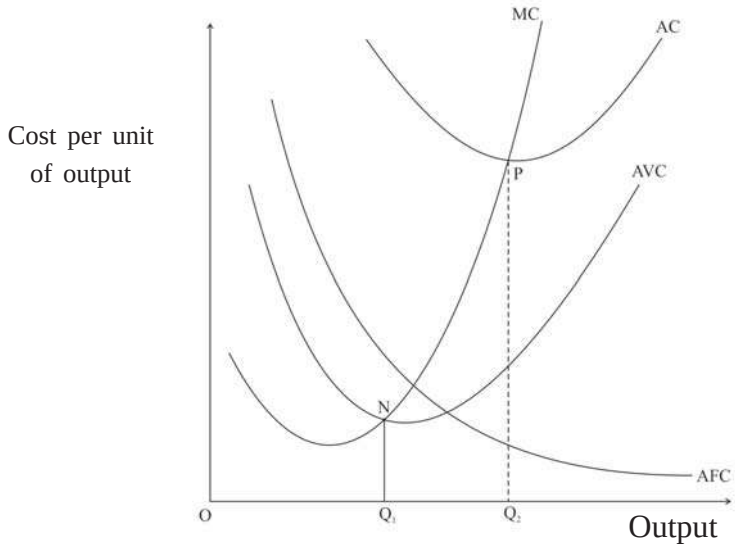


Fig 4(g) : AVC, AFC, AC and MC curves

the minimum point of AC (that is, point P), MC passes through. We are to note that AVC and AC get closer as the level of output goes on rising since AFC declines continuously as output rises. Like AC and AVC, the MC curve is also U-shaped. MC equals both AC and AVC at their minimum points. Besides, MC lies below both AVC and AC throughout the range over which these curves fall. However, it lies above them when they are rising. It is to be noted that all the cost curves except the AFC curve are U-shaped.

RELATIONSHIP BETWEEN AC AND AVC.

There are two ways for getting the AC :

(a) AC = AFC + AVC.

and (b) AC = $\frac{TC}{Q}$.

But AVC = $\frac{TVC}{Q}$.

That is, AVC is a part of AC, given AC = AFC + AVC. Moreover, both AVC and AC have U-shapes because of the operation of the law of variable proportions. However, the minimum point of AC lies to the right of the minimum point of AVC. The basic reason behind this is that AC not only includes AVC but also AFC which continuously falls as output rises. Besides, initially, with the rise in output, AVC falls. So AC must fall. AVC starts rising after OQ amount of output is being produced, its rise over a certain range is offset by a fall in AFC. That is why, AC goes on falling over the range of output even if AVC rises. That is why, AC's minimum point comes at a later range of output than AVC's minimum point. But once OQ_2 amount of output gets produced, the influence of the rise in AVC becomes stronger than the fall in AFC, so AC starts rising.

RELATION BETWEEN AVERAGE COST (AC) AND MARGINAL COST (MC)

We know that total cost (TC) = Average cost (AC) Quantity of output produced.

Thus, TC = AC × Q.

Also, marginal cost (MC), MC = $\frac{d(TC)}{dQ}$.

That is, MC = $\frac{d}{dQ}(AC.Q)$, = $\frac{d(AC)}{dQ}$. Q + AC$\frac{dQ}{dQ}$.

That is, MC = $\frac{dAC}{dQ}$. Q + AC .....(a)

Now, $\frac{dAC}{dQ}$ is the slope of AC. Since AC and Q are all non-negative,

MC > AC if and only if $\frac{d(AC)}{dQ} > 0$.

and MC < AC if and only if $\frac{d(AC)}{dQ} < 0$.

Finally, MC = AC if and only if $\frac{dAC}{dQ} = 0$.

The slope of AC will reduce to zero at its minimum point. That is why, at AC's minimum point, AC = MC.

That means, when AC is minimum, AC equals MC. If the slope of AC is positive, that is, if AC curve is rising, MC > AC. Further, if the slope of AC is negative, that is, if AC curve is falling, MC < AC.

From the above analysis we can summarize the relation between AC and MC as follows:

1. When AC falls, MC also falls and the rate of fall of MC is greater than the rate of fall of AC.

2. When AC is minimum and remains constant, MC and AC are equal or identical.

3. When AC rises, MC also rises and the rate of increase of MC is greater than that of AC.

The above-mentioned relationship between AC and MC can, graphically, be demonstrated in the following Fig. 4(h).

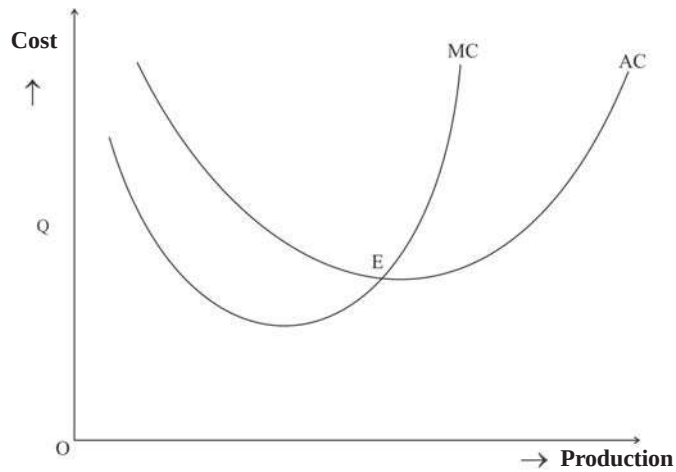


Fig. 4(h)

In the above diagram (Fig. 4h) we measure production horizontally and cost vertically. At point E, AC has become minimum and constant for a while. At this point both AC and MC have become identical. Before reaching point, E, AC is continuously decreasing and MC is also continuously decreasing. However, the rate of decrease of MC is greater than that of AC. After the point E, AC is increasing and so is MC. However, the rate of increase of MC is greater than that of AC.

4.6 Derivation of Long Run Average Cost Curve (LAC)

Let us now deal with the procedure for deriving the long run average cost curve.

We are to remember that long run is said to be that time period when a firm under consideration can change all its inputs. In reality, in the long run there are no fixed inputs–all inputs are variable. This implies that in the long run there is no fixed cost–all costs are variable. In the long run, therefore, a firm can change its scale of production according to its necessity. In the short run, the size of a plant or scale remains fixed while, in the long run changes in plant size or in scale can be made. A firm, in the long run, can move from one plant to another one. It can also build up a large-sized plant or a smaller one, if the situation so demands. The long run, is a "planning horizon" in the sense that it acts as a guideline to the firm with regard to the future decision of output. We know that production takes place in the short period. Every firm aims at producting for a future date and chooses many aspects of the short run in which it operates in the future. Therefore, long run consists of all short run situations among which the firm may choose.

Thus, LAC is derived from SAC curves. LAC depicts the lowest possible average cost for the production of various possible levels of output. In order to derive the LAC curve, we assume that there are three different sizes of plants in an industry, namely small, medium and large. Three SAC curves, namely SAC_1, SAC_2 and SAC_3 represent respectively small-sized, medium-sized and large-sized plants. These have been demonstrated in figure 4(i). These SAC curves are also called plant curves. As we are considering the long run situation, the firm can choose any plant size where it will operate in the future to produce a given level of output at the least possible cost.

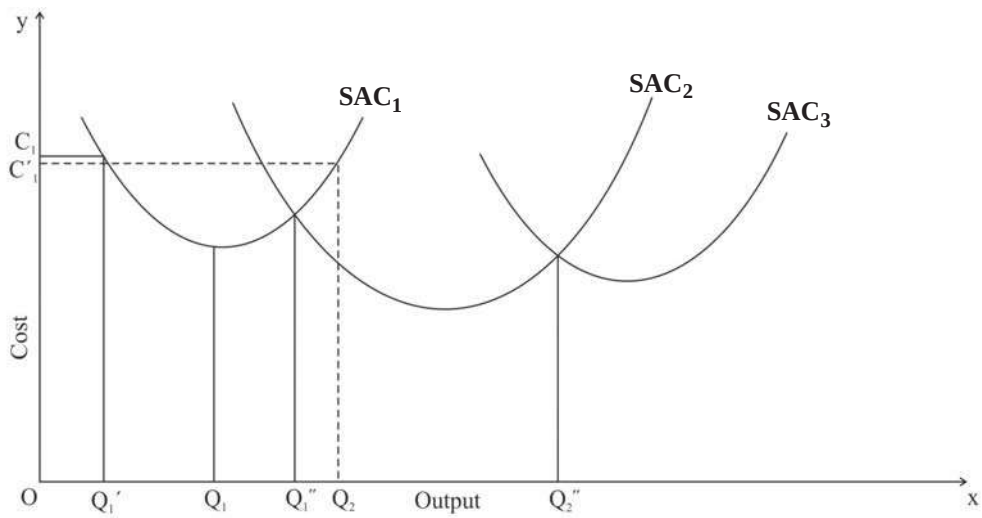


Fig. 4(i) : SAC curves for three plant sizes

If the firm decides to produce OQ_1 amount of output, it is sure to choose the plant size denoted by SAC_1. A lower level of output (say OQ_1) can also be produced on SAC_1 though at a higher cost. However, the same plant size, (i.e., SAC_1) enables a firm to produce larger output at a lower cost. If OQ_2 amount is considered to be the most profitable amount of output, the firm will choose SAC_2, that is, the medium-sized plant. However, taking such a decision is not so easy as it appears to be at the first sight. Assume the firm operates at SAC_1 and demand for its product rises gradually. However, it can produce OQ_1 amount of output at the possible lowest cost even operating on SAC_1. But production beyond OQ_1 will need a larger cost. If the firm expects to produce OQ_1 (seen in the figure) its choice of plant size becomes difficult as costs are the same for both the plant sizes, namely SAC_1 and SAC_2. However, the choice of the optimal plant size depends on the firm's expectation with regard to its demand for product in the coming years. At this level of output, cost can not be the determinant of the choice of a plant size. Now, the firm expects its demand for the product to increase in future. Therefore, the firm will instal the plant number SAC_2 rather than SAC_1. With a lower level of cost larger output can now be produced. In the same way though the output level OQ_1 can be produced by both the plant sizes, SAC_2 and SAC_3, it is economic to use the plant size represented by SAC_3.

Let us now assume that the industry faces a larger number of plant sizes represented by five SAC curves shown in figure 4(j). These curves will generate a smooth and continuous curve called the planning curve or LAC curve. Each point on this curve shows the least possible cost for producing the corresponding level of output. Thus the LAC curve is a planning curve since it is the curve which helps a firm to decide what plant is to be established for producing an output level consistent with the opitmal cost. Naturally, the firm selects that short run plant which yields the minimum cost of producing the anticipated output level. To produce a particular amount of output in the long run, the firm must select a point on the LAC curve corresponding to that output and it will then build a relevant short-run plant and operate on the corresponding SAC curve.

Let us now suppose that the firm thinks that for producing output OQ_1, point A on SAC_1 becomes the most profitable one. Naturally, it will build up a plant at the lower cost represented by SAC_1 (At point A, SAC_1 is tangent to the LAC curve). But the firm could reduce its cost by expanding output to the amount associated with point B,

the minimum point on SAC_1. We now assume that the firm anticipates that demand for its product would be rising in future. So, it would construct a new plant represented by SAC_2 and will operate at point D on SAC_2, thereby lowering its unit cost. (Corresponding to the output level OQ_2, SAC_2 is tangent to LAC). In the same way, for output OQ_3, the firm will construct SAC_3 and operate at point E as here the unit cost has become the least. For output level OQ_4, the firm will construct plant size SAC_4 and will operate at point F, though the minimum point of SAC now lies to the left of the operational point F. Each point of the LAC is, therefore, the point of tangency with the corresponding SAC curves. The LAC curve is the locus of all the tangency points. Consequently, the LAC curve has been called the envelope curve as it envelopes a family of SAC curves.

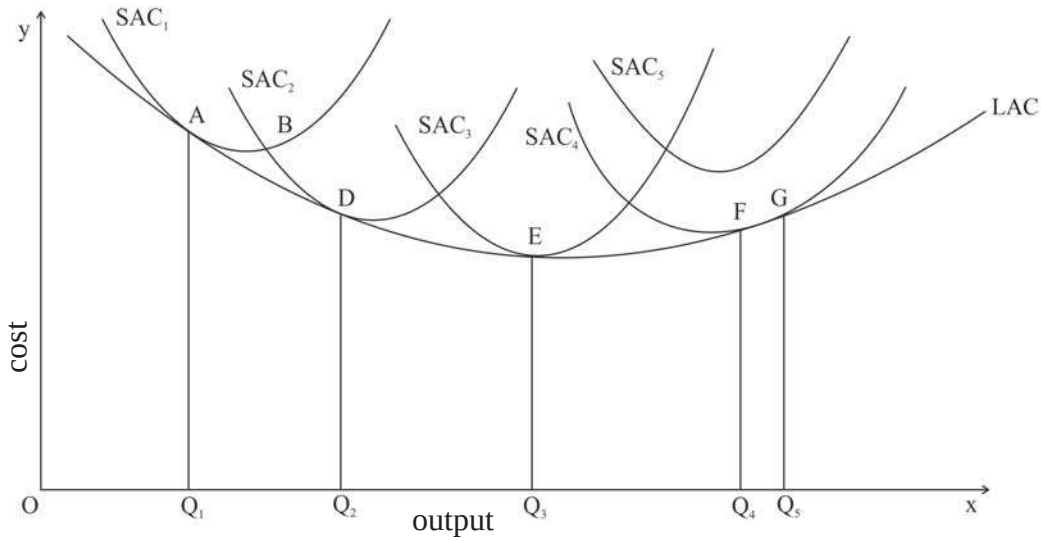


Fig 4(j) : LAC Curve

However, the LAC curve throughout its length is not tangent to the minimum points of all SAC curves. Only at point E, the minimum point of LAC, is tangent to the minimum point of SAC.

We are to note here that at the points of tangency SAC = LAC. But to the right or to the left of the tangency point SAC > LAC. Therefore, we can say that the LAC curve is U-shaped, that is, initially it falls, reaches a minimum point and rises afterwards as level of output gets expanded. However, the U-shapedness of the LAC curve is less pronounced than the U-shapedness of the SAC curve.

4.7 Some Problems and Their Solutions

1. Let a firm's cost function be $C = 20 + 3Q + 0.2Q^2$. Show that MC equals AC when the latter is minimum.

Solution : Here AC = $\frac{C}{Q} = \frac{20}{Q} + 3 + 0.2Q$

Also, MC = $\frac{dC}{dQ} = 3 + 0.4Q$

AC will be minimum when $\frac{dAC}{dQ} = 0$.

Now. $\frac{dAC}{dQ} = \frac{d}{dQ}\left(20Q^{-1}\right) + \frac{d}{dQ}(0.2Q) = \frac{-20}{Q} + 0.2 = 0$.

That is, $0.2 = \frac{20}{Q^2}$, that is, $0.2Q^2 = 20$.

That is, $Q^2 = \frac{20}{0.2} = \frac{\overline{20}}{\frac{2}{10}} = 100$.

That is, Q = 10.

When Q = 10, MC = 3 + 0.4 × 10 = 7.

Also, when Q = 10, AC = $\frac{20}{10} + 3 + 0.2 \times 10 = 2 + 3 + 2 = 7$.

Therefore, MC = AC when the latter is a minimum.

2. If the total cost function is given by $TC = 100 - 2q + 0.5q^2$,

show that the slope of the average cost (AC) curve is negative when output is less than 10.

Solution : Here AC = $\frac{TC}{q} = \frac{100 - 2q + 0.5q^2}{q} = \frac{100}{q} - 2 + 0.5q$.

Now, the slope of AC = $\frac{dAC}{dq} = \frac{d}{dq}$ $[100q^{-1} - 2 + 0.5q]$.

$= (-1)\ 100q^{-2} + 0.5 = 0.5 - \frac{100}{q^2}$.

That is, the slope of AC = $0.5 - \frac{100}{q^2}$.

When q = 10, the slope of AC is 0.5. Thus when q < 10, say 9, the slope of AC will be negative, i.e., less than –0.5.

3. Show mathematically that the marginal cost has no relation to fixed costs.

Solution : Let the total cost function be C = K + f (Q) where K = TFC = Constant and f(Q) is TVC.

Now, AC = $\frac{C}{Q} = \frac{K}{Q} + \frac{f(Q)}{Q}$ = AFC + AVC.

Also, MC = $\frac{dC}{dQ} = \frac{d}{dQ}\left(\frac{K}{Q}\right) + \frac{d}{dQ}$ $[f(Q)] = f^1(Q)$ since K is a constant.

That means, marginal cost has no relation to fixed costs.

4. Suppose $C = a + bx + cx^2$ is the total cost function of a firm, x being the quantity of output. Show that $\frac{d(AC)}{dx} = \frac{1}{x}(MC - AC)$.

Solution : Since $TC = a + bx + cx^2$,

AC = $\frac{c}{x} = \frac{a}{x} + b + cx$.

and MC = $\frac{d}{dx}\left(a + bx + cx^2\right) = b + 2cx$.

Now, $\frac{d(AC)}{dx}$ = slope of the AC function

$= \frac{d}{dx}\left[\frac{a}{x} + b + cx\right] = \frac{-a}{x^2} + c$ .....(i)

Thus, $\frac{1}{x}(MC - AC) = \frac{1}{x}\left[(b+2cx)-\left(\frac{a}{x}+b+cx\right)\right] = \frac{1}{x}\left[cx-\frac{a}{xc}\right]$

$= \frac{1}{x}c - \frac{a}{x^2} = c-\frac{a}{x^2}$ .....(ii)

From (i) and (ii) the result follows.

5. Suppose the total cost function is given by $C = 15q - 6q^2 + q^3$. Obtain the equations of AC and MC curves. Find the ouput level at which AC is minimum and show, further, that when AC is minimum, AC equals MC.

Solution : $AC = \frac{c}{q} = 15 - 6q + q^2$ and $MC = \frac{dc}{dq} = 15 - 12q + 3q^2$.

When AC is minimum, $\frac{d\,AC}{dq}=0$.

Now, $\frac{d\,AC}{dq}=0$. Thus $\frac{d\,AC}{dq}=0 \Rightarrow -6 + 2q = 0$. That is, $q = 3$.

That is, $-6 + 2q = 0$. That is, $q = 3$.

Further, $\frac{d^2AC}{dq^2}=2 > 0$. Hence AC is minimum when $q = 3$.

When $q = 3$, $AC = 15 - 6.3 + 3^2 = 15 - 18 + 9 = 6$.

When $q = 3$, $MC = 15 - 12 \times 3 + 3(3^2) = 15 - 36 + 27 = 6$.

Thus, when AC is minimum, AC = MC.

6. Let the total cost be given by $C = 5000 + 1000q - 500q^2 + \frac{2}{3}q^3$.

Answer the following questions :

(i) Find the MC function, (ii) Find the expression for the slope of the MC function, (iii) Find the average total cost function, (iv) At what value of q does MC equal AVC?

Solution : (i) since the cost function is $C = 5000 + 1000q - 500q^2 + \frac{2}{3}q^3$,

MC will be $\frac{dc}{dq} = \frac{d}{dq}$ [5000 + 1000q – 500q^2 + $\frac{2}{3}$q^3].

= 1000 – 2 × 500q + $\frac{2}{3}$3q^2 = 1000 – 1000q + 2q^2 Ans.

(ii) The slope of the MC function is $\frac{d}{dq}\left(\frac{dc}{dq}\right)$ = –1000 + 4q Ans.

(iii) Average total cost function = $\frac{c}{q} = \frac{1}{q}$ [5000 + 1000q – 500q^2 + $\frac{2}{3}$q^3].

= $\frac{5000}{q}$ + 1000 – 500q + $\frac{2}{3}$q^2 Ans.

(iv) TVC = TC – TFC = 5000 + 1000q – 500q^2 + $\frac{2}{3}$q^3 – 5000.

= 1000q – 500q^2 + $\frac{2}{3}$q^3.

Hence, AVC = $\frac{TVC}{q}$ = 1000 – 500q + $\frac{2}{3}$q^2.

Now, MC = AVC means 1000 – 1000q + 2q^2 = 1000 – 500q + $\frac{2}{3}$q^2.

That is, –1000q + 500q + 2q^2 – $\frac{2}{3}$q^2 = 0.

That is, –500q + $\frac{4}{3}$q^2 = 0. That is, $\left[\frac{4}{3q} - 500\right] = 0 \because q \neq 0$.

That is, $\frac{4}{3}$q = 500, that is, q = $\frac{500 \times 3}{4}$ = 375 units. Ans.

7. Given the total cost function $C = 1000 + 10q - 0.9q^2 + 0.04q^3$, determine the rate of output that leads to minimum AVC.

Solution : Let us, first of all, find the MC.

$$MC = \frac{dc}{dq} = \frac{d}{dq}\ [1000 + 10q - 0.9q^2 + 0.04q^3] = 10 - 1.8q + 0.12q^2.$$

Now, $TVC = TC - TFC = 10q - 0.9q^2 + 0.04q^3$.

$$AVC = \frac{TVC}{q} = 10 - 0.9q + 0.04q^2.$$

Since the minimum point of AVC occurs at its intersection with MC, let us equate AVC and MC.

$10 - 0.9q + 0.04q^2 = 10 - 1.8q + 0.12q^2$.

That is, $1.8Q - 0.9Q + 0.04q^2 - 0.12q^2 = 0$.

That is, $0.9q - 0.08q^2 = 0$.

That is, $q\ (0.9 - 0.08q) = 0$.

That is, $q = 0$ and $q = \frac{0.9}{0.08} = \frac{90}{8} = 11.2 = 11$ (say) units.

Therefore, the required amount of output is 11 units.

8. Express mathematically the relation between AC and MC.

Solution : We know that when AC is falling, MC < AC.

Now, $AC = \frac{C}{Q}$. When AC is falling, $\frac{d\left(\frac{C}{Q}\right)}{dQ} = \frac{Q\frac{dc}{dQ} - c}{Q^2} < 0.$

That is, $\frac{\frac{dc}{dQ}}{dQ} - \frac{c}{Q^2} < 0,$

That is, $\frac{dc}{dQ} - \frac{c}{Q} < 0.$

That is, $\frac{dc}{dQ} < \frac{c}{Q}$. That is, MC < AC.

Again, when AC is constant, $\frac{d\left(\frac{c}{Q}\right)}{dQ} = 0$.

That is, MC = AC.

Also, when AC is increasing, $\frac{d\left(\frac{c}{Q}\right)}{dQ} < 0$ and MC < AC.

Therefore, the relation between AC and MC can be stated in plain language as:

(a) When AC is constant, AC = MC.

(b) When AC is increasing, AC < MC.

(c) When AC is decreasing, AC > MC.

9. The total cost, C, of a firm is given by $C = 1000 + 100q - 80q^2 + \frac{1}{3}q^3$ where q is the quantity produced.

(i) Find the marginal cost of production.

(ii) At what value of q does marginal cost equal average variable cost?

Soltion : (i) $MC = \frac{dc}{dq} = \frac{d}{dq}\left[1000 + 100q - 80q^2 + \frac{1}{3}q^3\right] = 100 - 160q + q^2$.

(ii) In the given problem $TVC = 100q - 80q^2 + \frac{1}{3}q^3$.

Thus $AVC = \frac{TVC}{q} = 100 - 80q + \frac{1}{3}q^2$.

Now, $MC = AVC \Rightarrow 100 - 160q + q^2 = 100 - 80q + \frac{1}{3}q^2$.

$\Rightarrow q^2 - \frac{1}{3}q^2 - 160q + 80q = 0.$

$\Rightarrow \frac{2}{3}q^2 - 80q = 0 \Rightarrow \frac{2}{3}q - 80 = 0.$

$\Rightarrow 2q = 240 \Rightarrow q = 120.$ Thus when $q = 120$, MC = AVC.

10. Express, mathematically, the relation between AC and MC. Show, further, that MC has no relation to fixed costs.

Solution : The total cost function of a firm can be expressed as C = K + f(Q) where K is TFC = Constant and F(Q) is TVC, a function of output.

The average cost function is $\frac{c}{Q} = \frac{K}{Q} + \frac{f(Q)}{Q}$ = AVC + AVC.

The marginal cost function is $\frac{dc}{dQ} = \frac{d}{dQ}(K) + \frac{d}{dQ}\left[f(Q)\right] = f^1(Q)$ as K is constant.

Therefore, the marginal cost has no relation to fixed costs.

11. Suppose a firm's cost function is $C = 20 + 3Q + 0.2Q^2$.

Prove that MC equals AC when the latter is a minimum.

Solution : Here MC = $\frac{dc}{dQ} = 3 + 0.4Q$ and AC = $\frac{20}{Q} + 3 + 0.2Q$.

For minimization of cost let us differentiate AC w.r.t. Q and set the first derivative equal to zero.

Now, $\frac{dAC}{dQ} = \frac{d}{dQ}\left[\frac{20}{Q} + 3 + 0.2Q\right] = -\frac{20}{Q^2} + 0.2 = 0.$

Hence, $0.2 = \frac{20}{Q^2} \Rightarrow 0.1 = \frac{10}{Q^2} \Rightarrow Q = 10.$

When Q = 10, MC = 3 + 0.4(10) = 7 and AC = $\frac{20}{10} + 3 + .2 \times 10 = 2 + 3 + 2 = 7.$

Therefore, MC = AC when the latter is a minimum.

12. Suppose that the total cost function is given by TC = 100 – 2q + $0.5q^2$. Show that the slope of average cost (AC) is negative when output is less than 10.

Solution : The AC is defined as $\frac{TC}{q} = \frac{100 - 2q + 0.5q^2}{q}$

Now, $\frac{d(AC)}{dq} = \frac{(-2+q)q - (100 - 2q + 0.5q^2)}{q^2}$

$= \frac{-2q + q^2 - 100 + 2q - 0.5q^2}{q^2}$

$= \frac{0.5q^2 - 100}{q^2} = 0.5 - \frac{100}{q^2}$.

That is, the slope AC is 0.5. When q < 10, say, 9 the slope of AC will be negative, i.e. less than –0.5. Hence proved.

4.8 Sample Questions

A. Objective-type Questions

(a) What kind of a curve is average fixed cost (AFC)? : (i) U-shaped, (ii) rectangular hyperbola, (iii) downward falling and (iv) upward rising.

Ans. (ii)

(b) The marginal cost curve cuts the average cost curve at its : (i) starting point, (ii) end point, (iii) minimum point, (iv) maximum point.

Ans. (iii)

(c) The shape of the short run average cost curve is : (i) rectangular hyperbola, (ii) inverted U-shaped, (iii) U-shaped and (iv) square shaped.

Ans. (iii)

(d) The marginal cost can be derived from : (i) total cost, (ii) variable cost, (iii) fixed cost, (iv) average cost.

Ans. (ii)

(e) In the long run all costs are —— Fill in the gap. : (i) fixed, (ii) variable, (iii) partly fixed and partly variable, (iv) uncountable.

Ans. (ii)

(f) Which cost increases continuously with the increase in production?
(i) average cost, (ii) marginal cost, (iii) fixed cost and (iv) variable cost.

Ans. (ii)

(g) Which of the following cost curves is never 'U'-shaped?

(i) average cost curve,

(ii) marginal cost curve,

(iii) average variable cost curve,

(iv) average fixed cost curve.

Ans. (iv)

(h) Which of the following statements is correct?

(i) when the average cost is rising the marginal cost must also be rising,

(ii) when the average cost is rising, the marginal cost must be falling,

(iii) when the average cost is rising, the marginal cost is above the average cost,

(iv) when the average cost is falling, the marginal cost must be rising.

Ans. (iii).

B. Short-type Questions :

1. Explain the relationship between average cost and marginal cost.
2. Prove mathematically or graphically that the marginal cost cuts the average cost at the minimum point of the latter.
3. What is the shape of the average fixed cost curve of a firm? Explain the nature of this curve.
4. Explain the concepts of fixed cost and variable cost. Do your think that these concepts are always valid ?
5. In the short run what will be total cost if total production is zero ? Give reasons for your answer.

6. Let the equation of the total variable cost curve be $TVC = aQ - bQ^2 + cQ^3$. Prove that marginal cost equals AVC at the minimum point on the average variable cost curve.

C. Essay-type Questions :

1. Define AVC, AFC, AC and MC. Draw a graph showing all these curves. Also explain their shapes, giving suitable reasons.
2. Why is the short run average cost curve of a firm U-shaped ? Explain your answer graphically, giving sufficient reasons.
3. Explain the procedure for deriving the LAC curve from the SAC curves.
4. Suppose the total cost function of a firm is given by $C = 2Q - 2Q^2 + Q^3$.

(i) Find out the average cost and marginal cost.

(ii) At what level of output will AC be minimum?

(iii) Show, further that at the minimum point of AC, AC and MC are identical.

5. In the context of cost analysis of a firm complete the following table :

Output (Unit)	Total fixed Cost	Total Variable Cost	Total Cost	Marginal Cost
1			45	
2				30
3		90		
4			120	
5				10
6	10			

Unit 5(a) ❐ Market and Market Structures

Structure

5a.1 Objectives

The basic objectives of this unit are as follows :

i. Meaning and features of a market, classification of markets.

ii. Various aspets of Revenve in different markets.

iii. Perfectly competitive market and its various aspects.

iv. Supply curve of a firm under perfectly competitive market.

5a.2 Introduction

Market occupies a very important position in the modern complex monetary organization. Adam Smith, the Father of Economics, in his great book "Wealth of Nations" published in 1776 in two volumes had, very rightly, pointed out that division of labour is limited by the extent of the market. Market helps the producers to take the economic decisions which are of basic nature : what to produce and how to produce. Naturally, market highly affects the producers, consumers and various other sections of a society. Professor Wicksteed had, rightly, recommended : "The market is the characteristic phenomenon of economic life. The constitution of markets and market prices are the central problems of Economics."

5a.3 Meaning of a Market

In ordinary language, "market" refers to a place where commodities are bought and sold. In this sense we talk of Hong Kong Market, Mumbai Market, Howrah Market and the like. However, in Economics, "market" has no reference to a place but to a particular commodity which is being bought and sold. In reality, we very often speak of tea market, bullion market, labour market, share market and so on.

In Economics "market" refers to a group of buyers and sellers taking part in the exchange of a commodity. The buyers and sellers may be scattered within a country or may be abroad. However, there must be some contact by means of meetings, fairs, published price lists or through telephone, e-mail, whatsup, fax or some other forms of communication.

5a.4 Features of a Market

Thus, in Economics, there are four basic features of a market :

(i) Buyers and sellers : In a market there must have the existence of buyers and sellers. However, there is no definite limit to these numbers.

(ii) There must be some contact : There must have some contact among the buyers and sellets. This contact may be either direct or indirect.

(iii) There must be the existence of a commodity : In a market there should be a commodity to deal with. In case there are different commodities there will be different markets.

(iv) The existence of a price : In the market there should be price for the commodity that is bought and sold.

5a.5 Classification of Markets

Markets may be classified into various ways :

1. On the basis of Area : According to the area covered, markets may be classified into local markets, regional markets, national markets and international markets.

(a) **Local Market :** Local markets refer to a market in a particular village or locality. Generally, perishable goods like milk, butter, vegetables have local markets.

(b) **Regional Market :** Regional market refers to a market which covers a particular region. Generally, bulky articles like bricks and stones have regional markets.

(c) **National Market :** If the demand for and supply of a commodity is spread over the entire country, the market is said to be a national market. Generally, jute, cotton, wheat, rice and tobacco etc. have national markets.

(d) **International Market :** When the demand for and supply of a commodity is at the global level, the market is said to be an international market. Normally, valuable metals like gold, diamond, silver etc. have an international market.

2. On the basis of Time : On the basis of time element, a market has been classified into very short period, short period and long period markets.

(a) **Very short period markets :** A very short period is that in which supply can neither be increased nor can it be decreased to adjust to demand. This period is also known as market period. Vegetables market, fruit market and markets for perishble goods are examples of such markets.

(b) **Short Period :** Short period refers to a period of time in which the rate of production is variable. An increase in production is possible by an intensive use of capital or by enlarging the size of the plant.

(c) **Long Period :** By long period we mean a period of time in which the supply of the commodity can be varied according to the conditions of demand. Long period involves many years.

3. On the basis of a Commodity : On the basis of a commodity concerned markets may be classified as

(a) **Share Market :** A share market is a place in which shares are bought and sold. Thus Calcutta share market, Mumbai share market are examples of a share market.

(b) **Bullion Market :** A bullion market is a place in which valuable metals like diamonds, gold, silver etc. are traded.

(c) **Money Market :** A money market is a place in which short term credit instruments are bought and sold. London money market, Mumbai money market, New York money market, etc. are some of the money markets in the world.

(d) **Capital Market :** A capital market is meant for very long period loans of capitals.

4. On the basis of the volume of Business : On the basis of the volume of business also sometimes markets have been classified into wholesale markets and retail markets. In the former market goods are transacted in large quantities whereas in the latter market goods are transacted in small quantities.

5. Classification on the basis of transactions : Markets may be classified, on the basis of transactions, into the spot market and the future market. A spot market is that market where goods are physically exchanged on the spot. On the other hand, future market is one where an agreement is made on the spot while exchange of goods takes place in future, as agreed upon by the parties.

6. Classification on the basis of the "status of sellers" : On the basis of the status of sellers, markets have, broadly, been classified into primary markets, secondary markets and terminal markets.

A primary market is a market where goods are transacted from the producers to the wholesalers. A secondary market is a market where the wholesalers act as an intermediate link between the producers and the retailers. A terminal market is one where the goods are trasnferred from the retailers to the ultimate consumers.

7. Classification on the basis of Government Regulations : On the basis of the Govt. regulations, markets may also be classified as regulated markets and unregulated markets. A regulated market is one where transactions are governed by various rules and regulations as framed by the Government. On the other hand, a market where transactions of goods and services are left to market forces of demand and supply is known as an unregulated market.

8. On the basis of Competition : According to the degree of competition, markets may be classified into perfect markets and imperfect markets.

Perfect markets : A perfect market is one where there is perfect competition. Sometimes a distinction is made between pure competition and perfect competition. Perfect competition is a wider concept which includes not only the conditions of pure competition but also a few more imperfect market.

A market is imperfectly competitive if the actions of one or more buyers or sellers have a perceptible influence on price.

An imperfect market may assume the following forms :

(a) Monopoly market.

(b) Monopolistically competitive market.

(c) Oligopoly market.

(d) Duopoly market.

(e) Bilateral monopoly.

(f) Monopsony.

(g) Duopsony.

and (h) Oligopsony.

5a.6 Factors Affecting the Size of a Market

The size or extent of a market depends upon a number of factors :

(i) Nature of the commodity and (ii) General Conditions.

NATURE OF THE COMMODITY

(a) **Wide demand :** A commodity which is of universal demand will have a wide market. The best example of a commodity having wide demand is fuel oil which has wide demand and has unlimited market. On the other hand, a commodity which is local in character will have a limited market. Thus Malayalm novels and Tamil films have only a regional market.

(b) **Durability :** A commodity being durable in nature will have wide market. Thus goods like gold, silver etc. have very wide markets since these commodities can be preserved for a long time and can also be transported without any damage. On the other hand, perishable commodities like vegetables, milk and flowers have only limited markets. However, the development of processing facilities, refrigeration and air lifting have greatly increased the size of the market for perishable goods also.

(c) **Portability :** An important pre-requisite for wide markets is that the commodity should be easily portable. It must be small in size but possess a high value. Goods like machines and cotton textiles are easilly portable and thus they command wide markets. On the other hand, bricks and stones enjoy only a narrow market.

(d) **Grading and Sampling :** Grading means dividing the goods on the basis of

quality. Sampling means a part reprsenting the whole. Goods which are amenable to grading and sampling normally command an international market. Grading helps the purchasers to identify the quality of the product without making the inspection.

(e) **Adequate Supply :** To have a wider market the commodities should have an adequate supply. However, rare pictures and paintings though they could not be supplied in plenty, enjoy international markets because of their prestige value.

5a.7 General Conditions Affecting the Demand for a Commodity

The demand for a commodity is highly affected by the general conditions prevailing in the country :

1. Peace and security : The existence of politial stability, peace and security are of vital importance for a smooth trade. If the law and order situation in a country due to political unrest and revolutionary attitudes of some groups become very bad, markets will get hampered.

2. Stable Currency and Efficient Financial Institutions : With the help of well-developed banks and other financial institutions, marketing can be profitable and easily carried on over wide and extensive areas. In the modern world, international trade depends on the value of the home currency and the confidence it inspires among foreigners.

3. Government Policy : The import and export policy of the Government of a country has also a considerable impact on the size of the market. If the Government imposes a lot of restrictions in the form of quotas and duties, trade will be affected. Likewise restrictive policies of the importing country will also have an adverse effect on the expoting country.

4. Transport and Communication : A properly developed transport system in the form of roads, railways, steamships and aeroplanes is highly useful for the widening of both internal and international trade. Likewise, efficient means of communications, like post office and telegraphs, telephones, email etc. help the traders to keep in touch with their distant markets and send information regarding market trends.

5. Degree of Division of Labour : Adam Smith, the Father of Economics, said “Division of Labour is limited by the extent of a market”. However, the extent of the market also depends upon division of labour. Division of labour by reducing the average cost of production increases the size of the market.

REVENUE ANALYSIS

A very important component of the equilibrium position of a firm under any form of a market is the revenue, the other component being cost which has already been dealt with, in detail, in a previous chapter (i.e., chapter 4) Hence in this section we will concentrate our attention on revenue analysis only.

5a.8 The Concept of Revenue

By revenue we mean receipts derived from sales. Revenue depends upon the price at which, quantities of output are sold by the firm. A firm's revenue can be categorised as : (a) total revenue (TR), (b) average revenue (AR) and (c) marginal revenue (MR).

Total Revenue : Total revenue is the total sales receipts of output produced over a given period of time. Total revenue depends on two factors : (a) the price of the product and (b) the quantity sold of the product. By multiplying the quantity sold (Q), by its selling price (P), we get total revenue. Thus TR = PxQ.

It may be illustrated with the following example. Suppose a producer intends to sell 1,00 units of a commodity at a price of Rs. 300 each. In this case total revenue will be TR = Rs. 300 × 100 = Rs. 30000/-.

AVERAGE REVENUE : By 'average revenue' we mean revenue obtained per unit of output sold. This is obtained by dividing total revenue by the number of units of output sold. That means,

$AR = \frac{TR}{Q}$, Q being total units of output sold.

In our previous example, total revenue is Rs. 30000/- and total output sold is 100 units. Hence, Average revenue $= \frac{Rs.30000/-}{100\,units}$ = Rs. 300/-

By definition, 'average revenue' is the price. Price is always expressed per unit. Per unit sales receipt is also called average revenue. Since sellers receive according to price, price and average revenue are one and the same thing. It may be proved in the following manner :

$AR = \frac{TR}{Q} = \frac{P \times Q}{Q} = P.$

MARGINAL REVENUE : Marginal revenue is the additional or extra revenue that a firm gets by selling one more unit of an item. It may simply be stated as the revenue

or sales receipt of the marginal unit of the firm's output. Algebraically, the marginal revenue of the n-th unit per period of time of a given product is the difference between the total revenue earned by selling (n – 1) units from the total revenue by selling n units per period of time.

That is, $MR_n = TR_n - TR_{(n-1)}$.

Here n stands for the number of units of output sold. It may be illustrated in the following manner. If a firm sells 10 units of output X for Rs. 300 each, its total revenue will be Rs. 3000/-. If it were to sell one more unit, i.e., a total of 11 units of X, it could do it at a lower price, say, Rs. 280 per unit. Its total revenue in that case will be Rs. 3080/-. In other words, the eleventh unit has made a net addition to only Rs. 80/- to its previous total revenue of Rs. 3000/- from the sale of 10 units.

By applying the above formula we get $MR_{11} = TR_{11} - TR_{10}$.

That is, MR = Rs. 3080/ – Rs. 3000/ = Rs. 80/-.

Marginal revenue is also defined to be the rate of change in total revenue when the increment in the sale of output is assumed to be one unit more. Hence, marginal revenue may also be defined as the ratio of change in total revenue to a unit change in output sold. It may, symbolically, be expressed as $MR = \frac{dTR}{dQ}$.

When the revenue function is given in the algebraic form, then to find marginal revenue we can, more easily, apply this form.

Example : Suppose the revenue function of a firm is given by $R = 500Q - 2Q^2$.

Then the marginal revenue will be $MR = \frac{dR}{dQ} = \frac{d}{dQ}[500Q - 2Q^2]$.

$$= 500\frac{dQ}{dQ} - 2\frac{d}{dQ}Q^2 = 500 - 2 \times 2Q = 500 - 4Q.$$

5a.9 The Relationship Between Average Revenue, Marginal Revenue and the Price Elasticity of Demand

We know that total revenue = Price × Quantity sold

That is, TR = P × Q.

Now, $MR = \frac{dTR}{dQ} = \frac{d}{dQ}[P \times Q] = P\frac{dQ}{dQ} + Q\frac{dP}{dQ} = P + Q\frac{dP}{dQ} = P\left[1 + \frac{Q}{P}.\frac{dP}{dQ}\right]$.

$$= P\left[1-\frac{1}{ed}\right] \quad \because ed = \frac{-dQ}{Q} \Big/ \frac{dP}{P}.$$

That means, MR = Price $\left(1-\frac{1}{ed}\right)$.

Some observations on this relation :

Observation 1 : Let $_{ed}$ = 1. That is, the commodity has unitary elasticity.

In this case, MR = $P\left(1-\frac{1}{1}\right) = 0$.

That is, MR is zero. So when the commodity has unitary elasticity, MR is zero. That means, total revenue is maximum. Thus, for maximization of total revenue, the commodity concerned should have unitary elasticity of demand.

Observation 2 : Let ed be greater than 1 (one), say, 3. In this case

MR = $P\left(1-\frac{1}{3}\right)$ = P × a positive fraction = Positive.

So, MR is positive. MR is positive means with more sale, total revenue will go on increasing.

Observation 3 : Let ed be less than one, say, $\frac{1}{3}$.

In this case, MR = $P\left(1-\frac{1}{\frac{1}{3}}\right)$ = P(1 – 3) = Negative.

That is, MR is negative. MR is negative means with more sale, total revenue will go on decreasing. This is, however, an unrealistic case.

Of the above three observations, only the first two are realistic.

However, many firms now-a-days intend to maximise revenue. For this the necessary condition is that the commodity concerned should have unitary elasticity of demand. This is also the condition for maximising the sales or revenue of a firm.

5a.10 Relation Between MR And AR Curves Under Different Forms of a Market

1. Since price elasticity of demand (ed) is, by definition, non-negative, the MR will always be less than price (= average revenue) at all levels of output. This is obvious because $MR = P\left(1-\frac{1}{ed}\right)$ and MR = AR = P.

Thus, graphically, the MR curve always lies below the AR curve.

2. Under a perfectly competitive market ed is infinity (∝). Hence in this form of market

$$MR = P\left(1-\frac{1}{ed}\right) = P\left(1-\frac{1}{\infty}\right) = P(1 - 0) = P.$$

Therefore, under perfect competition MR = Price = AR, that is, under perfect competition MR = Price = AR, that is, under perfect competition MR curve = AR curve and both are horizontal straight lines, being parallel to the horizontal axis. This has been demonstrated in the following diagram. (Fig. 5.1)

Fig : 5.1 Perfect competition

Also, since under perfect competition MR = AR = Price, we have

$$MR = \text{price}\left(1-\frac{1}{ed}\right).$$

That is, MR = Price = AR. This means that under perfect competition $1-\frac{1}{e}=1$. This implies that $\frac{1}{e}=0$. This is possible, if e = ∝.

Therefore, under perfect competition the commodity concerned has perfectly elastic demand.

3. Under monopoly, price or average revenue is always greater than marginal revenue.

We have $e = \frac{AR}{AR-MR}$.

However, under monopoly e is greater than zero but less than 1. Therefore, monopoly demand is inelistic. Hence, monopoly demand curve is always downard sloping. This has been presented in the following diagram. (Fig 5.2)

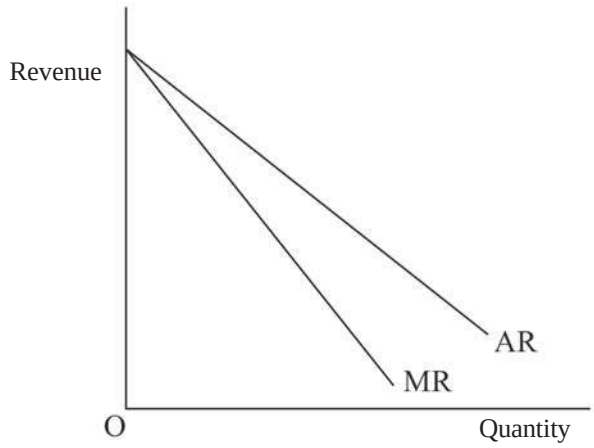


Fig. : 5.2 Monopoly

RELATION BETWEEN AVERAGE REVENUE (AR) AND MARGINAL REVENUE (MR)

From the above analysis let us now summarise the relation between AR and MR as under :

1. When AR is constant, MR is also constant and AR and MR are identical.

2. When AR increases, MR also increases. However, the rate of increase of MR is greater than the rate of increase of AR.

3. When AR decreases, MR also decreaes. However, the rate of decrease of MR is greater than the rate of decrease of AR. These three relations may be illustrated graphically in the following manner :

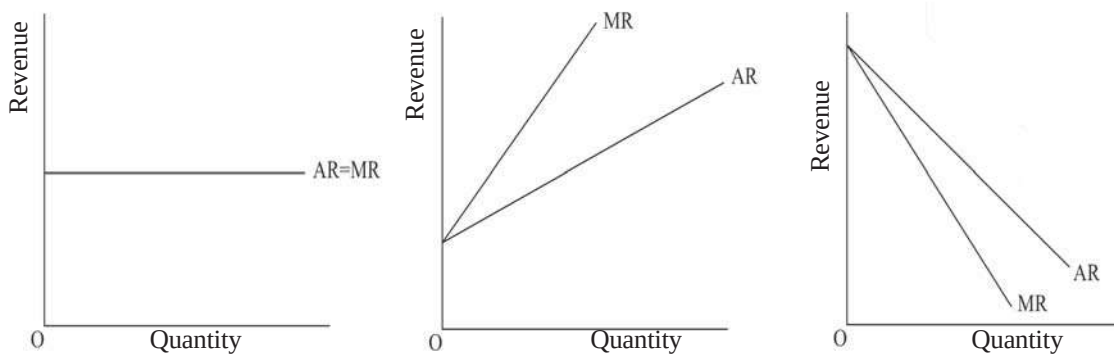


Fig : 5.3

GENERAL DISCUSSION OVER EQUILIBRIUM OF THE FIRM AND INDUSTRY

In Economics, equilibrium means a state of balance or a state of no change. This means that whenever an economic unit attains equilibrium, it does not want any further change from its position already attained.

Thus a producer is said to be in equilibrium when he gets maximum output with the least cost combination of various inputs.

THE CONCEPT OF A FIRM

A firm is, generally, a decision making business unit which organizes and directs the production of goods and services. Professor Watson in his “Price Theory and Its Uses” stated, “A firm is a unit engaged in production for sale at a profit and with the objective of maximizing the profit”. Naturally, a firm is in equilibrium when it has no intention either to expand or to contract its output. According to Professor Hanson, “A firm is said to be in equilibrium when the entrepreneur has no motivation to change its organization or its scale of production”.

ASSUMPTIONS MADE FOR DISCUSSING EQUILIBRIUM OF A FIRM

A. APPROACH OF TOTAL REVENUE AND TOTAL COST :

It is known to us that

Total profit = Total Revenue – Total Cost.

That is, π = TR – TC.

Thus, profits are the difference between total revenue (TR) and total cost (TC) per period of sales. If TR and TC curves of the firm are given, the profit maximizing level of output will be that level where the vertical distance between TR and TC is the greatest. Moving either to the left or to the right will reduce the amount of profit that will be earned by the firm. Naturally, the firm will alter its production upto the point where the difference between total revenue and total cost will be maximum. At this level of output the firm will have no intention to change its output and, naturally, the equilibrium will be a stable one.

EQUILIBRIUM OF THE FIRM UNDER THE PERFECTLY COMPETITIVE MARKET

Because of the various characteristics of a firm under a perfectly competitive market each firm in this market is just a price taker, not a price maker. Since the firm is simply

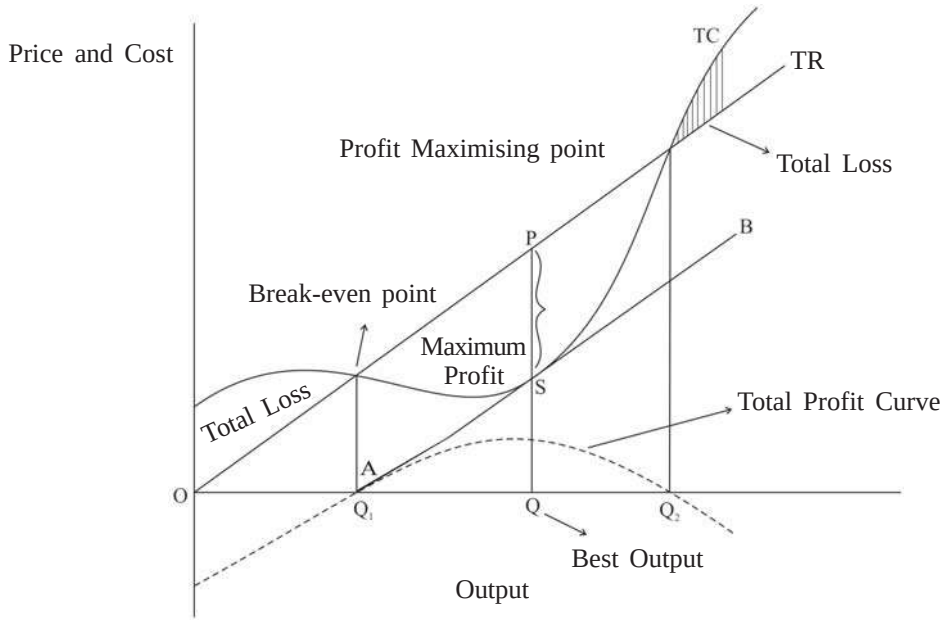


Fig. 5.4

a price taker, a firm can sell as much as it wishes though only at a given price. This is the reason why the total revenue curve of a firm in this market is an upward sloping straight line drawn at an angle of 45°. The shape of the TR curve shows that the total revenue of the firm is increasing with an increase in production at the same ratio. It has been demonstrated in the above diagram (Fig 5.4) measuring output horizontally while revenue and cost have been measured vertically.

In the above figure 5.4 TR is the total revenue curve while TC is the total cost curve of a firm under the perfectly competitive market. Initially, the firm has been incurring loss upto OQ_1 level of output. At point A where OQ_1 level of output gets produced, total revenue just covers total cost and so A is the break-even point. Beyond the point A any further increase in production leads to an increase in profits. Hence between outputs OQ_1 and OQ_2, total revenue is in excess of total cost. So profitable range of production is Q_1Q_2. However, profit is not maximum at all levels of output in the range between OQ_1 and OQ_2. Total profit is maximum at that point where the vertical distance between TR and TC becomes maximum. Such an output can be obtained by drawing a tangent to TC and that to TR. These two tangents are parallel to each other. The maximum distance between TR and TC is PS. The equilibrium level of output is OQ where total profit is maximum.

DEFECTS OF THE TR AND TC APPROACH

However, according to many economists, explanation of equilibrium of a firm by total revenue and total cost approch is cumbersome because of the following two reasons :

1. The maximum vertical distance between total revenue and total cost is not always easy to see at a glance.

2. Besides, it is not possible to discover the price per unit of the product at the first sight. To get the price per unit of output, we are to divide total revenue by total output sold.

5a.11 Marginal Revenue And Marginal Cost Approach

Because of the above-mentioed defects of TR and TC approach for explaining the equilibrium position of a firm another approach, that is, i.e., Marginal Revenue and Marginal Cost Approach has been introduced.

Mrs. Joan Robinson in her book, ‘The Economics of Imperfect Competition’ had used the tools of marginal revenue and marginal cost to explain the equilibrium of a firm.

We know that Total Profit = Total Revenue – Total Cost.

In symbols, π = TR – TC.

= R(Q) – C(Q), Q being the quantity of output.

For the maximisation of profit there are two conditions :

(a) Necessary or first order condition :

According to this condition, $\frac{d\pi}{dq} = 0$.

That is, $\frac{dR(Q)}{dQ} - \frac{dC(Q)}{dQ} = 0$.

That is, $R^1(Q) - C^1(Q) = 0$.

That is, MR – MC = 0.

That is, MR = MC.

Therefore, the equality between the marginal revenue and marginal cost is the necessary or the first order condition for the maximisation of profit.

(b) Sufficient or the second order condition

According to this condition, $\frac{d^2\pi}{dQ^2} < 0$.

This means that $\frac{d^2R(Q)}{dQ2} - \frac{d^2C(Q)}{dQ2} < 0$.

That is, $\frac{d^2R(Q)}{dQ2} < \frac{d^2C(Q)}{dQ2}$.

In language, this means that at the profit maximising point the rate of change in marginal revenue is less than the rate of change in marginal cost. To simplify this further, this condition states that just before attaining maximum profit, marginal revenue is more than marginal cost and just after attaining maximum profit marginal cost is greater than marginal revenue. To be more specific, at the profit maximising point the marginal cost curve will cut the marginal revenue curve from below.

A FIRM UNDER PERFECTLY COMPETITIVE MARKET

A perfectly competitive market is one in which there are a large number of buyers and sellers of a homogeneous product and neither a buyer nor a seller has any control over the price of the product. Such a market is characterised by the following features :

FEATURES OF A PERFECTLY COMPETITIVE MARKET

1. INNUMERABLE SELLERS AND BUYERS : Under perfect competition the number of sellers is assumed to be so large that the share of each seller in the total supply of the product is very small. Therefore, not a single seller can have any influence on the market price by changing his supply or can charge a higher or lower price. Obviously, firms are mere price takers, not price makers. In the same way, the number of buyers is so large that the share of each buyer in the total demand for the product is very small and that not a single buyer or a group of buyers can influence the market price.

2. HOMOGENEOUS PRODUCT : The commodities supplied by all the firms of an industry are assumed to be homogeneous or almost identical. Homogeneity of the product implies that buyers do not make any distinction between product supplied by various firms of an industry. Product of each firm is regarded as a perfect substitute for the products of all other firms. Hence, no firm can gain any competitive advantage over other firms. This assumption eliminates the power of all the suppliers of the firms to charge a price higher than the market price.

3. PERFECT MOBILITY OF FACTORS OF PRODUCTION : Factors of production are freely mobile between the firms. Labour can freely move from one firm to another or from one occupation to another, as there is no barrier to labour mobility. Similarly, capital can also move freely from one firm to another.

4. FREE ENTRY AND FREE EXIT : There is no legal or market barrier on the entry of new firms to the industry, nor is there any restriction on exit of the firms from the industry.

5. PERFECT KNOWLEDGE : Both the buyers and sellers have perfect knowledge about the market conditions. It means that all the buyers and sellers have full information regarding the prevailing and future prices and availability of the commodity concerned.

6. NO GOVERNMENT INTERFERENCE : There is no Government intervention in the market. This means that tariff, subsidies, rationing of production or demand and so on are ruled out.

7. ABSENCE OF COLLUSION AND INDEPENDENT DECISION MAKING BY FIRMS : In this market there is no collusion between the firms; they are not in league with one another in the form of guild or cartel. Nor are the buyers in any kind of collusion among themselves.

5a.12 Graphical Presentation of the Short Run Equilibrium of a Perfectly Competitive Firm.

Under the perfectly competitive market,

Total Revenue = price × quantity sold.

or, TR = pq.

Hence, Average Revenue, AR = $\frac{TR}{q} = \frac{p \times q}{q} = p$.

Thus, under perfect competition, AR is identical with price.

Now, marginal revenue, MR = $\frac{dTR}{dq} = p$.

Hence under perfect competition, p = AR = MR will be simply a horizontal straight line parallel to the horizontal axis as has been shown in figure 5.5(a)

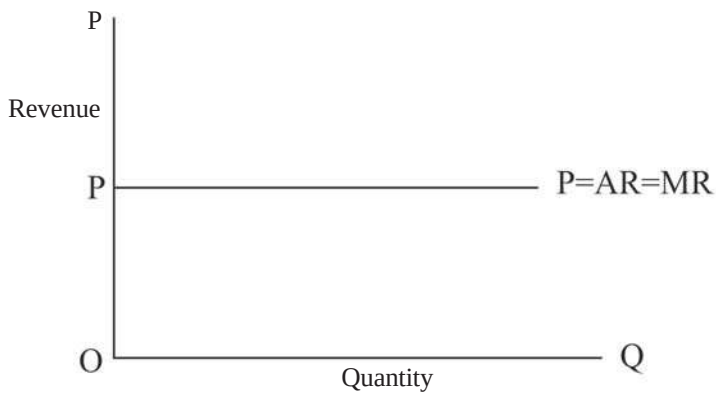


Fig. 5.5(a)

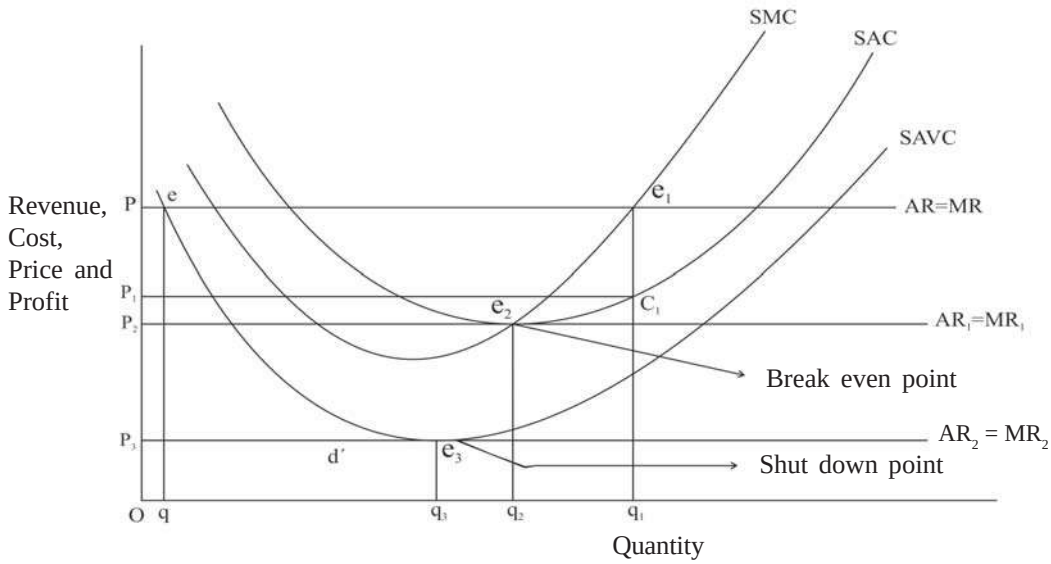


Fig. 5.5(b)

Since AR is a horizontal straight line, MR is also so and AR and MR coincide. In diagram 5.5(b) PAR = MR represents this feature of the market. Besides, SAC, SAVC and SMC have all usual shapes. Naturally, the vertical distance between SAC and SAVC indicates the AFC (= the average fixed cost).

In figure 5.5(b) at points e and e_1, MR = MC. Hence by virtue of the necessary condition, the firm should attain equilibrium at both these points. However, it cannot attain equilibrium at point e where AC > AR = price as this will lead to loss to the firm. Hence it will proceed to point e_1 where it will produce and will sell oq_1 units of output. At this stage its AR = OP = e_1q_1 so that its TR = $e_1q_1 \times oq_1$ = ☐ oPe_1q_1. Here AC is C_1q_1 so that TC is $oq_1 \times C_1q_1$ = ☐ $op_1c_1q_1$. Obviously, the firm attains profit to the tune of the area ☐ $PP_1c_1e_1$ which is referred to as abnormal or supernormal profit.

The difference between points e and e_1 is that at e only the necessary condition for the maximization of profit is getting satisfied while at e_1 both the necessary and sufficient conditions are satisfied. Hence it is usually said that "for maximization of profit the equality between MR and MC is only the necessary condition but not the sufficient one". Since the firm does not attain equilibrium at point e, it is generally said that "perfect competition is incompatible (inconsistent) with increasing returns to scale".

The possibility of earning supernormal profit will attract other firms to enter into the industry and as a consequence supply of the commodity will go on increasing. This will, obviously, lead to a fall in price of the commodity. Hence the new average revenue curve will be AR_1 and $AR_1 = MR_1$ and the new equilibrium point is e_2 where both the conditions for maximization of profit get satisfied. At e_2 the firm produces and sells oq_2 units of output and it is here where AC = AR = MR = price. This means that at e_2 only normal profit is attained : at e_2, SAC is tangent to AR and the point e_2 is referred to as the "break-even point" (that is, the normal profit-earning point). The break-even point is the point of tangency between the SAC curve and the AR curve. Even at this stage new firms will make entry into the market, leading to more supply, causing less price. So long as P < AC, the firm will incur loss but it will stay in the market in the hope of earning profit in near future. If price < AC, but p > AVC, then the entire amount of SAVC will be covered and a part of AFC will be covered. The firm will consider worthwhile to stay in the market so long as $P \geqq AVC$. However, if P < AVC, it will be uneconomic on the part of the firm to stay in the market. Hence the firm will remain in equilibrium at point e_3 where SAVC has become tangent to the $AR_2 = MR_2$. However, if price < SAVC, the firm will totally close down the business. That is why, the point 'e_3' is referred to as the "shut down point" where only total variable cost gets covered. So the shut down point is the point of tangency between SAVC and the AR curve.

From the above analysis we may thus say that a firm under perfectly competitive market will attain equilibrium in the short run if the following conditions are simultaneously stastisfied :

1. MR = MC.
2. MC cuts MR from below.
3. Price $\begin{matrix} > \\ = \\ < \end{matrix}$ AC

but price $\not<$ AVC.

5a.13 Long Run Equilibrium

In the long run all costs are variable, there is no question of having fixed cost.

In the short run, there is the possibility of getting supernormal profit, normal profit and even incurring loss. In the long run, however, supernormal profit will be competed away and the inefficient firms will leave the industry. Only the efficient firms will survive in the long run when they will have to be satisfied with the attainment of only normal profit. This situation has been demonstrated in figure 5.6 where along the horizontal axis we meausre quantity of output produced and sold whereas along the vertical axis revenue, cost, price and profit have been measured.

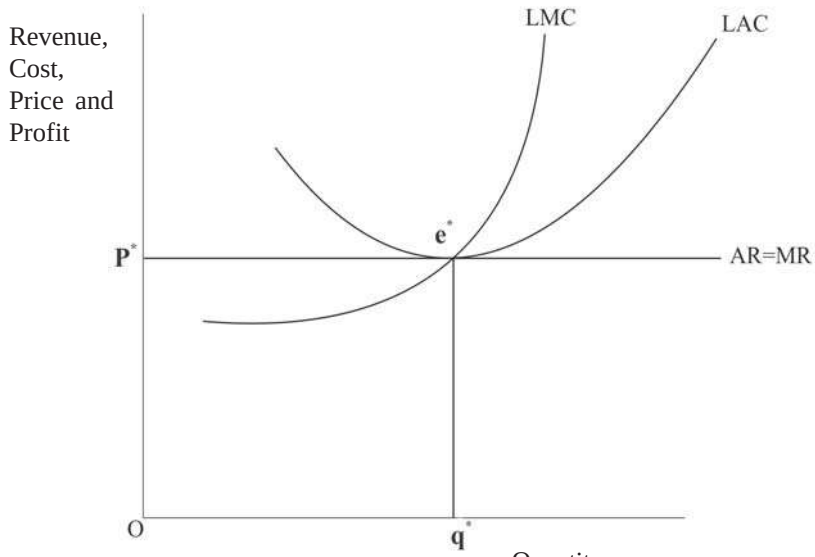


Fig. 5.6

In figure 5.6 the firm has attained equilibrium at point e^*, satisfying both the conditions for the attainment of maximum profit.

Hence in the long run a firm under perfectly competitive market will attain equilibrium when

1. MR = MC
2. MC cuts MR from below

and 3. MC = MR = AR = price = AC.

That means that in the long run, Price = Average cost. Therefore, in the long run the firm attains only normal profit and at the same time it produces at the minimum cost.

THE SHORT RUN SUPPLY CURVE OF A FIRM UNDER PERFECT COMPETITION

After discussing the profit maximizing level of output of a firm under the perfectly competitive market now we are interested in discussing how a firm under this market derives its supply curve.

The supply curve of a firm under a perfectly competitive market can be derived from the principle of the maximisation of profit. It is known to us that under the ceteris paribus clause (other things remaining the same), a supply curve gives a unique relationship between price and quantity supplied and the supply function is given by

$$\mathbf{s = f(p)}$$

$$\text{with } \frac{ds}{dp} > 0,$$

that is, the slope of the supply curve is positive.

We know that a firm under the perfectly competitive market attains equilibrium at the points of intersection of its marginal cost curve with successive average revenue curves. This can be illustrated in diagram 5.7 where along the horizontal axis we measure quantity while along the vertical axis revenue and cost have been measured. The firm under consideration is initially in equilibrium at point e_3 where two conditions for attaining equilibrium have been satisfied. The firm produces and sells oq_3 amount of output at price op_3 and it attains supernormal profit. When price comes down from OP_3 to OP_2 the profit maximizing conditions are satisfied at e_2 and the firm attains equilibrium at e_2 producing and supplying oq_2 amount of output at price OP_2. At this point price = oP_2 = e_2q_2 = average cost. Thus at point e_2 the firm is in a position to cover its entire cost, there is neither profit nor loss. The point e_2 is called the "break-even" point where only normal profit is earned. However, if price becomes less than OP_2 but more than OP_0, the firm will be in a position to cover the entire amount of the fixed cost (the vertical distance between AC and AVC is the

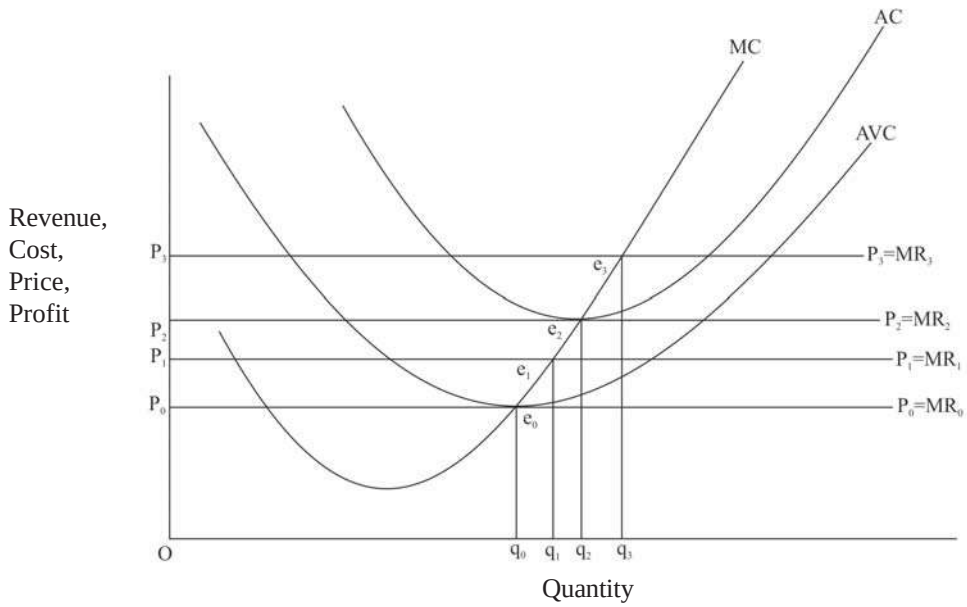


Fig. 5.7

averge fixed cost) and a part of the variable cost. The firm will continue production and supply

it in the market as it will be able to cover the total variable cost. Thus when price falls from OP_2 to OP_1 the firm produces and supplies oq_1 amount of output. In the same way when price comes down to OP_0 the firm attains equilibrium at point e_0 where both the conditions of attaining equilibrium get satisfied. So at price OP_0 the firm produces and supplies oq_0 amount of output. The point e_0, that is, the point of tangency between AVC and AR = MR line is known as the shut down point where the firm is able to cover only the variable cost (not even a portion of the fixed cost). If price comes below OP_0 the firm will not produce anything and will not supply any amount also.

From what has been stated above it is clear that the points e_0, e_1, e_2 and e_3 on the MC curve of a firm under the perfectly competitive market, in the short run, are on the supply curve. By joining these points, we get a curve known as the supply curve. This curve corresponds to SMC curve. Therefore, under the perfectly competitive market in the short run, the MC curve that lies above the AVC curve is the firm's supply curve in the short run. This may also be stated as the MC curve above the shut down point is the supply curve of a firm under perfect competition in the short run. Naturally, supply falls to zero if price of the commodity falls below AVC.

We may sum up the above analysis by stating that

$$S = S(P)$$

such that $S(P) > 0$ when $P \geq AVC$

and $S = 0$ when $P < AVC$.

However, there are many firms in a competitive industry, each producing that quantity of goods where its MC curve intersects AR = MR curve. All firms sell the product at the same price. Naturally, it is apparent that the industry supply curve, of course, in the short run under the perfectly competitive market, is equal to the horizontal summation of all firms' supply curve or MC curve above the minimum point of the AVC. Naturally, like the individual supply curve, the short run industry supply curve is upward sloping.

We end our discussion of the supply curve by saying the statement of Professor R. G. Lipsey : "The supply curve for a competitive industry is the horizontal sum of the marginal cost curves of all individual firms in the industry".

5a.14 Sample Questions

1. What are the main features of a perfectly competitive market?
2. What is the relation between Average Revenue, Marginal Revenue and Elasticity of demand?
3. What are the conditions of profit maximization by perfectly competitive firm?
4. How do you get the supply curve of a competitive firm? What is the relation between a short run supply curve and a long run supply curve?

Unit 5(b) ❐ Monopoly

Structure

5b.1 Objectives

The basic objectives of this unit are to discuss :

i. The Characteristics of Monopoly

ii. Causes for the Growth of Monopoly

iii. Objectives of the Monopolist

iv. Equilibrium of A Monopolist

5b.2 Introduction

The word 'monopoly' has been derived from the Greek word monopolein which means "alone to sell". This single seller is called a 'monopolist'.

Thus monopoly has been defined to be a market situation in which there is a single seller of a commodity for which there are no very close substitutes.

5b.3 Characteristics of Monopoly

The following are the main features of monopoly :

1. There is only one seller while there are a large number of buyers. This single seller is called a monopolist. However, there may be a group of sellers who may be called monopolists. For example, OPEC (Organisation of Petroleum Exporting Countries), an association of eleven "major producers of oil, has the monopoly power".

2. There is absence of competition.

3. There are no close substitutes of the product of the monopolist.

4. In case of monopoly the distinction between a firm and an industry disappears.

5. It is difficult for new firms to enter the market.

6. The monopolist has full control over the supply of the product.

7. Cross price elasticity of demand for a monopolist's product is zero in case of pure monopoly and very low in the case of simple monopoly (who will charge the same price for all the buyers).

5b.4 Causes for The Growth of Monopoly

There are a number of causes leading to the growth of monopoly :

1. There may be ownership of the strategic raw materials or exclusive knowledge of production techniques.

2. There may be patent rights for a particular product or for a production process.

3. There may be Government licensing or the imposition of foreign trade business to exclude foreign competition.

4. The size of the market may be such as not to support more than one plant of optimum size.

5. The monopolist may set limit price to prevent entry of new firms.

5b.5 Objectives of The Monopolist

The natural objective of the monoplist is, obviously, to maximize profit which is the difference between total revenue (TR) and total cost (TC).

This objective of the monopolist will, obviously, be fulfilled if two conditions are simultaneously fulfilled : (i) Necessary or first order condition : This condition is the equality between marginal revenue (MR) and marginal cost (MC) and (ii) Sufficient or the second order condition which states that at the profit maximizing point the marginal cost curve should cut the marginal revenue curve from below.

5b.6 Equilibrium of A Monopolist

(A) SHORT RUN EQUILIBRIUM : A monopolist faces a negatively sloping demand curve or average revenue (AR) curve. The reason behind this is that though he is a monopolist, if he wants to sell more he must lower down the price of his product. Since the average revenue curve is downward sloping, the corresponding MR curve lies below the AR curve, that is, AR > MR.

Since the monopolist is a price maker, he should take both price and output decisions. However, given the downward sloping demand curve, these two decisions are interdependent : fulfilment of one decision leads to the fulfilment of another decision. That is why, the monopolist will either set the price and sell the amount that the market will absorb or determine output which will be sold at the corresponding

price. However, for the monopolist the essential point for profit maximization is the fulfilment of both the first order and the second order conditions.

5b.7 Short Run Equilibrium

In the short run a monopolist can earn pure profit or economic profit as well as normal profit. In the short run a monopolist may also incur a loss. All these possibilities appear in the following figure where there are three diagrams : (a), (b) and (c). In each of these diagrams we measure output along the horizontal axis while revenue, cost and price have been measured along the vertical axis. The equilibrium point in each diagram is corresponding to this equilibrium point the monopolist produces OQ amount of output and sells it at a price OP.

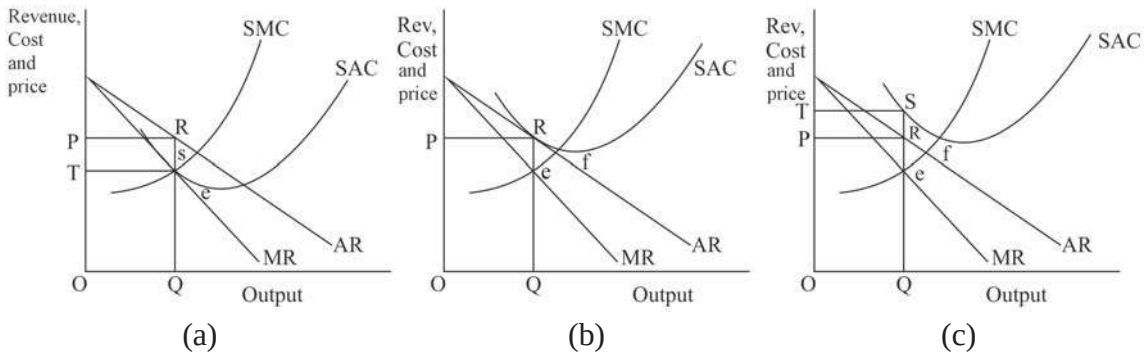


Fig. 5.8 : (a) Supernormal profit, (b) Normal profit and (c) Loss

From diagram (a) we see that the monopolist earns supernormal profit since at OQ level of output total revenue ☐ OPRQ is in excess of total cost ☐ OTSQ by ☐ TPRS which is the amount of supernormal profit. Thus the monopolist will earn supernormal profit after attaining equilibrium when MC = MR < AR > AC.

In other words, a monopolist will earn supernormal profit when the average cost curve will lie below the demand curve. Diagram (b) demonstrates the situation of earning normal profit since here total revenue ☐ OPRQ equals total cost ☐ OPRQ. Therefore, the condition for getting normal profit, that is, MC = MR < AR = AC has been satisfied.

That means that a monopolist will be satisfied with normal profit when the AC curve at a particular point coincides with the AR curve.

In the short run, the monopolist may also incur loss. This case has been demonstrated in figure (c) above where at the equilibrium point e total revenue of the monopolist

is ☐ OPRQ while total cost is ☐ OTSQ and consequently, the amount of loss is ☐ TPRS. Thus a monopolist will incur loss in the short run when the AC curve lies above the AR curve at the equilibrium level of output. But the monopolist will continue production in the short run if he can cover up only average variable cost. This implies that if the loss of the monopolist equals total fixed cost, the monopolist will stay in business. However, if the monopolist fails to cover up the variable cost, that is, if P < AVC, only then he will stop production. It may be noted that there cannot be loss in the long run.

5b.8 Long Run Equilibrium

In the long run, the monopolist can change his scale of production in such a way that he may attain the greatest amount of profit. However, since entry is blocked, excess profit will be enjoyed by the monopolist. In the long run a monopolist may also earn normal profit. However, in the long run a monopolist will not incur any loss.

5b.9 Some Interesting Points About Monopoly Equilibrium

1. The monopolist will never attain equilibrium in the inelastic portion of his demand curve or AR curve. This means that a monopolist attains equilibrium at the elastic range of his demand curve.

The primary goal of the monopolist is to attain maximum profit. This can be attained when MC = MR. But $MR = AR\left(1-\frac{1}{ed}\right)$. If ed > 1 at a particular point on a straight line demand curve, then naturally, MR becomes positive (that is, MR > 0). Costs are always positive (costs can never be negative or it cannot be zero also). Equilibrium must occur when positive MC equals positive MR.

However, if ed < 1, MR will be negative. Negative MR can never be equal to positive MC. Therefore, in this range of the demand curve, a monopolist can not attain equilibrium.

Finally, when ed = 0, MR will be zero. Equilibrium can not be attained even in this case because positive MC cannot be matched with zero MR.

From what has been stated above, it follows that a monopolist always operates and attains equilibrium only at the elastic zone (ed > 1) of his demand curve in order to attain maximum profit.

2. A monopolist never operates at the lowest point of his average cost curve where cost comes to be the least. More specifically, the monopolist operates somewhere to the left of the minimum point of his AC curve, whether in the short run or in the long run. This happens because of the nature of the demand curve. Since the demand curve is downward falling, it must be tangent to the falling portion of the Ac corve. Since the monopolist operates at this point, it implies that the monopolist firm cannot utilise its plant optimally. This also implies the under-utilisation of resource. Since entry is closed, there is no urgency on the part of the monopolist to operate at the lowest point of its AC curve. On the contrary, a competitive firm always operates at the lowest point of its AC curve since in this case the demand curve is perfectly elastic.

5b.10 Monopoly Supply Curve

It is interesting to note that a monopoly firm has no well-defined supply curve. This may be due the fact that there is no unique supply curve for the monopolist to be derived from his marginal cost curve. We know that under perfect competition, the short run MC curve above the shut down point is the supply curve which shows a unique relationship between price and quantity. At a particular price, a particular amount of the commodity will be supplied. However, under monopoly, there is no such one-to-one correspondence between price and quantity supplied. This is due to the fact that output decision of a monopolist does not only depend on MC but also on the shape of the demand curve. Consequently, shifts in demand do not trace out a series of prices and quantities as happens with a competitive supply curve.

It may happen that a monopolist, given his MC curve, may supply a particular quantity at different prices depending on the elasticity of demand. Hence, the construction of a supply curve from the marginal cost curve is not possible under monopoly and for that matter under any branch of imperfect competition.

5b.11 Price Discriminating Monopolist

A monopolist may, very often, charge different prices for his product to different customers without making any qualitative change for his product. This practice of charging different prices for the same product to different customers is known as price discrimination. A monopoly firm being the only one seller in the market is free to charge different prices to different markets or buyers when conditions are appropriate for this. If the firm follows such a policy in practice, we will call it price discrimination. Instead of selling all output at a single price, the monopoly firm charges higher price for a part of its output and thereby increases its revenue.

5b.12 Forms of Price Discrimination

Price discrimination may take many forms. The common forms of price discrimination may, briefly, be stated as follows :

(A) PERSONAL DISCRIMINATION : Generally, depending upon the economic status of buyers, different prices may be charged to different buyers in providing similar service. For instance, a surgeon may charge a high operation fee to a rich patient and a lower one to a poor patient.

(B) AGE DISCRIMINATION : On the basis of age of the buyers price discrimination may be possible. For example, a barber may charge lower rates for children's hair cuts than those for adults. In railways and bus transport services, it is a commonly adopted form of price discrimination that persons below 12 years of age are charged at half the rate.

(C) SEX DISCRIMINATION : In selling some goods, producers may discriminate between male and female buyers by charging low prices to females. For example, an organizing firm may provide seats to ladies at concessional rates.

(D) SIZE DISCRIMINATION : On the basis of size of quantity of the product, different prices may be charged. For example, an economy size tooth paste tube is relatively cheaper than a small size tube. Similarly, a product is sold in the retail market at a higher price than in the wholesale market by the producer.

(E) USE DISCRIMINATION : Sometimes, depending upon the kind of use of the product, different rates may be charged. For example, an electricity distribution company may charge low rates for industrial use as compared to the high rates for AC.

(F) NATURE OF COMMODITY DISCRIMINATION : Sometimes, because of the nature of a commodity, price discrimination may be made. For example, freight charges by the railways are different for coal and iron for the same distance.

(G). TIME DISCRIMINATION : On the basis of time of service, different rates may be charged. For example, cinema houses charge lower rates of admission for mornings and matinee shows than for regular shows.

5b.13 Necessary Conditions For Making Price Discrimination Possible

Under price discrimination different buyers pay different prices. So, price in one market must be higher and lower in another market. The following two necessary condtions must be satisfied for the system of price discrimination to function.

(i) For making price discrimination possible, goods must not be transferred from one market to another. That means, price discrimination must preclude resale of the goods. To be more specific, price discrimination will break down if buyers purchase the monopoly product from the low-priced market and sell it in the high-priced market.

(ii) Demand must not be shifted from the dearer market to the cheaper one. This means that price discrimination will not be possible if buyers of the product transfer themselves from one market to another market.

(iii) It may happen that buyers do not know what prices that the monopolist charges are prevailing in different markets. Such ignorance on the part of the buyers may cause price discrimination to exist. Besides, laziness of consumers may lead to differences in prices of the monopoly product. Customers may care little for the difference in prices though they know the difference in prices. Moreover, customers may care little for the differences in prices though they know the difference in prices. Nature of the product may demand price discrimination to stay. Personal serives of say, doctors, lawyers are such that price discrimination is supposed to exist. Finally, there may be legal sanction for price discrimination. For illustration, the railway charges different fares to different passengers.

5b.14 Conditions for Successful Price Discrimination

We may point out that even if price discrimination is possible, it may not be profitable for the monopolist to discriminate prices. Charging two different prices to different buyers may not necessarily result in higher profit. To make price discrimination profitable, the following three conditions must be fulfilled :

(1) The seller must have a monopoly power. This means that it can be maintained only under monopoly. It can not be done either in perfect competition or in monopolistic competition.

(2) The elasticities of demand in different markets must be different. Under this situation the monopolist will charge higher price in the market having relatively inelastic demand and lower price in the market having relatively elastic demand.

Proof : Let there be two markets 1 and 2 having respective prices p_1 and p_2 and respective elasticities of demand e_1 and e_2.

Naturally, $MR_1 = p_1\left(1-\frac{1}{e_1}\right)$ and $MR_2 = p_2\left(1-\frac{1}{e_2}\right)$.

The condition for equilibrium is $MR_1 = MC$

and $MR_2 = MC$ so that $MR_1 = MR_2$.

Hence $= p_1\left(1-\frac{1}{e_1}\right) = p_2\left(1-\frac{1}{e_2}\right)$.

That means, $\frac{p_1}{p_2} = \frac{\left(1-\frac{1}{e_2}\right)}{\left(1-\frac{1}{e_1}\right)}$.

Now, assume $e_1 = 3$ and $e_2 = 5$ so that $\frac{p_1}{p_2} = \frac{1-\frac{1}{5}}{1-\frac{1}{3}} = \frac{\frac{4}{5}}{\frac{2}{3}} = \frac{6}{5}$.

Thus $5p_1 = 6p_2$. That is, $p_1 > p_2$.

Hence when $e_1 < e_2$, $p_1 > p_2$ proved.

5b.15 Relation Between Elasticity Of Demand And The Policy Of Price Discrimination

Let there be two markets 1 and 2 having respective elasticities e_1 and e_2, respective prices p_1 and p_2 and respective marginal revenues MR_1 and MR_2.

Hence, $MR_1 = AR_1\left(1-\frac{1}{e_1}\right)$ and $MR_2 = AR_2\left(1-\frac{1}{e_2}\right)$.

For equilibrium, $MR_1 = MR_2 = MC$.

That means $AR_1\left(1-\frac{1}{e_1}\right) = AR_2\left(1-\frac{1}{e_2}\right)$.

That is, $\frac{AR_1}{AR_2} = \frac{P_1}{P_2} = \frac{1-\frac{1}{e_2}}{1-\frac{1}{e_1}}$.

If $e_1 = e_2$, then $\frac{P_1}{P_2} = 1$. That is, $P_1 = P_2$.

This amounts to saying that if elasticities of demand are of uniform nature in all markets, it will not be possible and also profitable too, to discriminate between buyers.

That means, the monopolist will charge the same price to all buyers.

On the other hand, if $e_1 > e_2$ then $\frac{1}{e_2} > \frac{1}{e_1}$.

That is, $\left(1-\frac{1}{e_1}\right) > \left(1-\frac{1}{e_2}\right)$. That is, $\frac{1-\frac{1}{e_2}}{1-\frac{1}{e_1}} < 1$. That is, $\frac{p_1}{p_2} < 1$, that is, $p_2 > p_1$.

That means, if $e_1 < e_2$, then $p_1 > p_2$.

That is, the monopolist will charge a higher (lower) price in the market in which demand is less (more) elastic.

But it must be noted that the coefficient of elasticities of demand for the product in all markets should have a value greater than one. This is because of the fact that the monopolist does always operate at the elastic zone (ed > 1) of his demand curve.

5b.16 Equilibrium Conditions Under Price Discrimination

Let us assume that the monopolist concerned produces q units of a commodity. Of these q unts he intends to sell q_1 units in market 1 and q_2 units in market$_2$ so that $q_1 + q_2 = q$. Naturally, profit of the monopolist will be

$\pi = TR - TC = TR_1 + TR_2 - TC$

$= R_1(q_1) + R_2(q_2) - C(q_1 + q_2)$.

For maximization of profit we are to set $\frac{\partial \pi}{\partial q_1} = 0$ and $\frac{\partial \pi}{\partial q_2} = 0$.

Now, $\frac{\partial \pi}{\partial q_1} = \frac{\partial R_1(q_1)}{\partial q_1} - \frac{\partial C(q_1 + q_2)}{\partial q_1} = 0$.

That is, $R^1(q_1) - C^1(q_1 + q_2) = 0$.

That is, $R^1(q_1) = C^1(q_1 + q_2)$.

That is, $MR_1 = MC$.....(a)

Again, $\frac{\partial \pi}{\partial q_2} = \frac{\partial R_2(q_2)}{\partial q_2} - \frac{\partial C(q_1 + q_2)}{\partial q_2} = 0$.

That is, $R_2^1(q_2) - MC = 0$.

That is, $MR_2 = MC$.....(b)

Hence, for equilibrium under price discrimination,

$MR_1 = MR_2 = MC$.

This is the necessary condition for equilibrium.

We are to remember that at the stage of production no discrimination had been made.

The sufficient condition is that at the profit maximizing point the MC curve should cut the MR curve from below.

5b.17 Graphical Presentation of Equilibrium of a Discriminating Monopolist

The equilibrium position of a price discriminatory monopolist can, graphically, be, presented in the following diagram where there are three diagrams : 1, 2 and 3. In diagram 1 the AR and MR curves show the average and marginal revenue curves of the monopolist in market 1 where a part of his total output is being sold. Diagram 2 reveals the corresponding AR and MR curves of the monopolist in market 2. Finally, diagram 3 reveals

the equilibrium situation of the discriminatory monopolist. From the diagrams it is clear that the demand curve in market 1 is relatively inelastic while that in market 2 is relatively elastic.

In diagram 3, the combined marginal revenue curve, that is, $\sum_{i=1}^{2} MRi$ has been obtained by the lateral summation of the two marginal revenue curves in the two markets. In this diagram the equilibrium point has been attained where $\sum_{i=1}^{2} MRi = MC$ and the MC cuts the $\sum_{i=1}^{2} MRi$ from below. Therefore, the monopolist sells oq units of his output

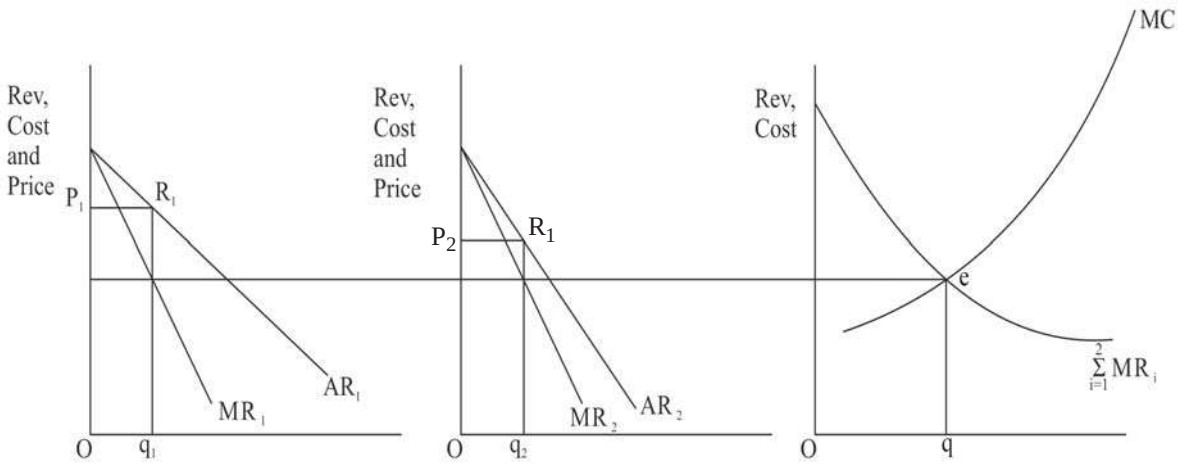


Fig. 5.9

The monopolist has to allocate total output oq in the two sub-markets in such a way that MR in each of the two sub-markets is equal to eq. From the diagrams it is clear that the monopolsit sells oq_1 units of output oq in market 1 at price OP_1 and oq_2 units of output oq in market 2 at price OP_2. Naturally, $oq_1 + oq_2 = oq$. It is to be noted here that demand for the monopolist's output in market 1 is relatively inelastic while that in market 2 is relatively elastic, that is, $e_1 < e_2$. Obviously, $OP_1 > OP_2$.

5b.18 Degree of Monopoly Power

An important feature of Monopoly is that the monopolist enjoys some power commonly known as monopoly power. By this power is meant the amount of discretion which a monopolist enjoys with regard to price and output policy. That means, the monopoly power indicates the degree of control which a seller exerts over the price and output level of his product. It may also indicate the deviation from perfect competition. There are a number of measures of this degree of monopoly power.

1. Lerner's Measure : The divergence between price and marginal cost is the main source of monopolistic exploitation. Professor Abba P. Lerner has expressed this gap per unit of price as

$$\text{DMP} = \frac{P - MC}{P} = \frac{P - MR}{P} = \frac{P - P\left(1 - \frac{1}{e}\right)}{P} = \frac{1}{e}.$$

Thus the index of monopoly power varies inversely with the elasticity of demand. The greater the absolute price elasticity of demand the lower the degree of monopoly power and vice versa.

Since under perfect competition, P = MC, the degree of monopoly power is zero under this form of market.

Alternatively, we know that under perfect competition $|e| = \infty$. Hence under perfect competition, index of monopoly power = $\frac{1}{|e|} = \frac{1}{\infty} = 0$. On the other hand, when the monopolist has no cost of production, that is, MC = 0, then the index of monopoly power is unity. Hence Lerner's measure of the index of monopoly power lies between zero (0) and unity (1). It is zero under perfect competition while unity when there is costless monopoly.

2. Size of supernormal profits : Another indicator of the strength of monopoly power is the size of supernormal profits. Supernormal profit, under perfect competition, in the long run, is zero. However, it is, generally, positive under monopoly. Hence the size of supernormal profits has been a measure of the degree of monopoly power. We know that supernormal profits arise when P > AC. Hence (P – AC) may be regarded as an indicator of monopoly power.

3. Rate of profit : The rate of profit may also be used as a meausre of the degree of monopoly power. The rate of profit has been defined as the ratio of total profit to total capital employed. The higher the rate of profit, the greater is the degree of monopoly power. The rate of profit is zero in the long run under perfect competition and hence the monopoly power is zero, though it is positive under monopoly.

4. Cross price elasticity as a measure : The concept of cross price elasticity of demand as a measure of the degree of monopoly power has been used by Professor Triffin.

The cross price elasticity of demand has been defined as the percentage change in sales of the i-th firm due to one percentage change in price charged by the j-th firm. In symbols,

$$e_{ij} = \frac{dq_i}{q_i} \Big/ \frac{dp_j}{p_j} = \frac{dq_i}{q_i} \times \frac{p_j}{dp_j} = \frac{p_j}{q_i} \cdot \frac{dq_i}{dp_j}.$$

Obviously, $\frac{dq_i}{dp_j}$ will be zero under pure monopoly while for perfect competition $\frac{dq_i}{dp_j}$ is infinity (∞). Hence e_{ij} will be zero for pure monopoly whereas e_{ij} is infinity under perfect competition.

Therefore, the lower the value of cross price elasticity of demand, the greater the degree of monopoly power and vice versa. However, if this cross price elasticity is finite, there is neither pure competition nor is there pure monopoly.

5b.19 Multiplant Monopoly

Sometimes a monopolist may produce his output to be sold in the market in several plants with, obviously, different cost conditions. In this case the market is said to be the multiplant monopoly. Let us assume that the monopolist produces his entire output in two different plants with different cost structures. We also assume that the monopolsit knows his market demand. In such a situation the monopolist has to solve two problems : (a) how much total output is to be produced and sold and (b) how much total amount of output is to be distributed between the two plants. We assume that the monopolist wants to maximize his profits.

Let us assume that q_i is the amount of output to plant, be produced in the i-th plant, i = 1, 2. Let ci be the total cost to be incurred by the i-th plant i = 1, 2. Hence the cost functions of the two plants are $c_1 = c_1(q_1)$ and $c_2 = c_2(q_2)$. The revenue function can be written as $R = R(q_1 + q_2)$.

Hence total profit of the monopolist will be

$\pi = R(q_1 + q_2) - C_1(q_1) - C_2(q_2)$.

The first order condition for maximization of profit needs:

$\frac{\partial \pi}{\partial q_1} = 0$. That is, $\frac{\partial R}{\partial q_1} - \frac{\partial C_1}{\partial q_1} = 0$, that is, $MR = MC_1$.

Also, $\frac{\partial \pi}{\partial q_2} = 0$. That is, $\frac{\partial R}{\partial q_2} - \frac{\partial C_2}{\partial q_2} = 0$ That is, $MR = MC_2$.

Hence the first order condition for the maximization of the monopoly profit is

$$MR = MC_1 = MC_2.$$

The second order condition for the maximization of profit needs

$$\frac{\partial^2 \pi}{\partial q_1^2} < 0 \text{ or, } \frac{\partial^2 R}{\partial q_1^2} - \frac{\partial^2 C_1}{\partial q_1^2} < 0.$$

That is, $\frac{\partial^2 \pi}{\partial q_1^2} < \frac{\partial^2 C_1}{\partial q_1^2}$.

Also, $\frac{\partial^2 \pi}{\partial q_2^2} < 0$. That is, $\frac{\partial^2 R}{\partial q_2^2} - \frac{\partial^2 C_2}{\partial q_2^2} < 0$.

That is, $\frac{\partial^2 R}{\partial q_2^2} < \frac{\partial^2 C_2}{\partial q_2^2} < 0$.

In language, the second order condition for profit maximization requires that the monopolist's MC in each plant should be steeper than MR.

The equilibrium of the multiplant monopolist can, clearly, be explained with the help of the following diagram where along the horizontal axis quantity sold is measured while revenue and cost are measured along the vertical axis.

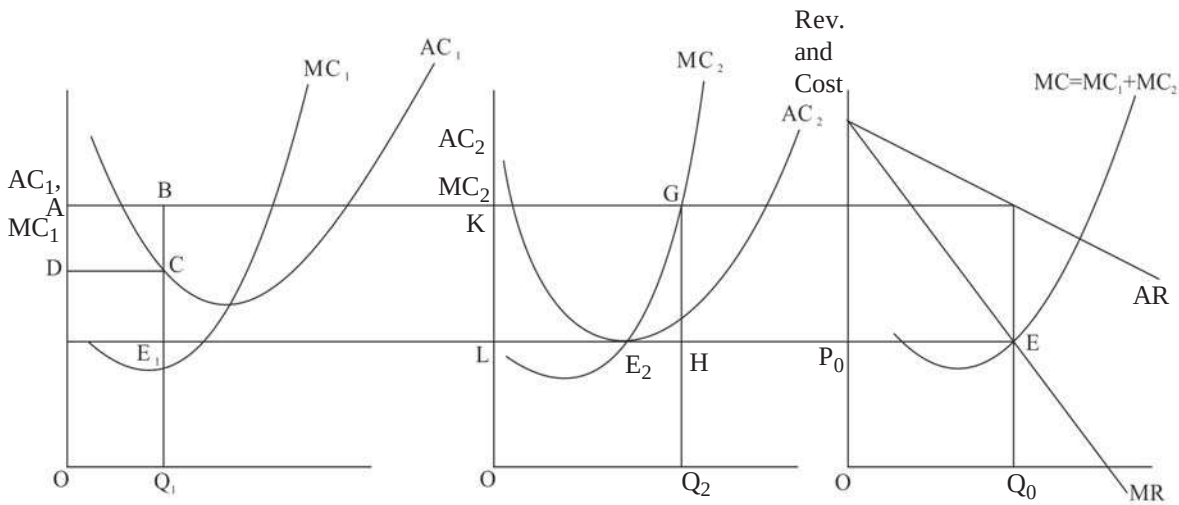


Fig. 5.10

In the first panel of the above diagram we have drawn the AC and MC curves for plant A in panel A and those for plant B in panel B. By the horizontal summation of the MC curves, we get the combined MC curve in panel C. In this panel AR and MR curves have also been drawn. The equilibrium price and output levels have been determined at the intersection point of the MR and MC curves, that is, at point E. Obviously, the equilibrium level of output is OQ_0 while the equilibrium price is OP_0. The multiplant monopolist, in this way, will decide how much output is to be produced and at what price it is to be sold.

The next task of the monopolist will be to determine the levels of output to be produced in each plant. For that we have drawn a horizontal line from E towards the left. This line has cut MC_2 at H while MC_1 at E_1. At these points the profit maximizing conditions (that is, $MR = MC_1 = MC_2$) have been satisfied. From the diagram we note that OQ_1 amount of output will be produced in plant 1 while OQ_2 amount of output will be produced in plant 2. Naturally, $OQ_1 + OQ_2 = OQ_0$. The multiplant monopolist will, in this way, distribute his output among his plants for the maximization of profit.

Finally, by drawing the respective AC curves for the two plants we can find out the profits of the monopolist from the two plants. From the diagram we note that the monopolist will get profit equalling to the area ☐ ABCD from the first plant while he will get profit equalling to the area ☐ KLHG from the secondplant.

5b.20 Degrees of Price Discrimination

Professor Pigou has distinguished between different types of price discrimination in the following three manners :

1. Price discrimination of the first degree.
2. Price discrimination of the second degree. and
3. Price discrimination of the third degree.

Let us discuss these three degrees of price discrimination one by one.

1. **Price discrimination of the first degree :** This type of price discrimination is said to occur when the monopolist is able to sell separate unit of the output at a different price to the same buyer. That means, it involves maximum possible exploitation of each buyer. That means, it involves a complete shift of consumers' surplus to the seller. Such discrimination is also called perfect price discrimination. In negotiating with each buyer the monopolist charges him the maximum price he is willing to pay under the threat of denying selling of any quantity to him. That means, he offers every buyer a "take-it-or-leave it" choice. This type of price discrimination has been presented in the following figure where the demand curve also becomes the MR curve of the monopolist whose total revenue for OX amount of output has been shown by the shaded area.

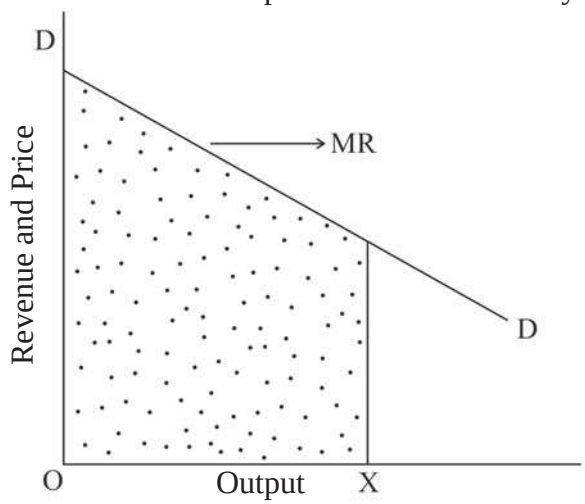


Fig. 5.11

2. **Price discrimination of the second degree :** In such cases the goods are divided into different blocks of units and for each block a different price is charged. Such a price discrimination is possible if each individual buyer has perfectly inelastic demand for the product below and above a certain price. Such a case has been shown in figure 5.11 where DD is the market demand curve. This market is divided into four groups. In

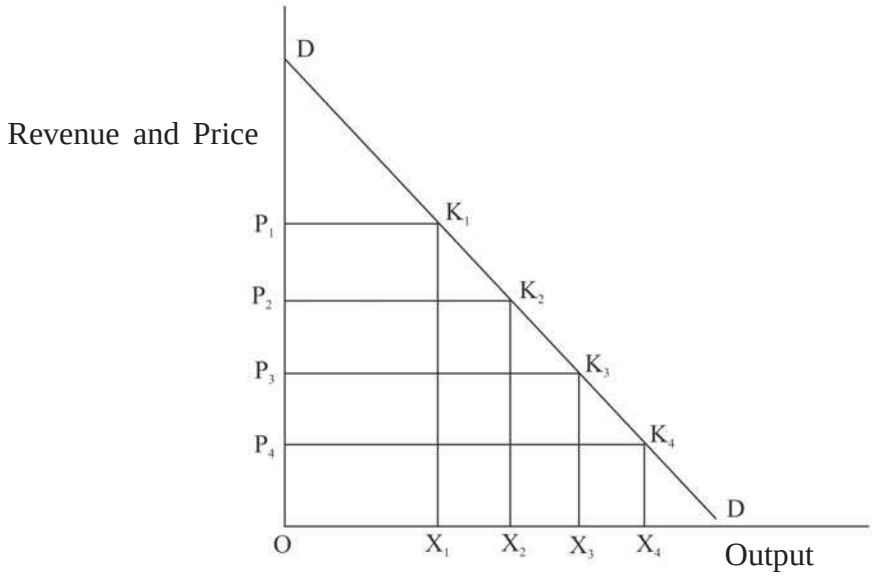


Fig. 5.12

the first group the X_1th unit of the good has demand price equal to OP_1. So all the buyers in this group pay OP_1 price and the group gets DK_1P_1 as consumer's surplus. Similarly, for group 2, 3 and 4 consumers pay OP_2, OP_3 and OP_4 and get the consumers' surplus equivalent to DK_2P_2; DK_3P_3 and DK_4P_4 respectively. In this manner the seller divides the total output into different groups and sub-groups in such a way that for each group/sub-groups he charges a different price.

3. Price discrimination of the third degree : This is said to occur when the seller divides buyers according to their income, location, types of uses of the product etc. and charges different prices from each group of buyers. Thus the seller is at an advantage segregating the buyers into different groups and charging them different prices so long as the demand elasticities of these groups differ. Naturally, this is the most common type of price discrimination.

5b.21 Sample Questions

A. Objective-type Questions :

(a) (i) A monopolist can charge any price to maximise profits, (ii) A monopoly firm can fix its price any where along the demand curve, (iii) If monopoly's

MC = 0, it fixes its price where e = 0, (iv) The slope of the monopoly's MR curve is twice that of AR curve.

Ans. (iv)

(b) Under monopoly the supply curve has (i) A positive slope, (ii) A negative slope, (iii) A slope = 0, (iv) None of the above.

Ans. (iv)

(c) What is the shape of the demand curve faced by a firm under perfect competition? (i) Horizontal, (ii) Vertical, (iii) Positively sloped, (iv) Negatively sloped.

Ans. (i)

(d) Which is the other name that is given to the average revenue curve? (i) Profit curve, (ii) Demand curve, (iii) Indifference curve, (iv) Average cost curve.

Ans. (ii)

(e) Price discrimination will be profitable only if the elasticity of demand in different markets in which the total market has been divided is (i) Uniform, (ii) Different, (iii) Less, (iv) Zero.

Ans. (ii)

(f) Discriminating monopoly implies that the monopolist charges different prices for his commodity : (i) From different groups of consumers, (ii) For different uses, (iii) At different places, (iv) Any of the above.

Ans. (iv)

(g) Which is the first order condition for the profit of a firm to be maximum? (i) AC = MR, (ii) MC = MR, (iii) MR = AR, (iv) AC = AR.

Ans. (ii)

(h) Under which of the following forms of market structure does a firm have no control over the price of its product? (i) Monopoly, (ii) Monopolistic competition, (iii) Oligopoly, (iv) Perfect competition.

Ans. (iv)

B. Short-type Questions :

(a) Explain the shape of the demand curve of a firm operating under a perfectly competitive market.

(b) Calculate the ratio of prices charged by a discriminating monopolist in two markets, A and B, having price elasticities of demands as –0.5 and –1.5 respectively.

(c) A monopolist's marginal cost is MC = 50 + 3Q and the demand for its product is Q = 150 – P. Find out the profit maximizing output and also the equilibrium price.

(d) A monopolist's demand curve is P = 200 – 5q. Find the MR function of the monopolist.

(e) Suppose the demand function of a firm is given by P = 20 – Q and the total cost function is given by $C = Q^2 + 8Q + 2$. Find out the output level that maximizes total profit. Also find out the corresponding price.

(f) Show that the profit-maximizing quantity for a monopolist will always lie in the elastic region of the demand curve.

(g) If perfect competition is rare in the world of business, why do we study it?

(h) Explain the differet types of market structure.

(i) Make a clear distiction between break even point and shut down point.

(j) Explain the reasons behind the growth of monopoly.

C. Essay-type Questions :

1. Analyse the conditions for equilibrium of a perfectly competitive firm in the short run.

2. Explain how the short run supply curve of a perfectly competitive firm is related to its marginal cost curve.

3. "The equality between marginal revenue and marginal cost is only necessary but not a sufficient condition for maximization of profit in a perfectly competitive market". Explain, in detail.

4. What is monopoly? What are its basic features? How does monopoly arise?

5. What is price discrimination? When is price discrimination possible and when is it profitable?

6. Discuss how the equilibrium price and quantity are determined under monopoly.

7. A monopolist faces the demand curve given by p = 20 – q and his cost function is given as $c = q^2 + 8q + 2$. Determine the profit maximizing output and the corresponding price of the monopolist.

Unit 6(a) ❐ Monopolistic Competition

Structure

6a.1 Objectives

The primary objectives of this unit are :

i. To define Monopolistic Competition, indicate its characteristic features and explain its equilibrium.

ii. To explain the concept of product differentiation.

iii. To discuss selling cost and its implications.

6a.2 Introduction

Till 1920's perfect competition and monopoly had been the two favourite analytical models of the economists. However, these two extreme market forms are far from reality. What we find in reality is some sort of a middle group between these two extreme forms. Most of the producing firms are not monopolists since they are to face competition against each other. On the other hand, these firms are not price-takers since they do not operate under perfectly competitive conditions. This sort of intermediate market structure was built up independently by late Professor Mrs. Joan Robinson of England in 1933 and Professor E. C. Chamberlin of the USA also in 1933. It may be mentioned

that Professor Robinson called this newly built but realistic market form as “Imperfect Competition” whereas Professor Chamberlin called his market as “Monopolistic Competition”. However, we are interested in going through Chamberlin’s model of “Monopolistic Competition”.

DEFINITION : Monopolistic Competition refers to a market structure in which a large number of sellers sell differentiated products which are close substitutes for one another. It may be recalled that a close substitute is one whose cross elasticity is unity or greater. This form of market combines the basic elements of both perfect competition and monopoly. The element of monopoly in monopolistic competition arises from the fact that each firm has an absolute right to produce and sell a branded or patented product. Other firms are prevented by laws from producing and selling a branded product of other firms. This gives a firm monopoly power over production, pricing and sale of its own branded product. The element of competition comes from the fact that each branded product is a close substitute for another and firms selling branded products of the same generic category have to compete for the market share.

6a.3 Characteristics of Monopolistic Competition

The monopolistically competitive market has the following characteristics :

1. There are a large number of buyers and sellers. However, the number of sellers is not as large as that under perfect competition.

2. Product differentiation : Each seller sells a differentiated product. This implies that the product of each seller is different from that of others. The difference between the products of different sellers is either in quality or in labels or trade marks. Hence, the product of every seller is different. Viewing from this angle, each seller is a monopolist. However, under pure monopoly the product has no close substituties. But under monopolistic competition the products of different sellers are close substitutes of one another. One product can be used in place of the other. Hence, under monopolistic competition, firms have monopoly in production but competition in selling. Hence, the market has been termed as “Monopolistic Competition”.

3. Free entry and free exit : In this market there is no barrier on the entry of new firms and exit of old firms from the industry. Entry of new firms reduces the market share of the existing ones and exit of firms does the opposite. The consequences of free entry and exit of firms lead to intensive competition among the firms for both retaining and increasing their market share.

4. As each seller is a monopolist, his AR curve is downward sloping and naturally, MR < AR. Each seller is a price maker, not a price taker. This implies that every seller can influence the price. Evidently, he can sell more at a lower price or sell less at a higher price.

5. Selling Costs : Since each firm produces a differentiated product, each firm has to make some heavy expenditure on advertisement and other sales promotional schemes for his product. However, there is no such advertising costs under monopoly or under perfect competition.

In a perfectly competitive market we discuss the equilibrium of a firm and equilibrium of the industry. By industry we mean the whole set of those firms which produce a homogeneous product. However, under monopolistic competition the firms do not produce a homogeneous product. Hence the concept of an industry is not applicable here. Instead, the concept of "product group" has been used by Chamberlin. By product group we mean as a collection of firms producing almost similar goods but not identical goods.

Under monopolistic competition each firm is a monopolist seller. He is producing a single product. No one is producing exactly that commodity. For example, there are tooth pastes of various brands in the market. However, only one firm is producing Colgate tooth paste. Thus the producer of Colgae toothpaste is a monopolist and the equilibrium of this producer will be determined in the same manner as in a monopoly market.

6a.4 Proportional Demand Curve Vis-a-vis Perceived Demand Curve

We know that monopolistic competition is an amalgam of perfect competition and monopoly. Naturally, a monopolistically competitive firm does not face a horizontal demand curve. However, a competitive firm has experience of facing a horizontal demand curve since goods produced by all firms are homogeneous in nature. Product differentiation is one of the vital assumptions of monopolistic competition where goods are not perfect substitutes like those under perfect competition. Goods produced by a large number of monopolistic competitors are very closely related to each other. Each product is a very close substitute for the product of others.

It is because of this product differentiation that every firm enjoys some sort of monopoly power since each product is of unique nature. Therefore, the demand curve faced by a monopolistically competitive firm is negatively sloping.

For explaining his theory of monopolistic competition Professor Chamberlin has considered two types of demand curves : (a) proportional demand curve or actual sales curve and (b) perceived demand curve or anticipated demand curve. The first type of demand curve can be drawn by assuming that all the monopolistically competitive firms charge the same price while the latter demand curve is drawn on the assumption that the competitor sellers will not change the original price. In the following diagram, Fig 6.1 D_P is the proportional demand curve faced by a particular firm when all the sellers

charge the same price whereas DA is the anticipated demand curve facing the firm if all sellers maintain the original price.

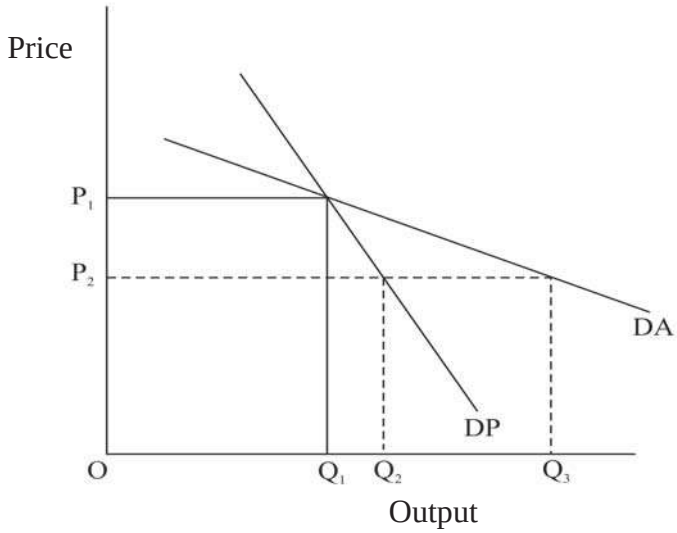


Fig. 6.1 Two types of demand curve

We start with a price OP_1 at which OQ_1 amount of output gets sold. Let us now assume that a typical firm wishes a cut in price from OP_1 and he will expect his sales to go up. This, according to Chamberlin, is due to the fact that all other firms will keep this price at OP_1. In the same way, if the particular seller wishes to raise the price of the product, he can expect that there will be a drastic fall in sales because of the fact that all other sellers will keep their price at OP_1. Therefore, the individual firm perceives a demand curve DA at OP_1 price since every seller expects that his action to go, will remain unnoticed by his rivals. In the diagram the DA curve is more elastic than DP because each firm, in this model, believes that no other firms will react to changes in its price.

Expecting elastic demand, each firm has an incentive to lower the price of its product for capturing a lower share of the market. That means, if a particular firm reduces price of its product from OP_1 to OP_2, it can expect its sales to increase to OQ_3. Since every firm expects that no other firms will cut price, the gain in the ultimate analysis will be smaller. If all firms reduce price to OP_2, the actual sales will be OQ_2, not OQ_3. Because of this reason, the D_P curve is more steep or less elastic than the DA curve.

The D_P curve shows the actual sales as it takes into account the effects of the actions of the competing sellers to the price changes by the firm. However, the DA curve shows anticipated changes in quantity sold when it contemplates a change in price. Under this form of market, each seller has the belief that his actions will go unnoticed. Hence every

firm ignores the reaction of rival sellers and, thus he behaves independently. All firms, acting independently, do actually sell less than what they do anticipate.

But in the present context we have not used anticipated or perceived demand curves to describe the equilibrium situation of a firm under monopolistic competition.

6a.5 Equilibrium situation of a Firm under Monopolistic Competition

We know that the basic condition for attaining maximum profit is to produce that level of output at which MR = MC and at the same time at which the MC curve cuts the MR curve from below. In this market product differentiation gives rise to a negatively sloped demand curve or AR curve. In this form of market the demand curve or the AR curve facing the firm has been drawn on the assmumption that all the rival firms do not change their prices. However, the demand curve is relativly more elastic. The corresponding MR curve lies below it.

6a.5.1 Short Run Equilibrium

The short run equilibrium of a firm under monopolistic competition has been illustrated in the following diagram fig. (6.2) where quantity of output has been measured horizontally whereas revenue, cost, price and profit have been measured vertically. In the diagram both the conditions for the attainment of maximum profit have been satisfied at point e which is the point of equilibrium.

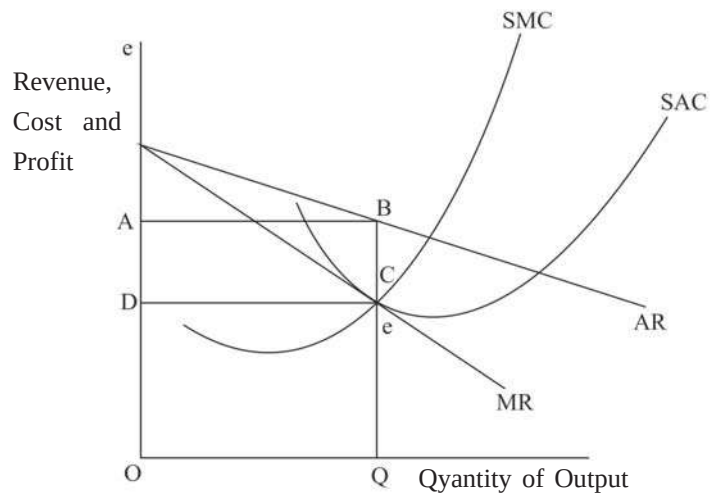


Fig. 6.2

From the diagram it is seen that the firm sells OQ units of output at price OA so that its total revenue becomes OA × OQ = ▭ OABQ. For producing OQ units of output the average cost is CQ so that total cost becomes OQ × CQ = ▭ ODCQ. Hence the amount of pure profit is ▭ OABQ – ▭ ODCQ = ▭ ABCD.

A firm, in the short run, may earn only normal profit if MC = MR < AR = AC happens. A loss may take place in the short run also if MC = MR < AR < AC takes place.

6a.5.2 Long Run Equilibrium

In the long run, the monopolistic competition comes closer to the perfect competition because the freedom of entry and exit allows firms to enjoy only normal profit. Whenever some firms earn pure profit in the long run some other firms may be attracted to join this product group thereby driving excess profit to zero. Experiencing losses for a long time, losing firms may be induced to leave the product group, thereby eliminating losses. Therefore, all firms in the long run are to be satisfied only with normal profit.

In the following diagram (Fig. 6.3) the long run equilibrium has been attained at point E where LMC equals MR.

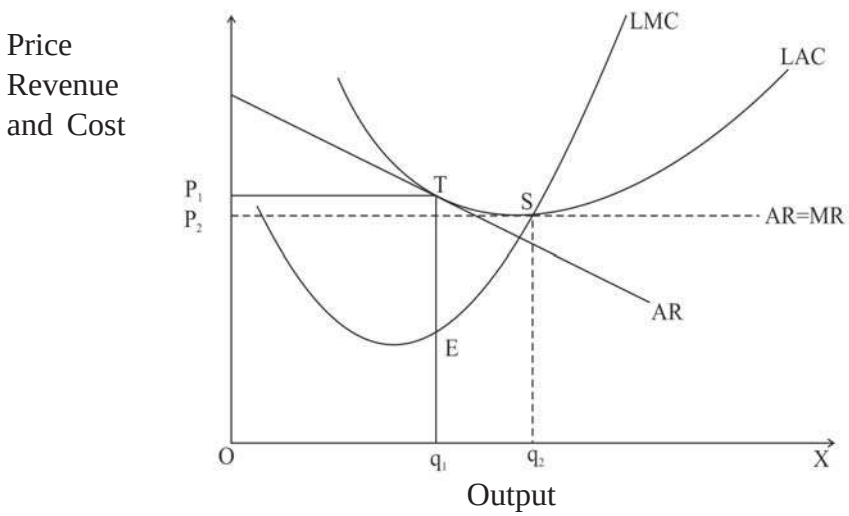


Fig. 6.3

The equilibrium output determined in this way is oq_1 at which AR equals AC. The firm will get normal profit by selling oq_1 units of output at price OP_1. We may note here that a monopolistically competitive firm always operates somewhere to the left of

the minimum point of its AC curve. It may be noted that as the demand curve is not perfectly elastic or as the demand curve is negatively sloping, the AR curve becomes tangent above the lowest point of the AC curve (i.e., point T). Thus each firm produces at a cost higher than the minimum one and attains only normal profit.

Under perfectly competitive market, the long run equilibrium is attained at that point at which MR = MC = AC = AR = Price. Since the AR curve is perfectly elastic, a tangency between AR and AC takes place only on the lowest point of the AC curve. In the above diagram fig. 6.3, the dotted AR = MR curve is the demand curve faced by a competitive firm. Equilibrium has been attained at point S where LMC = MR = AR = lowest point of LAC. The competitive output determined in this way is oq_2 and this will be sold at the price op_2. Therefore, we may conclude that the monopolistically competitive output (oq_1) is less than the perfectly competitive output (oq_2) and monopolistically competitive price (op_1) is higher than the competitive price (op_2). Naturally, the difference in output to the extent of q_1q_2 measures excess capacity or unused capacity faced by a monopolistically competitive firm. Production at a higher cost means wastage of resources or underutilization of resources.

Under perfect competition since production takes place at the lowest point of the AC curve, there does not occur any wastage of resources in this market. The society gets the larger amount of output and the consumers also get output at a lower price. Therefore, as perfect competiton maximizes social welfare, it is an ideal form of market. However, social welfare is not maximised under monopolistic competition since the society gets lower output compared to the perfectly competitive output and buyers purchase the differentiated products at a high price.

From what has been stated and verified above, it may be observed that “perfect competition is the ideal form of market”.

6a.6 Product Differentiation

We know that in the modern business goods are not identical or homogeneous in quality. We notice that goods are very closely related or slightly differentiated. Professor Chamberlin’s monopolistic competition is featured by the existence of commodities that have fairly good, though not perfect substitutes. This trend has put its shadow in modern markets. In our daily life we see various types of detergent, toilet soaps etc. Each variety is close to the other and that is why buyers do not mind much if they fail to get their chosen item, because very close alternatives appear before them.

Product differentiation is usually of two varieties : (a) real and (b) spurious. A real product differentiation is noticed when the inherent characteristics of the products (for

example, chemical composition of say, margo soap is different from that of Haman soap) are different. Again, product differentiation may be spurious or fancied. That means the product is basically the same but through advertisements or packaging differences get established.

The above-mentioned product differentiation makes the sellers enjoy some degree of monopoly power. Buyers get usually fascinated to buy a particular variety of a product. It is the brand loyalty that creates some sort of monopoly. Naturally, product differentiation leads to a negatively sloping demand curve. But the monopoly power is limited since rival firms' products are in the market and the customers may move to these products. That means, the monopoly sellers face some sort of competition. This implies that the demand curve faced by a monopolistically competitive seller will be really elastic.

Moreover, modern sellers very often engage themselves in price war for promoting their sales. So, Professor Chamberlin, besides stressing on price competition, had given stress on non-price competition. Product differentiation, in modern business, has justified the importance of selling expenses or selling costs. According to Chamberlin, product differentiation and selling costs are two basic forms of non-price competition. In modern days by spending on attractive advertisement expenditure many buyers may be induced to purchase the product even without price-cutting. This implies that selling expenses have a great influence on demand for a product. That means in Chamberlin's model, demand is influenced not only by the price of the product but also by the style of the product, the services associated with it and the selling activities of the firm. Just to attract more customers in their favour, sellers very often compete with each other that may have high influence on demand for the product. These demand influencing variables (not the price of the product) are the elements of non-price competition.

6a.7 Analysis of Selling Cost or Advertisement Cost in Monopolistic Competition

There is a vital difference between monopolistic competition and perfect competition. Under perfect competition since products of all firms are homogeneous or identical in quality, no individual firm has any incentive for making any sort of advertisement for its product. Gains from advertisement by an individual firm have spill over effects in the sense that the gains will get distributed among all firms in the industry. However, such is not the case for monopolistic competition as different sellers produce slightly different goods. It is the very product differentiation that enables a monopolistically competitive seller to exercise some type of monopoly power. Naturally, sales promotional activities are undertaken by sellers to persuade many buyers to buy their products. Sales promotion,

in this way, involves selling cost which is, obviously, to be distinguished from production cost. Naturally, under monopolistically competitive market, total cost of producing a product is the sum total of production cost and selling cost.

Professor Chamberlin was the first economist to introduce the concept of selling cost. Through various sales promotional activities and advertising every firm claims that his product is better than his rival's one. Naturally, selling costs are those costs which are incurred by a firm to influence or persuade buyers to buy its product in place of products of others. In the language of Professor Chamberlin, selling costs are those costs which are incurred by a monopolistically competitive firm in order to alter the position or shape of the demand curve for a product. The obvious purpose of selling cost is, therefore, to capture the saleable markets in order to increase total revenue and, ultimately, profit. As the buyers have imperfect knowledge about the firm, price, quality of the product, existence of rival sellers etc. and since there is the possibility of altering wants of buyers through advertisement and other sales promotional activities, sellers are tempted to incur selling costs. Naturally, selling costs increase and shift the demand curve to the right.

Let us now consider the effect of advertising on price-output decision of a monopolistically competitive seller. We are interested in finding out the optimum output, price and also selling cost. This has been demonstrated in the following diagram, Fig 6.4.

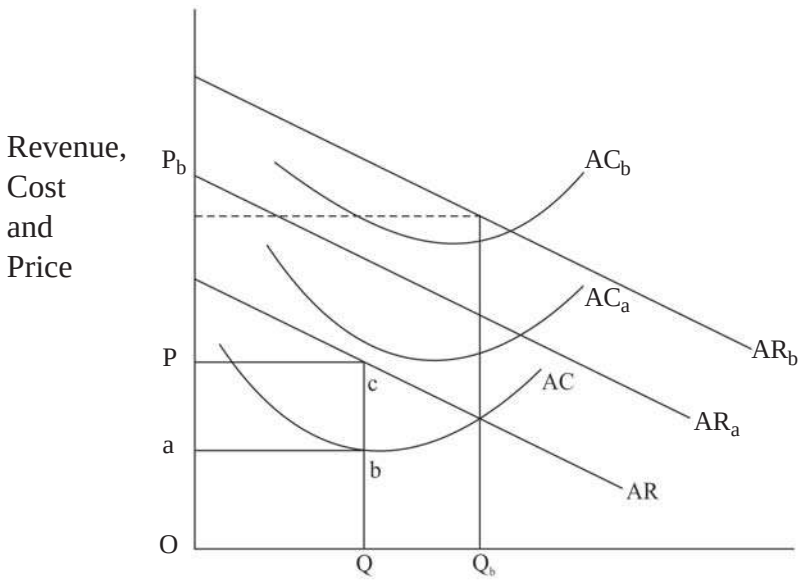


Fig. 6.4 Selling Cost and Equilibrium

We know that advertisement has got two roles to paly : (a) It leads to an increment in cost and (b) it leads to a shift in the demand curve in the upward direction. In the diagram AR is the average revenue of the firm before the introduction of any selling expenses. AC is the average cost of production and this does not take into account advertisement costs or selling costs. The profit maximizing output is determined at the point of intersection between MR and MC, which have not been shown in the figure.

6a.8 A Comparative study among Perfect Competition, Monopoly and Monopolistic Competition

Perfect competition and monopoly are two extreme forms of market. In reality, neither of these markets is found. In the real world many industries fall in between. One realistic market is monopolistically competitive market, named by Professor Chamberlin. It is so called because it has elements of both monopoly and perfect competition. But the fundamental aim of all sellers of all these market forms is the maximisation of profit which gets satisfied when MR equals MC and MC cuts MR from below.

The fundamental differences between perfect competition, monopoly and monopolistic competition may be pointed out in the following paragraphs :

(a) Structural Diffrences : Under the perfectly competitive market, there are innumerable number of firms each producing homonogeneous goods. However, each firm in the market is so small that it fails to exert any influence on either price or quantity of output. Naturally, the firms are merely price-takers. However, under monopolistic competition there are a large number of sellers, each selling slightly differrent products. Product differentiation enables a firm to exercise some power over price and output. That is, sellers in this model, behave as “price-makers”. A monopolist seller has full control over his price-output decision.

In both perfect competition and in monopolistic competition there is complete freedom of entry and exit of firms. This condition is true only during the long period. In the short period, however, entry or exit of new firmg is legally prohibited in monopoly.

(b) Behavioural Differences : A firm under perfect competition behaves as a price-taker. So the demand curve faced by it is a horizontal straight line. Under perfect competition since price is fixed, the AR curve coincides with the MR curve. A monopolist faces a negatively sloped demand curve and it can have perceptible

influence over price and output. Naturally, in this market the MR curve is negatively sloping and it lies below the AR curve. Under monopolistic competition this is also true. From this point of view the difference between monopoly and monopolistic competition is that the demand faced by a monopolistically competitive seller is relatively more elastic.

Under perfect competition price is fixed and so the firm has to undertake only output decisions. Besides, goods sold by competitive firms are perfect substitutes. Because of absolute product homogeneity, no firm reveals any eagerness to spend money on any kind of promotional activity.

Similar is the case with a monopolist who does not find any urgency to spend money on advertisement as there is no rival seller. However, a monopolistically competitive seller has to incur some kind of "selling costs" in order to give information about its product or rivals' products. In reality, in order to attract more and more customers in its favour, additional expenditure on selling cost is a mere necessity.

In every market sellers adopt the technique of independent price decision. However, all sellers of all markets form a basic principle which is that at the point of equilibrium there will be equality between MR and MC. Under perfect competition MR = AR, MC = MR = AR = P = price.

However, in both monopoly and monopolistic competition this behavioural rule is slightly modified to MC = MR < AR = price, simply because in both these two markets AR > MR.

Both a monopoly firm and a monopolistically competitive firm produce at that region of its demand curve where the coefficient of elasticity of demand is greater than one. However, under perfect competition the coefficient of elasticity of demand is infinite.

(C) Optimum Capacity and Sub-Optimal Capacity of Production : A competitive firm always produces at the minimum point of its AC curve. This amounts to saying that a firm utilises its plant optimally. As the AR curve is a horizontal straight line, a competitive firm will always produce at the lowest point of its AC curve. However, in both monopoly and monopolistic competition the demand curve is negatively sloping. Thus it is due to the nature of the demand curve that a firm fails to operate at the minimum point of its AC curve. It operates somewhere to the left of the lowest point of the AC curve. This implies that resources are not optimally utilised. Consequently, a higher price for the product is charged and a lower amount of output is produced. From

this point of view perfect competition is an ideal market where social welfare becomes maximum. However, this welfare gets reduced in both monopoly and monopolistic competition.

(D) Supply Curve : Under perfect competition the MC curve above the shut down point is the short run supply curve. However, under monopoly or monopolistic competition the supply curve remains indeterminate. This means that in these two forms of market the MC curve is not the supply curve.

6a.9 Sample Questions

1. What are the characteristics of monopolistic competition?
2. Distinguish between short run and long run equilibrium under monopolistic competition.
3. What is product differentiation? How does the producer make product differentiation?
4. What is the significance of selling cost under monopolistic competition? Explain.

Unit 6(B) ❐ Oligopoly

Structure

6a.1 Objectives

The basic objectives of this unit are to indicate :

i. Main features of oligopoly.

ii. Classification of oligopoly.

iii. Discuss equilibrium under different forms of oligopoly.

6b.2 Introduction

Many of the markets resemble monopolistic competition where sellers behave independently, that means, actions of one seller go unnoticed by his rival sellers. In contrast, many of the markets, in reality, are dominated by a few sellers where interdependence among themselves is the basic element. While taking price-output decision, each seller takes into account the actions of other sellers. Such a market form is popularly known as oligopoly. This term has been derived from 'Oligos' which means few and 'polis' which means sellers. This means that oligopoly is said to exist when there are few sellers. However, many economists opine that when the number of sellers varies between 2 and 20, the market is said to be an oligopolistic market. The extreme form of this oligopoly is 'duopoly' when the number of sellers is exactly two. However, there is no such precise 'number limit'. Precisely, it may be stated that the key issue is not numbers but rather interdependence among sellers. Thus it is said that oligopoly is said to prevail when there

are few firms or sellers in the market producing and selling a homogeneous or differentiated product. That is why it is often referred to as "competition among the few". Under oligopoly there is no fixed number of sellers. However, we may say that the number of sellers should be such that there will be interdependence among the sellers regarding their behaviour. If the product is homogeneous, we call it pure oligopoly or oligopoly without product differentiation. However, if products are differentiated, we call it imperfect oligopoly or oligopoly without product differentiation.

According to Professor Baumol, "An oligopoly is a market dominated by a few sellers at least several of which are large enough relative to the total market to be able to influence the market prices".

6b.3 Main Features of Oligopoly

There are some special characteristics of oligopoly. These are as follows :

(1) **Few sellers :** Oligopoly is a market structure in which a small number of rival firms dominate the industry. It is under oligopoly that rivalry among firms takes its most direct and active forms.

(2) **Product :** Products sold by oligopolistic firms can either be homogeneous or differentiated. For example, industries producing bread, cement, steel, petrol, cooking gas, chemicals, aluminium and sugar are industries characterised by homogeneous oligopoly. If firms of an oligopoly industry sell differentiated products, it is called differentiated or heterogeneous oligopoly. Automobiles, television sets, soaps and detergents, refrigerators, soft drinks, computers, cigarettes etc. are some examples of heterogeneous oligopoly.

(3) **Lack of Uniformity :** Lack of uniformity in the size of firms is a unique feature of oligopoly. Some firms may be small while others may be large.

(4) **Importance of advertising and selling costs :** Under oligopoly each seller has to inform his buyers about the quality of his product. Consequently, he has to incur advertising and selling costs.

(5) **Competition :** Competition implying struggle of rivals against rivals is a very unique feature in the oligopolistic market.

(6) **Uncertainty :** Uncertainty in the behaviour of firms is a special feature of oligopoly.

(7) **Elements of monopoly :** Under oligopoly there are only a few firms in the market. Ths existence of product differentiation creates "brand loyalty" on the part of the consumers,

which is the basic source of monopoly power. Again, through collusion, the existing firms can raise price and earn some monopoly income.

6b.3.1 Classification of Oligopoly

An oligopoly market is beset with the problem of price determination as the actions and reactions of rival firms vary from one industry to another. We shall classify oligopoly market on the basis of (a) specific assumptions about the actions and reactions between firms and (b) nature of the product.

On the basis of the nature of the product the oligopoly model may be either homogeneous or differentiated. In case of the homogeneous oligopoly or pure oligopoly the products of different firms are homogeneous. If oligopoly sellers sell homogeneous goods, price differences between products will be, to some extent, insignificant. That means, greater the homogeneity of the product, greater will be interdependence among sellers. This implies that whenever a seller changes the price of his product, the sales of the rival sellers will be affected. Naturally, the rival sellers will change their pricing policy.

In case of differentiated oligopoly sellers sell slightly different products. The change in price by one seller will now have less direct effect upon his rival sellers. Hence, under differentiated oligopoly the assumption of interdependence among few sellers becomes less significant. Insignificant rival consciousness greatly complicates the price-output determination model of oligopoly. However, most of our oligopoly models belong to this category.

Again, oligopoly market has been characterised by ‘Competition’ and ‘collusion’. Interdependence of a firm may encourage firms to compete with their rivals or may cause sellers to collude with each other. The former oligopoly market is known as non-collusive oligopoly while the latter is known as collusive oligopoly.

Non-collusive Oligopoly : In case of non-collusive oligopoly market, even though the firms are interdependent in the market, they go on behaving independently. Under this category of oligopoly the be haviour of a seller will depend upon how he thinks his rivals will react to his decision making. The seller just makes a guess about the reaction of the competing firms, selling homogeneous goods. Therefore, an autonomous or conjectural behaviour is studied by firms of rival firms while setting the price of his product, the firm believes that rival firms will not make any counter move when he changes price of his product. Such behaviour is called autonomous behaviour. However, for oligopoly market this kind of autonomous behaviour is self-contradictory. So the classical non-collusive models of oligopoly, namely the Cournot model, Bertrand’s

model and Edge Worth's model fail to determine equilibrium of a firm and an industry. Assuming the absence of collusion some economists have been successful in explaining equilibrium of a firm. Of these non-collusive oligopoly models the followings are worth mentioning : (a) Cournot's duopoly model, (ii) Kinked demand curve model of P. Sweezy and (iii) Stackleberg's leadership model.

6b.3.2 Collusive Oligopoly

In case the firms collude we get the case of collusive oligopoly. Very often non-collusive oligopoly may lead to a cut-throat competition among sellers, leading ultimately towards monopoly business. Sellers having influence may retaliate. Just to prevent this competitive price cutting or retaliation sellers very often make a collusive agreement which may be open or tacit. Formation of cartel is one form of open collusive agreement.

6b.4 Cournot's Duopoly Model

In Economics a market with two sellers and a large number of buyers is called a duopoly market. Duopoly is a special case of oligopoly where there are a few sellers. This model of duopoly is based on the following assumptions :

(i) There are only two interdependent sellers selling a homogeneous commodity.

(ii) There are a large number of buyers in the market.

(iii) Both the duopolists have identical cost curves. It is assumed that each duopolist has zero cost of production.

(iv) Each duopolist wants to maximize his total profit in each period.

(v) Each duopolist makes an output plan during a period and this can not be revised in that period.

(vi) No duopolist sets the price. However, each accepts the price of his product at which total planned output can be sold.

(vii) Each duopolist is fully aware of the mutual interdependence between their output plans. However, each is quite ignorant of the direction and magnitudes of the revision in his rival's plan that would be induced by any given change in his own. Each duopolist assumes that the output level of his rival is constant at a certain level and adjusts his own output in order to maximise his total profit.

Let q_i be the output level of the i-th seller, i = 1, 2. Hence total output is $q = q_1 + q_2$. We assume that the demand curve faced by each duopolist is a straight line,

given by $p = a + bq$, p being price, a and b are constants, $a > 0$, $b < 0$. Total profit of the first duopolist is given by

$$\pi = pq_1 = (a + bq)q_1 = [a + b(q_1 + q_2)]q_1$$
$$= aq_1 + bq_1^2 + bq_1q_2 \ldots\ldots(A)$$

That is, $\pi_1 = \pi_1(q_1, q_2)$.....(B)

We can get various combinations of q_1 and q_2 from which a fixed level of profit of the first duopolist can be obtained. Naturally, the locus of all such combinations of q_1 and q_2 will be an iso-profit curve or a profit indifference curve of the first seller. For each level of π_1 we get one such profit indifference curve of the first seller.

As per our assumption the objective of the first duopolist is to maximize π_1 for variations in q_1, assuming q_2 to be constant. To do that we are to set $\frac{d\pi_1}{dq_1} = 0$.

That is, $a + 2bq_1 + bq_2 = 0$....(C)

That is, $bq_2 = -a - 2bq_1$ whence $q_2 = -2q_1 - \frac{a}{b}$.

This is a straight line having slope – 2 and

intercept $-\frac{a}{b}$ (positive, assuming $a > 0$ and $b < 0$).

In the following diagram R_1R_1 is a straight line reaction function of the first seller. This function gives us those combinations of q_1 and q_2 for which the profit of the first duopolist will be maximum. Naturally, all points on R_1R_1 fulfil condition (C).

In the same way the profit level of the second duopolist is given by

$\pi_2 = pq_2 = [a + b(q_1 + q_2)]q_2 = aq_2 + bq_2^2 + bq_1q_2$.

That is, $\pi_2 = \pi_2 (q_1, q_2)$.

We can get the profit indifference curve of the second duopolist as

$\pi_2 = aq_2 + bq_2^2 + bq_1q_2$.....(D)

The objective of the second duopolist is to maximize π_2 for variations in q_2 assuming q_1 to be constant.

That is, $\frac{d\pi_2}{dq_2} = a + 2bq_2 + bq_1 = 0$ .....(E)

That is, $2bq_2 = -a - bq_1$,

whence $q_2 = -\frac{1}{2}q_1 - \frac{a}{2b}$.

This is also the equation of a straight line having slope $-\frac{1}{2}$ and an intercept of $-\frac{a}{2b}$. This straight line is R_2R_2 in the diagram Fig (6.5) and it is the reaction function of the second duopolist. This line will be downward sloping with a positive intercept. We see from the diagram that the intercept of R_2R_2 is less than that of R_1R_1. Also, the absolute value of the slope of R_2R_2 is less than that of R_1R_1. Thus the reaction function of the second duopolist is flatter than that of the first one. The first duopolist remaining on R_1R_1 intends to maximise his profit and the second one remaining on R_2R_2 will try to maximise his profit. F is the point of intersection of both the reaction curves and it is here that both the sellers maximise their profits. At this equilibrium point the first duopolist's output level is oq^*_1 while that of the second one is oq^*_2, which are, obviously, the equilibrium output levels. Both the conditions (C) and (E) get fulfilled at the intersecting point F. Thus there are two equations (C) and (E) and also there are two unknowns. By solving the equations we can determine the values of q_1 and q_2.

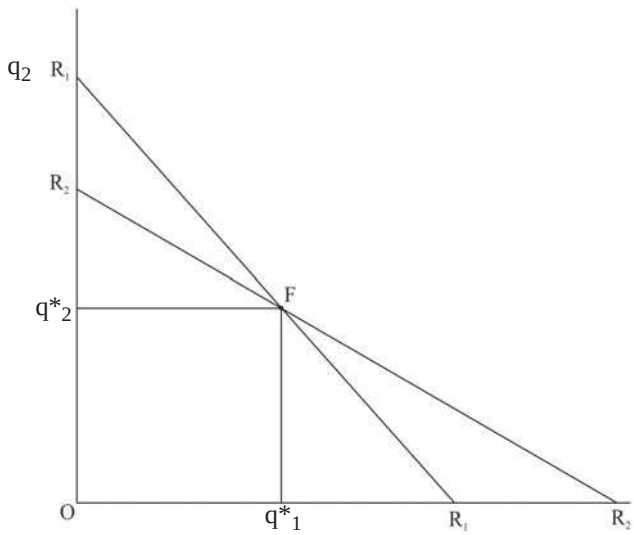


Fig. 6.5

We have already seen that

$a + 2bq_1 + bq_2 = 0$ and $a + 2bq_2 + bq_1 = 0$.

Hence, adding both sides we get $2a + 2b(q_1 + q_2) + b(q_1 + q_2) = 0$.

That is, $2a + 2bq + bq = 0$ $\quad \because q_1 + q_2 = q$.

That is, $2a + 3bq = 0$ whence $q = -\frac{2}{3}.\frac{a}{b}$.

If $q = -\frac{2}{3}.\frac{a}{b}$, then $p = a + bq = a - \frac{2}{3}.\frac{a}{b}.b = a - \frac{2a}{3} = \frac{a}{3}$.

Thus price in the duopoly market will be $\frac{a}{3}$.

Had there been perfectly competitive market having demand curve p = a + bq (assuming zero cost), the equilibrium will be attained where price equals marginal cost, i.e., price = 0, that is, $p = a + bq = 0$ or $q = -\frac{a}{b}$.

Now, assume that there was monopoly with zero costs and with the same demand function p = a + bq, equilibrium will be achieved when MR = 0, i.e., $a + 2bq = 0$ or $bq = \frac{-a}{2}$ or $q = \frac{-a}{2b}$.

Therefore, we see that the monopoly output is $-\frac{a}{2b}$. That means, we get the result that with zero costs and straight line demand curve the monopoly output is half of the competitive output and duopoly output is two-thirds of the competitive output.

It can be proved that equilibrium in the cournot model of duopoly is actually stable.

6b.5 Kinked Demand Curve Model of Oligopoly

The Kinked demand curve model of oligopoly was originally developed by Paul M. Sweezy in 1939 and later by Professors Hall and Hitch. However, this is not a theory of equilibrium price-output determination under oligopoly. It has been found, empirically, that price under oligopoly, if determined once, will remain rigid or sticky at that very level. This model of oligopoly seeks to explain this price rigidity in terms of a "Kink" on the demand curve. According to this model of oligopoly, the demand curve that any individual seller faces, has a Kink at the prevailing price-quantity combination. Due to the existence of such a "Kink" on the demand curve the marginal revenue curve becomes discontinuous. It is this very discontinuity in the MR curve that leads to price rigidity under oligopoly. This model of oligopoly does not deal with price and output determination. Rather, it seeks to establish that once a price-quantity combination is determined, an oligopoly firm does not find it profitable to change its price even when there is a considerable change in the cost of production and also a change in demand for the product.

The "Kink" in the demand curve stems from the asymmetric behavioural pattern of the sellers. According to Professor Sweezy each duopolist has the belief that if he lowers the price of his product, his competitors will not follow him. But, in reality, his rivals must react by reducing their prices. Again, he believes that if he raises the price above the prevailing price, his competitors will also raise their prices. However, in reality, the rivals will not follow his price-rise. It is this very particular assumption that gives rise to a "Kink" in the demand curve under oligopoly.

We assume that the prevailing price of an oligopoly product in the market is AE or OP in the following figure. If one particular seller increases the price above OP, the rival sellers will go on keeping the prices of their products at OP. In consequence of high price charged by the firm, the buyers will shift to products supplied by other sellers who have kept their prices at the old market rate. Naturally, sales of the first seller will sharply drop. It is because of this reason that the demand curve in the zone (dE) is relatively elastic. On the other hand, if a seller reduces the price of his product below AE, others will follow him so that demands for their products do not diminish.

Naturally, the demand curve in this region (ED) is relatively inelastic. It is this behavioural pattern that explains why in the oligopoly market prices are inflexible even when costs and demand change. The resulting demand curve is, therefore, dED shown in diagram 6.6.

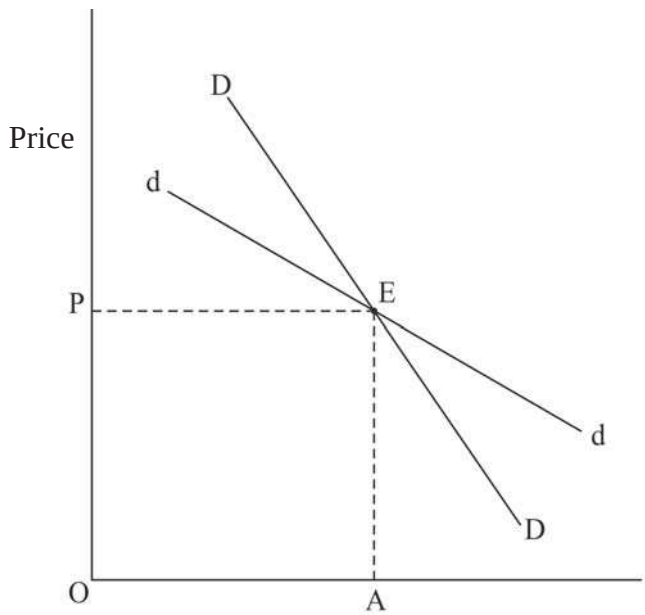


Fig. 6.6 Quantity

The ‘‘Kink’’ in the demand curve at point E results in a discontinuous MR curve. The marginal revenue (MR) curve has two segments at the level of output less than OA, the MR curve (that is dG) will correspond to dE portion of the AR curve and for output greater than OA, HMR the MR curve (that is, HMR) will correspond to the demand curve ED. Therefore, discontinuity in the MR curve has occurred between points G and H. This means that between these two points, the MR curve is vertical. The length of this discontinuity depends upon the relative elasticity at E on dE and the relative inelasticity of the demand curve at E on ED. We may mention here that dE is relatively elastic while ED is relatively inelastic.

Proof : We know that MR = price $\left(1-\frac{1}{e}\right)$.

Let price be Rs. 10 and e be 2 and 3. Hence, $MR_1 = 10\left(1-\frac{1}{2}\right) = 5$ and

$MR_2 = 10\left(1-\frac{1}{3}\right) = 6.7$.

Hence the gap between MR_1 and MR_2 is 1.7.

Further, let $e_3 = 5$ so that $MR_3 = 10\left(1-\frac{1}{5}\right) = 8$.

Hence the gap between MR_1 and MR_3 will be 3.

This proves that the larger the elasticity at E on dE and the smaller the elasticity at E on ED the greater is the length of the discontinuous zone on the MR curve. In diagram 6.7 the length of the discontinuity is GH.

The marginal rule tells us that both the conditions for profit maximisation, that is, (i) MR = MC and (ii) the slope of the MR curve is less than that of the MC curve are satisfied in the discontinuous range of the MR curve. Hence, the equilibrium level of price and output are determined as OP_o and OQ respectively.

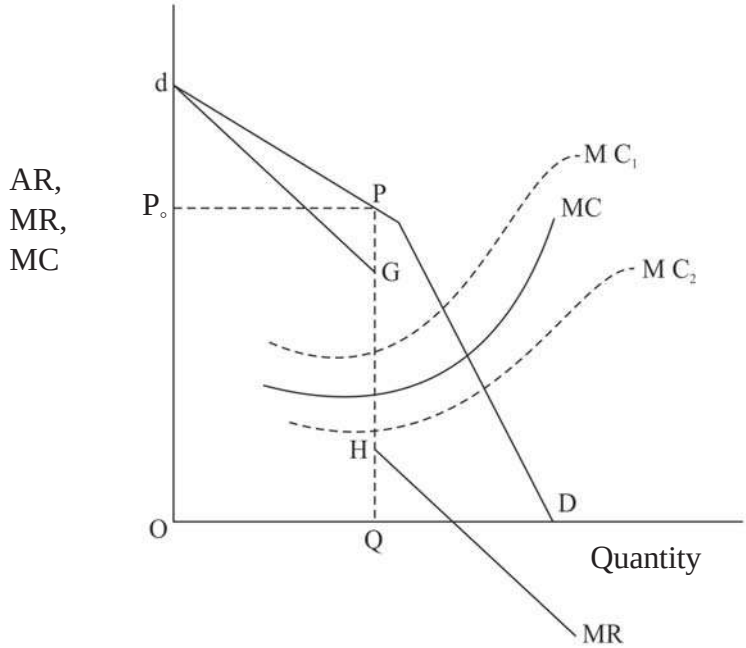


Fig. 6.7

We now suppose that the marginal cost curve of the firm passes through the discontinuous range of the MR curve. Naturally, the equilibrium of the oligopolistic firm is given by the Kinked point. The equilibrium price is OP_0 and quantity is OQ. For any level of output less than OQ, MR > MC. Hence the oligopolistic firm will raise the levels of output. Again, for any amount of output greater than OQ, MR < MC.

Naturally, the oligopolistic firm will reduce the level of output. Thus, total amount of output gets maximized at the Kink though the profit - maximizing condition (MR = MC) is not fulfilled. The discontinuous portion (that is, GH) of the MR curve implies that there is a range within which costs or demand may change without affecting the equilibrium price (i.e., OP_0) and output (that is, OQ).

For instance, if the cost condition of the oligopolist rises, the MC curve shifts to MC_1, though the equilibrium price remains unchanged at OP_0. Again, if the total cost falls so that the MC curve shifts to MC_2 but it still passes through the discontinuous portion of the MR curve, then also price remains unchanged at OP_0. Hence, for the existence of a discontinuous portion of the MR curve, price is rigid or sticky under oligopoly.

The main difficulty of the Kinked demand curve theory is that it does not explain

the price-output determination. It does not define the level at which price will be set to maximise profits.

The Kinked demand curve model of oligopoly can explain the rigidity of prices in a situation of changing costs and of strong rivalry. However, it fails to explain the level of the price at which the Kink will take place. That means, the Kinked demand curve model fails to explain the height of the kink. Thus, it is not a theory of pricing, but merely a tool for explaining why the price, once determined in some way or another, will tend to remain fixed there.

Sweezy had introduced the Kinked demand curve as an operative tool for the determination of equilibrium in oligopolistic markets. Later, Hall and Hitch used this model not as a tool of analysis for the determination of output and price in markets forming oligopoly but to explain why the price, once determined by virtue of average cost pricing principle, will remain sticky. Hall and Hitch had used the Kinked demand curve for explaining the stickiness of prices in oligopolistic markets and not as a tool for the determination of price.

6b.6 Sample Questions

A. Objective-type Questions :

(a) The monopolistically competitive market is associated with : (i) P. A. Samueslon, (ii) J. R. Hicks, (iii) J. M. Keynes, (iv) Chamberlin.

Ans. (iv)

(b) What does a monopolistic competition have in common with monopoly? (i) A large number of firms, (ii) A downward sloping demand curve, (iii) The ability to collect with respect to price, (iv) Mutual inter-dependence.

Ans. (ii)

(c) The key feature of oligopoly is : (i) Excess Capacity, (ii) Product differentiation, (iii) High profitability, (iv) Interdependence of firms.

Ans. (iv)

(d) Which one is true for a monopoly? (i) The product cannot be produced by small firms, (ii) There can be close substitutes for the product, (iii) There is a uniproduct with no close substitutes, (iv) Products are high-priced.

Ans. (iii)

(e) Under which of the following forms of market structure does a firm has no control over the price of its products? (i) Monopoly, (ii) Monopolistic competition, (iii) Oligopoly and (iv) Perfect competition.

Ans. (iv)

(f) To which of the market situations does the market for cars conform? (i) Monopoly, (ii) Perfect competition, (iii) Monopolistic competition and (iv) Oligopoly.

Ans. (iv)

(g) The length of the discontinuous zone of the MR curve corresponding to the demand curve under the Kinked demand curve model of oligopoly depends on (i) Elasticity of demand, (ii) Cost of production, (iii) Revenue and (iv) Price.

Ans. (i).

B. Short-type Questions :

1. Explain the features of a monopolistically competitive market.
2. Explain your understanding of "selling cost".
3. Make a clear distinction between profitable demand curve and perceived demand curve.
4. Write a brief note on "product differentiation". In which form of market it is found?
5. Explain the basic features of oligopoly.
6. Make a brief classification of oligopoly.
7. Explain the implication of the concept of "Kink" at the kinked demand curve model of oligopoly.
8. Whose name is associated with the Kinked demand curve model of oligopoly? What is the implication of the existence of the Kink?

C. Essay-type Questions :

1. What are the characteristics of a monopolistically competitive market? Why do firms under such a market have excess capacity even in the long run?

2. Mention the main features of monopolistic competition. Discuss the conditions of equilibrium of a firm under monopolistic competition in both the short run and the long run.
3. "Monopolistic competition is more like perfect competition than like monopoly". Do you agree with this view? If so, substantiate arguments in favour of your answer.
4. What are selling costs? Why are they peculiar to monopolistic competition?
5. Why does a Kink arise at the going price in the market demand curve under oligopoly? Explain the implications of this Kink.
6. What is an oligopoly market? Explain the main features of oligopoly. Make a brief classification of oligopoly.
7. Explain, in detail, the Cournot's model of duopoly.
8. What is the Kinked demand curve hypothesis? What is the implication of this Kink? Explain, graphically, how price and quantity are determined in this model of oligopoly.

Unit 7 ❐ Theories of the firm

Structure

7.0. Objective

The primary objectives of this unit are to :

i. Bring out alternative theories of the firm.

ii. Discuss Baumol's sales maximisation principle.

iii. Discuss Williamson's Expense Preference Model.

iv. Bring out Behavioural Theories.

7.0. Introduction

In this unit, we will study a number of topics which are outside the purview of the mainstream neo-classical analysis of the firm. Firstly, we will study some alternative theories of the firm. Next, we will study the case of a multi-product firm and how it settles prices of its products. Then, we will go into some special pricing techniques, like peak-load.

7.1 Alternative Theories of the Firm

Neo-classical theories characterised the firm as a 'black-box', which transforms inputs into outputs to earn the maximum profit. This is convenient for tackling issues like pricing, output, entry etc. as it ignores the details about the internal structure of the firm, along with its contractual and organisatioral basis. It also suggests that the optimal size of the firm is simultaneously determined with the number of firms in the industry. In case of a free entry situation, the size and number are determined by the extent of the demand for the product and the structure of technology. The economies of scale and scope are limited to the extent. If the question is asked: what limits the size of the firm? The focus on technology, demand, externalities suggest no obvious reason why a firm cannot monopolise the market. If there are diseconomies of scale associated with a single plant, the firm will surely go for multi-plant operation. So, it is necessary to spin the 'black belt'. In fact, in the last seventy years, the neo-classical methodology has tried to open the box. The early works on managerial discretion focussed on the organisational structure and the goal of the firm and its consequences. They recognised the separation of ownership from control in the corporate enterprise. This section gives a brief review of the earlier work on discretion and related theories.

7.1.1. Baumol's Sales Maximisation Model

Baumol (1958) argued that the goals pursued by managers (e.g., salaries, status, power, prestige, security etc.) are correlated with sales revenue. He suggested that the main objective of the firms is sales maximisation. For maximising sales revenue, the firm will set a lower price to produce a greater output than that would be under profit maximisation. The underperformance of a firm is visible by its decreasing level of profit and the firm has to take some action. If a firm earns below a minimum accepted level profit (MALP), then the management is at a risk from either shareholder action or take-over raiders. Thus the pursuit of sales is tempered by the need to earn enough profit.

This model assumes that MALP has defined a value, above which there would be no chance of intervention.

In the absence of any profit constriant, the firm maximises revenue by choosing price P_R and output $q_R$such that MR = 0, as per figure 7.1. This maximises revenue but pays no attention to the resulting profit in this figure, the profit is π_R. If MALP is $\bar{\pi}$, then the firm's objective is to choose P and q such as to maximise R subject to profit $\pi \geq \bar{\pi}$.

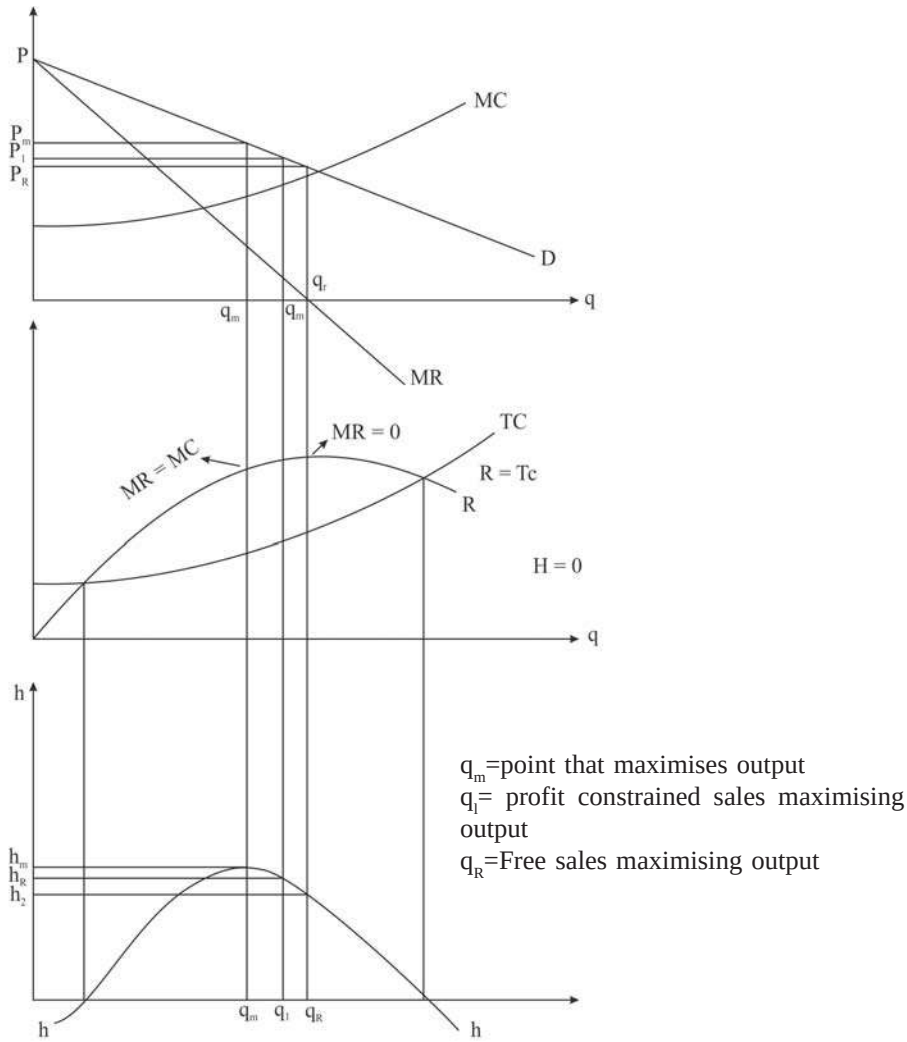
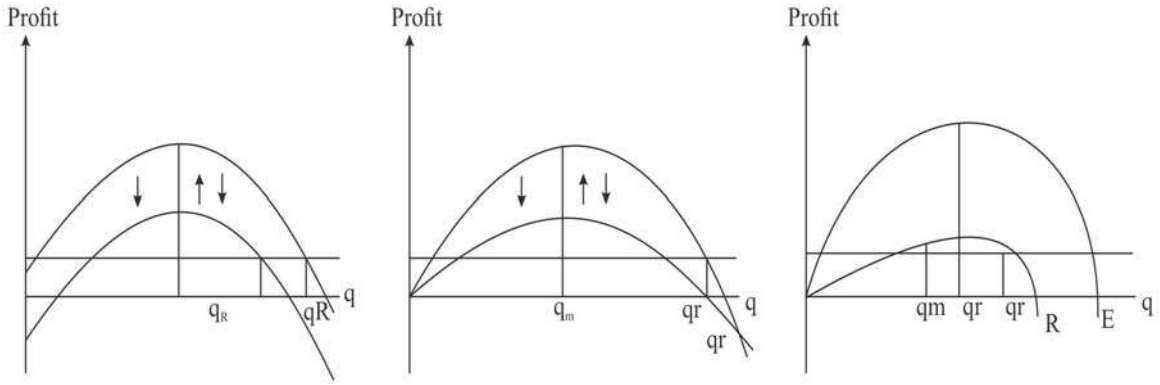


Figure 7.1 : Baumol's Sales Maximisation

If MALP is at the level $\bar{\pi}_2$, then the solution is unconstrained, i.e., revenue is maximised by setting MR = 0 where $\pi r \geq \bar{\pi}_2$ and the profit. So long as the MALP is less than the maximum attainable level of profit, the firm sets a price below the profit maximising level.

The sales revenue maximising firm (Baumol) and the profit maximising firm behave differently, when the environmental variables change. In figure 7.2, we will show some simple cases. In panel A, we have the consequence of an increase in fixed cost (or

lump sum tax); here the profit curve would fall everywhere by the amount of increase. Panel B shows the effect of an increase in corporate tax rate. In both of these cases, the Baumol firm reduces output and increases price. Lane C, on the other hand, shows the imposition of a per unit sales tax (or VAT). This leads to a fall in the would-be-net-profit function which is proportional to the level of output. The impact of this for both a profit maximising firm and a Baumol firm would be change in output, of course, and the Baumol firm would be more sensitive.



a) Effect of increse in fixed cost of increase in lumpsum tax

(b) Effect of corporation tax increase

(c) Effect of sales tax increase

Figure 7.2 : Some Comparative states of Baumol's firm

The firm would be cost-efficient if the MALP constraint binds, otherwise not. This is, again, not at all realistic, the managers will always have some preference for more profit rather than loss. From the welfare perspective, the Baumol firm is likely to be cost-efficient, the set price would be lower than that of the profit-maximising firm.

One of the weaknesses of the model is that it implies that the management prefers more sales to less and does not care about profit, till it falls below the MALP. At the point, profit becomes all important. In reality, of course, it is expected that the management would be concerned about both profit and revenue.

So, this model can be modified by assuming that the managers maximise a utility function of the form U(π, R). Here utility is increasing for both and R and the indifference curves have usual shape. In Figure 7.3, the locus of attainable combinations of π and R have been derived from Figure 7.1.

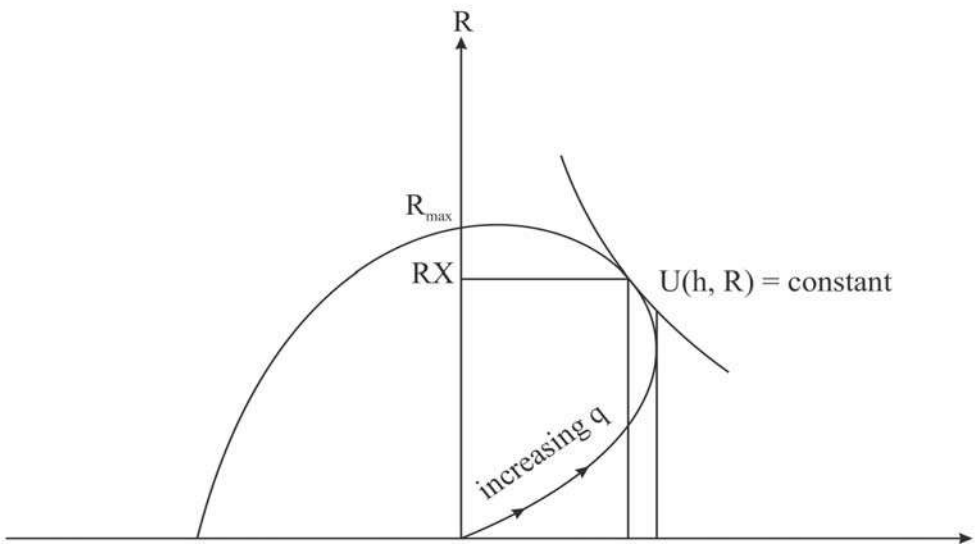


Figure 7.3 : Alternative Version of the Baumol Model

When output is zero, both profit and revenue are also zero. As output increases, both of them increase, until π reaches a maximum (R would be still increasing). As output further increases, π starts falling whilst R increases till it reaches a maximum value. Further increases in output lead to a fall in both π and R. The optimum output is characterised by the usual tangency solution, which is less than maximum revenue and also less than maximum profit.

7.1.1. Williamson's Expense Preference Model

Williamson (1964) suggested thal managers show preference for expense; they

derive utility from expenditure on staff, managerial emoluments and discretionary profits. Higher profits reduce the risk of embarrassment in case of adverse environmental fluctuations or take-over or profit-related payments. Excess may be partly attractive because not only it increases the dimensions of the managerial pyramid but also the pressure or manager concerned.

Here we will present a compact version of the model by focusing simply on the profit and expenses S (S is the composite expenditure on perquisites, on-line staffing etc.) The Williamson constraint of MALP necessary for survival and constraint on minimum expenditure necessary on stuff are ignored.

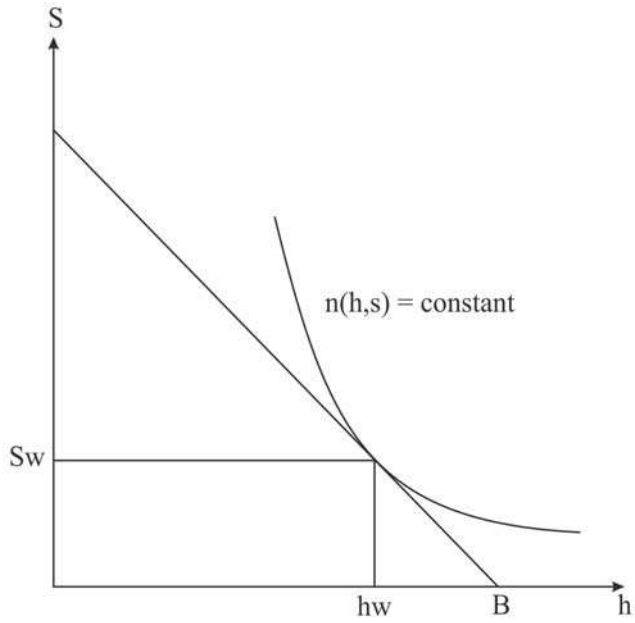


Figure 7.4 : Williamson's Expense Preference Model

The managerial utility function U(π, S) gives rise to a usual form of indifference curves, as depicted in Figure 7.4. If the gross profit is denoted by π_g = R(q) – C(q), then there must be a profit-maximising price and output for the firm to attain a maximum level of gross profit. This is denoted by π_m as shown in Figure 7.4. The actual net profit depends on the level of stalf and related expenses : $\pi = \pi_m - S$.

As π_m is a constant. the graph of the attainable combinations of (π, S) is the straight line AB in Figure 7.4. The optimal choice (π_w, S_w) is given by the point of tangency with the highest attainable managerial indifference curve. This implies cost inefficiency as expenditures on staff are positive but 'unnecessary',

The original Williamson model is richer than this simplistic version, but the message remains the same. Managers have utility by spending resources on themselves and so, such firms will not be cost-efficient.

The welfare implication of such a firm is more difficult to assess than in the Baumol case. The firm is maximising profit as it is cost-inefficient, for those who receive it. If S consists solely of staff perks and payments, this would imply a simple transfer of money from shareholders to managers; the consequence would be welfare neutral. But if S includes employment of excess staff, this involves loss of welfare.

7.2. Marris's Growth Maximisaiion Hypothesis

Marris (1963, 1964) developed a dynamic model of firm. He proposed that managers are concerned with growth maximisation, with expanding product demand and variety over time. Empirical evidences are there to show that managers prefer internal to external promotions as thc latter implied adopting new working practices and less job security. Marris's model implies a growing firm gives the managers the opportunity for internal promotions, career development and job security. It also examines the implication of the means of financing the growth.

We are not going into a detailed exposition of the model as it is quite complex and a number of *ad hoc* assumptions (specially with respect to pricing and demand for the product, alignment of manager's and owners' interest etc.) are also criticised. Its main merit lies in its focus on the dynamic factors.

7.2.1. Behavioural Theories

The behavioural theories (Simon 1957, Cyert and March 1963) viewed the firm as a coalition or participants or groups or participants (e.g. shareholders, creditors, suppliers, consumers and various categories of managers and workers). who work for the interest of the firm in a variety of payment The goals of the firms result in implicit bargain between various interested parties, influence of the parties on the overall direction of the firm is dependent on their bargaining power.

All these elements are consistent with the game-theoretic treatment of the firm, where the individuals are seen self-interested utility-maximisers. But the behavioural theories do not agree to the neo-classical paradigm of *homo economicus*. As for example, the employees may have aspiration levels rather than a reservation wage. Aspiration levels can increase upward with increase in experience but there will be no active job seeking until current wage is suifficiently below the aspiration level. The same idea applies to consumer behaviour, brand loyalty to mismatch between price qualily.

The difference between the payments required to keep the persons in their posts and the total possible revenues of the organisation is called organisational slack. In a stable environment, it may be thought that wages should converge on aspiration levels and competitive pressure may lead firms zero (or close to it) organisational slack. As the environments are not stationary (business) cycles, technological progress), the firms must strive to maintain themselves an ever-moving best-practice frontier.

The main features of the behavioural theory are as follows :

(i) it argues thal only individuals (not firms) have goals;

(ii) it argues that agents have bounded rationalily and satisfies rather optimise);

(iii) It recognises that information is costly;

(iv) It views firms as coalition of' agents.

Many of these ideas are incorporated in the 'optimising perspective' of neo-classical economics.

7.3. Multi-Product Firms and Multi-Product Pricing

Most of the text books in Economics focus the single-product firm, whereas, in practice, most firms produce more than one product The reasons for this can be traced to either technological or organisational advantages of the firm. Conglomerate firms may produce products which are not very related to one another and whose production processes may be largely independent. In such cases, where there is independence of the products both on the demand and the production side, the problem of pricing and marketing of the products is almost the same as the single-product firm. However, usually firms produce a product line where products are usually substitutes or complements. Apart from demand inter-dependencies, there are cost dependencies too. These arise when products utilise common production process, common inputs, common inventory etc. These interdependencies imply that theoretically, it is no longer possible to determine the profit-maximising price of each product in isolation. Pricing of the whole range of a firm's product has to be undertaken simultaneously in a system-wise solution.

7.3.1. Multi-product Firms

Multiple products are likely to be produced within a single organisational structure, if it is more profitable to do so. The reasons can, loosely, be categorised as demand or cost-related.

(a) Demand-related reasons :

The argument is that drawing products under umbrella of a single matter may help to facilitate price discrimination, to create barriers to entry to leverage monopoly power.

(i) Price Discrimination

A firm with monopoly power could improve its profitability by designing non-linear schedules by which it sells its products. Typically, it involves quantity discounts. Firms usually apply this technique when they bundle products. A car manufaturer will typically

offer a model with wide range of specifications, such that the higher specification model is cheaper than the cost of buying a lower-specification model and buying separately the additional specificiations. If we think of a situation where different firms provide different items having monopoly power over its item, then if they can be brought under the single umbrella, there may be no savings in cost, only in organisational benefits, pricing of different commodity bundles is likely to be better co-ordinated and controlled within a single firm.

(ii) Barries of Entry and Monopoly Leveraging

Monopoly leveraging of exploit monopoly power in one market either to prevent others being able to enter, or to achieve a dominant positiion in another market. Microsoft forced PC manufacturers to install Internet Explorer (its internet browser softwere) as a condition of obtaining licenses for windows 95 operating system. This internet browser is not superior to the others in the market, but as Microsoft has a stranglehold over the market for operating systems, it is able to leverage a significant market share for that product. The dominant firm gains by setting the price above its marginal cost and earning additional profit. Thus, tying together can be beneficial for the multi-product firm. In principle, separate firms can also sell their products too and such practices are quite common. But generally, such interfirm collusive agreements are illegal and subject to government intervention. But if it happened within the firm, it becomes difficult to track such processes. So this is also another rationale for multi-product firms along side other organisational economies.

(b) Cost-Relaled Reasons

The cost based rationale for multi-product firm is based on the concept of economies of scope, which is a special case of cost-sub-additivity. A cost function is sub-additive if it is cheaper to produce any given output vector in one batch than in a set of smaller batches. In the two-product case, this simply means that $C(q_1, q_2) < C(q_1) + C(q_2)$.

In general, economies of scope arise because input resources have the characteristics of a public good, i.e., the use of the input does not deplete its availability for other users. Examples can bc given of a database with many users of using elcctricity at night for other purpose, once it is installed day time production only. Of course, these examples of natural economies of scope associated with productions such as beef and hide or refining oil in a multiplicity of products. Multi-product firms facilitate a capacity utilisation over time by reducing indivisibilities or lumpiness in the production process.

Apart from overall firm economies of scope, one can distinguish the technological or plant economies of scope. These are with the production process itself while firm economies additionally include organisational economies.

(c) Financial Reasons

It is sometimes argued that by diversification and production of a wide range of outputs the firm spreads its risks, specially if the products are not close substitutes and thus likely to be affectcd by some environmenal products of market movements. This argument is not very tenable as such diversification does not add value to the company. As the shareholders are already diversified, they can get more advantage by diversifying their own portions, rather than the products.

(i) there may be some tax advantages associated with such diversification;

(ii) diversification facilitates dividend smoothing;

(iii) Diversification would benefit managers as it reduces a company's total risk;

(iv) larger the enterprise, the lower the costs of raising finance, thus benefiting the firm.

Thus, there is an assorted range of financial pressures which tend to encourage the formation of multi-product firms.

7.3.2. Multi-Product Pricing

Till now, we have shown that when there are economies of scope, there is an incentive to set up a multi-product firm. With common provision and use of common facilities, the cost function will feature some degree of interdependency in the sense that the marginal cost of production of each product may be affected not only by its output level, but also by that of the other products in the range.

Similarly, as the products in product range are often related and hence they have substitutes, demand.

Independence is also common. In such cases, profit-maximising prices cannot be determined independently. Table 7.1 shows the point.

Table 7.1 : The Multi-product pricing problem

	Demand		
		Independent (zero cross-price elasticties)	Independent (non-zero cross-price elasticties)
Cost	Independent	Price each product seperately	System solution required
	Independent	System solution required	System solution required

If there is interdependence either on the cost or on the demand side, then the theoretical problem of setting prices in order to maximise profits requires the solution of a set of simultaneous equations (equal to the number of products). It is not possible to identify this optimum by setting prices for the various products either sequentially or independently. If there is independence on both cost and the demand side, then the products may be priced in isolation. Here for optimising prices for multiple products, the demand and cost functions for each product are necessary. Here we would assume that demand and cost functions (also how cost varies with changes in output mix and volume) are known with certainty.

Suppose the firm manufactures n products. Ordinary demand functions would be of the form :

$q_i = f_i(p)$, i = 1, 2.....n.....(7.1)

where q^i = 1,.....n are the quantities demanded and

$p_i = (p_1, p_2.....p_n)$.....(7.2)

are the prices set by the firm. The prices of each of the firm's n products may have influence on the sales of each product although typically the own price would be the major determinant of an individual's sales of products. The total cost of producing output q is assumed to be a smooth cost function C(q). The firm's profits are thus given as

$$\pi = \left(\sum_{i=1}^{n} p_1 q_i \right) - C(q).....(7.3)$$

Given the ordinary demand function (7.1), the first-order necessary conditions for maximum profits are more easily obtained through analysis in price space. This treats n prices as choice variables and the first order conditions are thus :

$$\delta\pi|\delta p_j = q_j + \sum_{i=1}^{n} (p_i - \partial C(q)|\ \partial p_i)(\partial q_i\ |\ \partial p_j) = 0,\ j = 1.2,....., n.....(7.4)$$

In fact, so long as the demand and cost functions are reasonably well-behaved in conjunction with the demand equation (7.1), are sufficient to identify the optimum prices, outputs and attainable profit for the firm.

7.4. Some Special Pricing Techniques

7.4.1 Peak-load pricing

Many firms face a systematically varying demand for their product. The most common

examples are utilities like electricity, gas and telephone. Demand may fluctuate over a daily or weekly cycle or even an annual cycle. This implies that it may pay the firm to vary price systematically over time as well. The advantages of varying price will depend on

(i) substitutability of demand : the extent to which consumers can substitute demand at one point of time to another,

(ii) the extent to which the firm has to produce for demand just in time (rather than through accumulation of inventory).

Here we can show the example of electricity where demand during the daytime is considerably higher than at night. Consumers are unable to store electricity; if the firm charges a lower price electricily at night (the off-peak rate), this gives the consumers an incentive to purchase at night to use it during the day. If storage was costless, then purchase would have been done only at lower price (then it would be no more off-peak). So storage has to be difficult or costly for the consumers if peak pricing has to be sustainable. These are special types of perishable goods. This difficulty also affects the producer. If it is costly for the producer to store the product, he will also go for just-in-time production, This implies that peak level demand has to be dealt with directly, i.e., the installed capacity needs to cope with the anticipated peak level of demand and so, by opportunity cost reasoning, peak demand is more costly to provide than off peak. The marginal cost of providing one more unit of output at the peak is the sum of the marginal running costs plus the marginal cost of providing the additional unit of capacity, whereas since the capacity is already there, the marginal cost of off-peak period is simply the running cost. These differcnces in cost would also provoke the firm to differentiate in pricing between the peak and off peak periods, peak load pricing differs depending on the objective of the firm. Usually, peak-load pricing deal with welfare-maximising objective as in most of the cases. The firm is a public utility. However, following extensive privatisation all over the world, the profit maximizing firms should be studied.

SOME PROBLEMS AND THEIR SOLUTIONS :

1. The demand curve faced by a firm is $p = 110 - 2q$ and $c = q^3 - 3q^2 + 50q + 10$ is the cost function. Obtain the profit maximising output of the firm. Also obtain the Lerner's index of monopoly power, $\frac{p - mc}{p}$ in the concerned market.

Solution :

We know that $\pi = TR - TC = pq - c = (110 - 2q)q - q^3 + 3q^2 - 50q - 10$

That is, $\pi = 110q - 2q^2 - q^3 + 3q^2 - 50q - 10 = -q^3 + q^2 + 60q - 10$

To maximise π we are to set the first order condition $\frac{d\pi}{dq} = 0$.

But $\frac{d\pi}{dq} = -3q^2 + 2q + 60 = 0$

That is, $3q^2 - 2q - 60 = 0$.

That is, $q = \frac{2 \pm \sqrt{4 - 4(3)(-60)}}{2 \times 3} = \frac{2 \pm \sqrt{724}}{6} = \frac{2 \pm 26.9}{6} = 4.8$.

The second order condition needs $\frac{d^2\pi}{dq^2} = -6q + 2 = -6(4.8) + 2 = -26.8 < 0$.

Thus the profit maximizing level of output is 4.8 units.

The price is $p = 110 - 2 \times 4.8 = 100.4$

Also, $MC = \frac{dc}{dq} = 3q^2 - 6q + 50$. When $q = 4.8$, $MC = 3(4.8)^2 - 6(4.8) + 50 = 90.3$.

Now, monopoly power is $= \frac{p - mc}{p} = \frac{100.4 - 90.3}{100.4} = 0.10$

Thus, Lerner's indx of monopoly power = 0.10.

2. X, the monopolist, has the demand curve given by p = 20 – q and his cost function is $c = q^2 + 8q + 2$. Obtain the profit maximising output and the corresponding price of the monopolist.

Solution : We know that $\pi = TR - TC$

Now, $TR = pq = (20 - q)q = 20q - q^2$

Thus $\pi = 20q - q^2 - q^2 - 8q - 2 = 12q - 2q^2 - 2$

Naturally, the monopolist will want to maximise π.

For this we are to set $\frac{d\pi}{dq} = 0$.

Now $\frac{d\pi}{dq} = 0$ means $\frac{d}{dq}(12q - 2q^3 - 2) = 0$

That is, $12 - 4q = 0$, That is, $q = 3$.

Besides, $\frac{d^2\pi}{dq^2} = -4 < 0$. That is, the second order condition is also satisfied.

Hence, the profit maximising level of output of the monopolist is 3.

Naturally, $p = 20 - 3 = 17$.

3. The revenue function and the cost function of a firm are $R = 1000Q - 2Q^2$ and $C = Q^3 - 59Q^2 + 315Q + 2000$. Obtain the profit maximizing level of output.

Solution : Here $\pi = R - C = 1000Q - 2Q^2 - (Q^3 - 59Q^2 + 315Q + 2000)$

$= 1000Q - 2Q^2 - Q^3 + 59Q^2 - 315Q - 2000$

$= -Q^3 + 57Q^2 - 315Q - 2000$.

Now, $\frac{d\pi}{dQ} = 0 \Rightarrow \frac{d}{dQ}[-Q^3 + 57Q^2 - 315Q - 2000] = 0$

That is, $-3Q^2 + 114Q - 315 = 0$ That is, $-3(Q^3 - 38Q + 105) = 0$

That is, $Q^2 - 38Q + 105 = 0$ That is, $Q^2 - 35Q - 3Q + 105 = 0$

That is, $Q(Q - 35) - 3(Q - 35) = 0$ That is, $(Q - 35)(Q - 3) = 0$

Thus Q is either 3 or Q is 35.

Now, $= \frac{d^2\pi}{dQ^2} = -6Q + 114$ At $Q = 3, \frac{d^2\pi}{dQ^2} > 0$ and at $Q = 35, \frac{d^2\pi}{dQ^2} < 0$.

This means that the profit maximising level of output is 35 units.

7.7. Sample Questions

1. Consider a single-product firm for which the demand function is $p = 100 - q$ and total costs are $C(q) = q^2$. Find out the profit-maximising price, output and profitability of the firm. If the firm is a Baumol sales revenue maximiser, what is the choice of price and output if the MALP is 450? What is the difference in economic welfare in the two cases?

2. Critically discuss Williamson's Expense Preference Model.
3. Do you think that Marris's growth Maximisation Hypothesis gives a sufficient explanation of a firm's objective?
4. Why does a firm decide to produce multiple products?
5. Discuss the problem of multi-product pricing for demand-independent and demand-interdependent products. What are the problems of getting a solution for cost-dependent and cost-interdependent products?

SOME NUMERICAL PROBLEMS

4. You are given the following demand and cost functions $p = 250 - 3q$ and $c = 3q + 5q^2$ respectively. Obtain the profit maximizing price and output.

Solution : Let π stand for total profit and R stand for total revenue.

Now, $\pi = TR - TC = pq - TC = (250 - 3q)q - 3q - 5q^2 = 250q - 3q^2 - 3q - 5q^2 = -8q^2 + 247q$

For maximisation of profit $\frac{d\pi}{dq} = 0$. That is, $\frac{d}{dq}[-8q^2 + 247q] = 0$

That is, $-16q + 247 = 0$. That is, $q = \frac{247}{16} = 15.4$.

Also, $\frac{d^2\pi}{dq^2} = -16 < 0$.

Hence the second order condition is also satisfied. Therefore, the profit maximising level of output is q = 15.4 units = 15 units; price = $250 - 3 \times 15 = 205$ Ans.

5. A comptitive firm sells its output at a fixed price of Rs. 4 per unit. The cost function of the firm is given as $c = 0.04q^3 - 0.9q^2 + 10q + 5$. Find the profit maximising output level of the firm and determine the corresponding total profit, total revenue and total variable cost.

Solution : We have $\pi = TR - TC = 4q - [0.04q^3 - 0.9q^2 + 10q + 5] = 4q - 0.04q^3 + 0.9q^2 - 10q - 5 = -0.04q^3 + 0.9q^2 - 6q - 5$

Now, $\frac{d\pi}{dq} = -.12q^2 + 1.8q - 6 = 0$

$\Rightarrow -12q^2 + 180q - 600 = 0$

$\Rightarrow 12q^2 - 180q + 600 = 0$

$\Rightarrow q^2 - 5q + 50 = 0$

$\Rightarrow q^2 - 10q - 5q + 50 = 0$

$\Rightarrow (q - 5)(q - 10) = 0$

Thus q = 5 or 10.

The second order condition for maximisation needs $\frac{d^2\pi}{dq^2} < 0$.

That is, $-0.24q + 1.8 < 0$

When q = 5, $\frac{d^2\pi}{dq^2} = 0.6 > 0$. When q = 10, $\frac{d^2\pi}{dq^2} = -0.6 < 0$

Therefore, when q = 10, profit will be maximum

Here TR = Total Revenue = pq = 4 × 10 = 40

$\pi = -15$. That is, negative profit.

$TVC = 0.04q^3 - 0.9q^2 + 10q$

When q = 10, TVC = 50, TFC = 5. Thus, TC = 55 and TR = 40

Thus, π = TR – TC = 40 – 55 = –15.

6. A revenue maximising firm requries profit of at least Rs. 1500. Its demand and cost functions are : p = 304 – 2q and c = 500 + 4q + $8q^2$. Determine his output level and price. Compare these values with those that would be achieved under profit maximisation.

Solution : Here $TR = PQ = (304 - 2q)q = 304q - 2q^2$

$= 300q - 10q^2 - 500$

Now, $\pi = TR - TC$

$= 304q - 2q^2 - 500 - 4q - 8q^2$

$= 300q - 10q^2 - 500$

But the firm needs a minimum profit of Rs. 1500.

Hence, $\pi = 300q - 10q^2 - 500 = 1500$

$\Rightarrow 300q - 10q^2 - 2000 = 0$

$\Rightarrow 30q - q^2 - 200 = 0$

$\Rightarrow q^2 - 30q + 200 = 0$

$\Rightarrow q^2 - 20q - 10q + 200 = 0$

$\Rightarrow q(q - 20) - 10(q - 20) = 0$

$\Rightarrow (q - 20)\ (q - 10) = 0$

Thus, q is either 20 or 10.

But the firm needs maximisation of revenue. Hence q = 20.

Also, from $q^2 - 30q + 200 = 0$ we get $q = \frac{30 \pm \sqrt{900 - 800}}{2}$.

That is, q = 20 or 10.

Now, p = 304 2 × 20 = 264

TR = 304 × 20 – 2 × 20^2 = 6080 – 800 = 5280

C = 500 + 4 × 20 + 8 × 20^2 = 500 + 80 + 3200 = 3780.

Under profit maximisation, $\pi = -10q^2 + 300q - 500$

Now, $\frac{d\pi}{dq} = 0 \Rightarrow -20q + 300 = 0 \Rightarrow q = 15$

Also, $\frac{d^2\pi}{dq^2} = -20 < 0$ p = 304 – 2 × 15 = 274

c = 500 + 4 × 15 + 8 × 15^2 = 500 + 60 + 1800 = 2360

π = pq – 2360 = 274 × 15 – 2360 = 4110 – 2360 = 1750.

7. A fashion industry has the demand function p = –20 – Q and its total cost function is c = Q^2 + 8Q + 2. Determine the optimal output Q, price (p), total revenue R and total profit.

(a) under profit maximisation, (b) under sales revennue maximisatiion, (c) under sales revenue maximisation subject to the profit constraint of Rs. 10.

Solution : (a) $\pi = TR - TC = PQ - TC = PQ - TC$

$= (20 - Q)Q - Q^2 - 8Q - 2$

$= 20Q - Q^2 - Q^2 - 8Q - 2 = -2Q^2 + 12Q - 2$

Now, $\frac{d\pi}{dQ} = -4Q + 12 = 0 \Rightarrow Q = 3$

Also, $\frac{d^2\pi}{dQ^2} = -4 < 0.$

(b) Total sales or total sales revenue, $R = PQ = (20 - Q)Q = 20Q - Q^2$

Now, $\frac{dR}{dQ} = 0 \Rightarrow 20 - 2Q = 0$ that is, $Q = 10$

(c) $\pi = TR - TC = -2Q^2 + 12Q - 2 = 10$

That is, $\pi = -2Q^2 + 12Q - 12 = 0 \Rightarrow 2Q^2 - 12Q + 12 = 0$

That is, $Q = \frac{12 \pm \sqrt{144 - 4 \times 2 \times 12}}{4} = \frac{12 \pm \sqrt{144 - 96}}{4} = \frac{12 \pm 4\sqrt{3}}{4} = \frac{12 \pm 4\sqrt{3}}{4}$

$= 3 \pm \sqrt{3} = 3 \pm 1.7$

Thus Q = 4.7 and Q = 1.3 Actually, Q = 4.7 Ans.

8. The revenue function and the cost function of a firm are $R = 1000Q - 2Q^2$ and $C = Q^3 - 59Q^2 + 1315Q + 2000$, respectively. Find out the profit maximising level of output of the firm.

Solution : Here $\pi = R - C = 1000Q - 20Q^2 - (Q^3 - 59Q^2 + 1315Q + 2000)$

$= 1000Q - 2Q^2 - Q^3 + 59Q^2 - 1315Q - 2000$

$= -Q^3 + 57Q^2 - 315Q - 2000$

Now, $\frac{d\pi}{dQ} = 0 \Rightarrow \frac{d}{dQ}[-Q^3 + 57Q^2 - 315Q - 2000] = 0$

Thst is, $-3Q^2 + 114Q - 315 = 0$ That is, $-3(Q^2 - 38Q + 105) = 0$

That is, $Q^2 - 38Q + 105 = 0$ That is, $Q^2 - 35Q - 3Q + 105 = 0$

That is, $(Q - 35)(Q - 3) = 0$

Thus, Q is either 3 or Q is 35. Now, $\frac{d^2\pi}{dQ^2} = -6Q + 114$

At Q = 3, $\frac{d^2\pi}{dQ^2} > 0$ and at Q = 35, $\frac{d^2\pi}{dQ^2} < 0$.

That means, the profit maximizing level of output is 35 units.

7.5 Sample Questions

A. Objective-type Questions :

Choose the correct alternative :

(a) Sales maximisation hypothesis was introduced by : (i) Baumol, (ii) Simon, (iii) Williamson, (iv) Madrid.

Ans. (i)

(b) One of the alternative theories of the firm is : (i) Theory of Simon, (ii) Theory of Baumol, (iii) Theory of Samuelson, (iv) Theory of Joan Robinson.

Ans. (ii)

(c) The condition for maximisation of profit is attained at the point where : (i) MR = MC, (ii) MR = 0, (iii) MC cuts MR from below, (iv) MC cuts MR from above.

Ans. (i)

(d) The condition for maximisation of revenue is : (i) MC = 0, (ii) MR = 0, (iii) AR = 0, (iv) AC = 0.

Ans. (ii)

(e) Williamson's theory of firm deals with the theory of profit maximisation : (i) Correct, (ii) incorrect, (iii) uncertain, (iv) cannot be said a-priovin.

Ans. (ii)

B. Short-type Questions :

(a) Write a short note on peak-load pricing

(b) What do you know about multiple-product pricing?

(c) How can you decive the condition for getting maximum sales?

(d) Obtain, mathematically, the relation among MR, AR (= price) and the price elaoticity of demand. Hence derive the condition for attaining maximum revenue.

(e) Explain, in brief, the welfare implications of Williamson's model.

C. Essay-type Questions :

(a) Explain the Marris's model of managerial discretion. How would you compare it with Williamson's model?

(b) Explain why managers in many occasions consider sales maximisation as a more alternative goal vis-a-vis profit maximisation. "The objective of sales maximisation subject to a 'satisfactory' level of profit is attained at a level of output higher than that of profit maximisation." How can you arrive at such a statement? Give reasons for your answer.

(c) Explain, clearly, Williamson's expense preference model.

(d) Explain whether you think that Marris's growth maximisation hypothesis gives a sufficient explanation of a firm's objective.

(e) What arguments can you prescribe against profit maximisation hypothesis?

Unit 8 ❐ Inflation

Structure

8.1. Objectives

The primary objectives of this unit are :

i. To give the definition of inflation.
ii. To discuss different types of inflation.
iii. To explain the effects of inflation.
iv. To explain Phillips Curve.

8.2. Introduction

Inflation which now-a-days confronts the economic policy makers througout the whole world in the form of a dominant economic problem is, not at all, a new phenomenon since from the earliest days of recorded history, mankind has been puzzled by the rising prices. Besides, inflation is one of the major issues in Macroeconomics as it affects various aspects of an economy. It also assumes enormous importance in Managerial Economics because performance of business is closely linked to inflation. Naturally, it is essential to examine different types of inflation and various conditions that lead to inflationary situtations.

8.3. Definition

In Economics, the term "inflation" has been defined differently by different authors and there is no unanimity among these definitions. Some economists have defined inflation as a stage of rising price, not as a stage of rise in prices. There may be some situations where there is a rise in prices, but after a period of time this rise in prices gets stopped. In such a situation we say that this is not an inflationary situation. There may be different definitions of inflation. However, we shall give here the definition of inflation as had been advanced by Gardner Ackley in his Macroeconomic Theory. Accorcing to Professor Ackley, "Inflation is a persistent and appreciable rise in the general or average of prices". According to this definition, a sporadic price spurt or an imperceptible rise in prices will not be inflation. Professor Ackley has, further, stated that we define inflation as rising prices, not as "high" prices. In some sense inflation is, thus, a state of disequilibrium; it must be analysed dynamically rather than with the tools of statics.

8.4. Different Types of Inflation

Inflation is mainly caused by excess demand or decline in aggregate supply or output. The former leads to a rightward shift of the aggregate demand curve while the latter causes the aggregate supply curve to shift leftward. The former is called (1) demand- pull inflation (DPI) while the latter is called (2) the cost-push inflation (CPI).

8.5 Demand-pull Inflation Theory

There are two theoretical approaches to the demand-pull inflation : (a) the classical approach and (b) the Keynesian approach.

According to the classical economists or monetarists, inflation is caused by an increase in money supply which leads to a rightward shift in the negatively sloping aggregate demand curve. Given a situation of full employment, classical economists maintained that a change in money supply brings about an equiproportionate change in the price level. That is why monetarists argue that inflation is always and everywhere a monetary phenomenon. The Keynesians do not find any link between money supply and price level causing an upward shift in aggregate demand. According to Keynesians, aggregate demand may rise due to a rise in consumer demand or investment demand or Government expenditure or net exports or the combinations of these four components of aggregate demand. Given full employment, such increase in aggregate demand leads to an upward pressure in prices. Such a situation is referred to as "Demand-Pull Inflation" which can, graphically, be explained in the following diagram (Diagram 8.1).

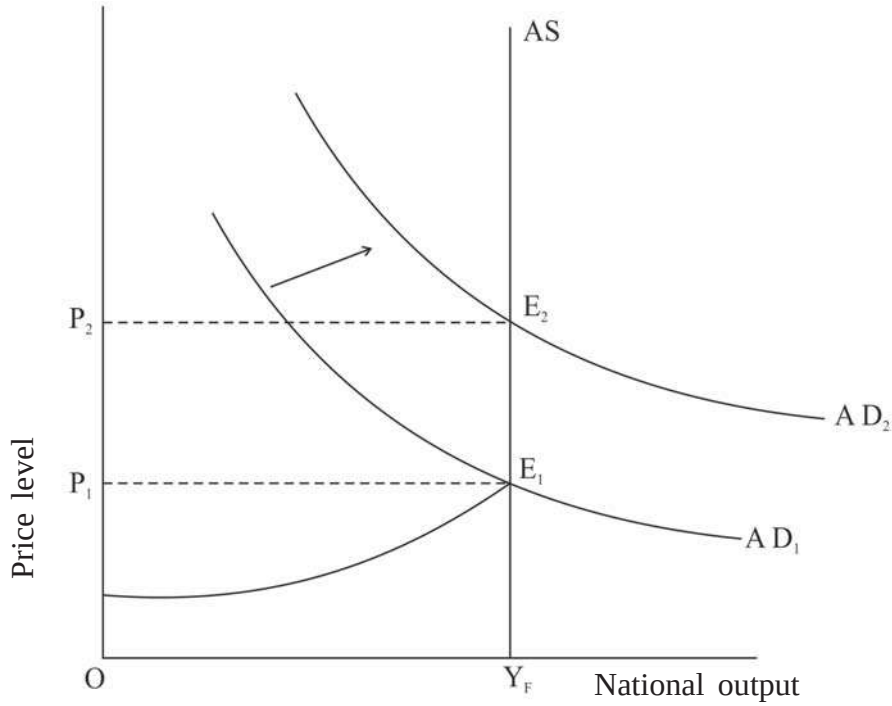


Figure 8.1 : Demand-Pull Inflation : Shifts in AD curves

Just like the determination of price of a commodity, the level of prices is determined by the interaction of aggregate demand and aggregate supply. In diagram 8.1, the aggregate supply curve before the full empolyment stage is positively sloping and becomes vertical after the full employment stage is reached. AD_1 is the initial aggregate demand curve that intersects the aggregate supply curve AS at point E_1. The determined price is thus OP_1. As the aggregate demand curve shifts from AD_1 to AD_2, price level rises to OP_2. Thus, an increase in aggregate demand at the full employment stage leads to an increase in the price level only, rather than the level of output. However, how much price level will rise following an increase in aggregate demand depends on the slope of the AS curve.

8.5.1 Causes of Demand-Pull Inflation

The DPI originates in the monetary sector. Monetarists' argument that "only money matters" is based on the assumption that at or near full employment excessive money supply will increase aggregate demand and will, thus, cause inflation.

An increase in nominal money supply shifts aggregate demand curve rightward. This enables people to hold excess cash balances. Spending of excess cash balances by them causes price level to rise. Price level will continue to rise until aggregate demand equals aggregate supply.

Keynesians argue that inflation originates in the non-monetary sector or the real sector. Aggregate demand may rise if there is an increase in consumption expenditure following a tax cut. There may be an autonomous increase in business investment or Government expenditure. Govenrment expenditure is inflationary if the needed money is procured by the Government by printing additional money. In brief, increase in aggregate demand, that is, increase in (C + I + G + X – M) causes price level to rise. However, aggregate demand may rise following an increase in money supply generated by the printing of additional money (Classical argument) which drives prices upward. Therefore, money plays a vital role. That is why Miltion Friedman aruges that inflation is always and everywhere a monetary phenomenon.

There are other reasons that may push aggregate demand and, hence, price level upwards. For instance, growth of population stimulates aggregate demand. Higher export earnings increase the purchasing power of the exporting countries. Additional purchasing power means additional aggregate demand. Purchasing power and hence, aggregate demand may also go up, if Government repays public debt. Again, there is a tendency on the part of the holders of black money to spend more on conspicuous consumption goods. Such a tendency fuels inflationary fire. So, demand pull inflation is caused by a variety of factors.

8.6 Cost-push Inflation

Like aggregate demand, aggregate supply also generates inflationary process. As inflation is caused by a leftward shift of the aggregate supply we call it the cost push inflation (CPI). The CPI is normally associated with non-monetary factors. The CPI arises due to the increase in cost of production. The cost of production may rise due to a rise in cost of raw materials or increase in wages. However, wage increase may lead to an increase in productivity of workers. If this happens, the aggregate supply (AS) curve will shift to the rightward, not leftward direction. Here we assume that productivity does not change in spite of an increase in wages. Such increase in costs are passed on to consumers by firms by raising the prices of the products. Rising wages lead to rising costs. Rising costs, again, lead to rising prices. Again, rising prices prompt trade unions to demand higher wages. Therefore, an inflationary wage-price spiral starts. This causes the aggregate supply curve to shift leftward.

The cost push inflation has been demonstrated in diagram 8.2 where national output has been measured horizontally while price level has been measured vertically. Below the full employment stage this AS curve is positively sloping and at full employment stage it becomes perfectly inelastic.

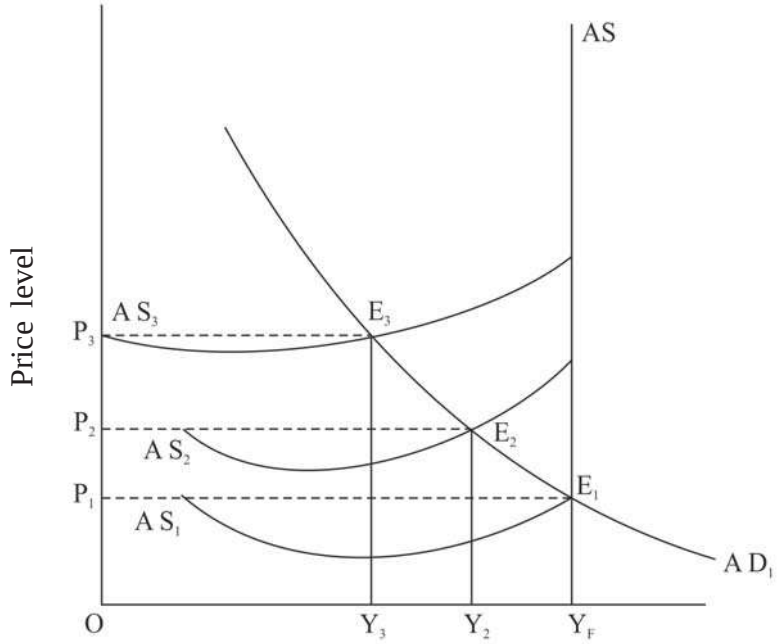


Figure. 8.2 : National Output

The point E_1, that is, the intersection point between AD_1 and AS_1 curves determine the price level (OP_1). There is now a leftward shift of aggregate supply curve to AS_2. With no change in aggregate demand, this causes price level to rise to OP_2 and output to fall to OY_2. With the reduction in output, employment in the economy declines or unemployment rises. Further shift in AS curve to AS_3 results in the higher price level, OP_3 and a lower volume of aggregate output (OY_3). Therefore, cost push inflation may arise even below the full employment (Y_F) stage.

8.6.1 Causes of Cost-Push Inflation

It is the cost factors that pull the prices upward. One of the important causes of price rise is the rise in price of raw materials. For instance, by an administrative order the Government may hike the price of petrol or diesel or freight rate. Now at higher prices the firms have to buy these inputs. This leads to an upward pressure on cost of production. Not only this the CPI is sometimes imported from outside the economy.

Increase in the price of petrol from OPEC compels the Govt. to increase the price of petrol and disel. These two important raw materials are needed by every sector, especially the transport sector. As a result, transport costs go up resulting in higher general price.

Besides, the CPI may be induced by wage-push inflation or profit-push inflation. Trade unions demand higher money wages as a compensation against inflationary price rise. If increase in money wages exceed labour productivity, aggregate supply will shift upward and leftward. Firms often exercise power by pushing prices up independently by pushing prices, up independently of consumer demand to expand their profit margins.

Fiscal policy changes, such as increase in tax rates also leads to an upward pressure in cost of production. For example, an overall increase in excise tax of mass consumption goods is inflationary. That is why, the Government is, then, accused of causing inflation.

Finally, production setbacks may result in decreases in output. Natural disaster, gradual exhaustion of natural reources, work stoppages, electric power cuts etc. may cause aggregate output to decline. In the midst of this output reduction, artificial scarcity of any goods created by traders and hoarders just simply ignite the situtation.

Inefficiency, corruption, mismanagement of the economy may also be the other reasons. Thus, inflation is caused by the interplay of various factors. A particular factor cannot be held responsible for any inflationary price rise.

8.7 Inflationary Gap

A very important concept in connection with the theory of inflation is “Inflationary Gap”. Moreover, the excess demand inflationary analysis has been described in terms of this inflationary gap. This gap is measured as the excess of the aggregate demand, that is, C + I + G over the full-employment aggregate supply Y_F as has been dimonstrated in diagram 8.3. That means, the situation $C + I + G > Y_F$ is an indicator of the presence of inflationary gap in an economy. On the other hand, if $C + I + G < Y_F$ there is presence of deflationary gap in the economy. Obviously, both the situations of inflationary gap and deflationary gap are situations of disequilibrium.

The concept of ‘inflationary gap’ may be defined as the positive difference between the actual level of aggregate demand which exists in an economy at the full employment level of income and the amount of aggregate demand which is needed to attain full employment. Again, since the amount of aggregte demand needed for attaining full employment should be equal to the amount of aggregate supply at full employment, we may also say that inflationary gap is the excess of the actual aggreate effective demand in the economy over the aggregate supply at the full employment level.

Let us draw, in diagram 8.3, a vertical line Y_FK at the full employment income Y_F

and we see that for full employment to exist in the economy, the aggregate demand function, C + I + G, in the said diagram should cut the 45º line at its intersection with the full employment line Y_FK. Since the aggregate demand schedule cuts the 45º line Y ≡ C + I + G at point E which is located to the right and above B, the point of intersection of the 45º line with the full employment line Y_FK, in the economy there is inflationary gap of the magnitude of AB present in the economy. This gap has been measured as a distance between the aggregate demand schedule C + I + G and the 45º-line indicating Y = C + I + G at the full employment income, Y_F.

In the figure aggregate consumption has been shown as a function of aggregate income Y, and the linear consumption function assumed to be of the form C = a + by, a being autonomous consumption while b being the marginal propensity to comsume (mpc), 0 < b < 1.

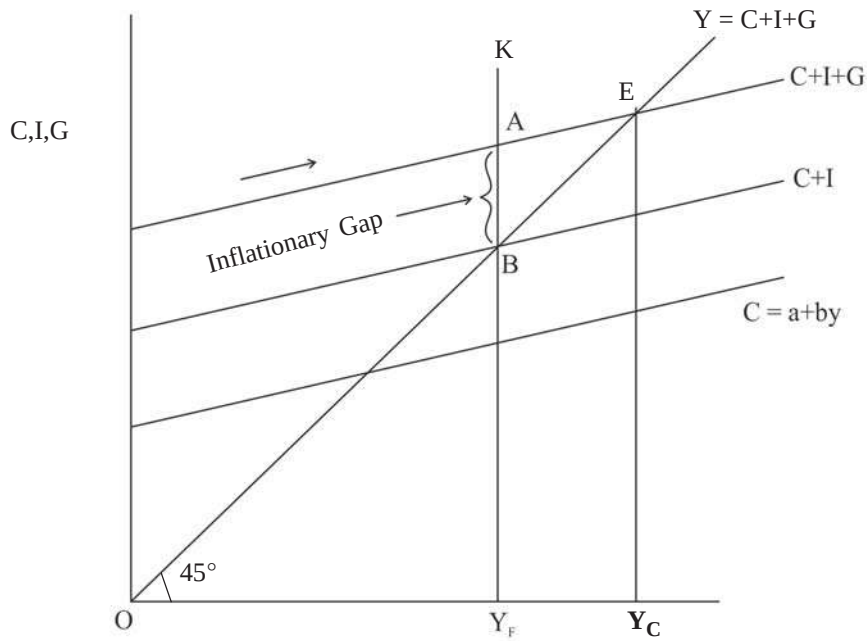


Figure 8.3 : Aggregate Income (Y)

Aggregate investment, I, and government spending, G, are both autonomously determined. Hence the aggregate demand function C + I + G has been drawn parallel to the consumption funtion, the distance between the two being I + G. If aggregate real output could increase without limit, then aggregate real income would increase to Y_C at which level the aggregate real output (that is, supply) and aggregate real expenditure will be in equilibrium. If, on the other hand, aggregate real output or supply cannot increase beyond Y_F, the real income cannot also exceed Y_F at which the

aggregate demand (C + I + G) exceeds the aggregate output by AB, that is, Y_F + AB = C + I + G. Naturally, AB is the magnitude of inflationary gap which, obviously, will raise the general price level. This gap will disappear only when the aggregate money in the general price level increases from Y_F to Y_C causing an increase in the general price level.

The increase in the money income is wholly due to the increase in prices as the aggregate real output is constant at the full employment output, Y_F. The magnitude of inflationary gap gets eliminated and the inflationary process is halted when the new equilibrium aggregate money income is established at Y_C. The change (increase) in the aggregate money income from Y_F to Y_C is wholly accounted for by the increase in prices with no increase having taken place in the aggregate real output remaining fixed at the full employment output, Y_F.

8.8 Effects of Inflation

From the standpoint of reality inflation is a very unpopular matter in any economy Various opinion surveys conducted in the USA, India and many other countries of the world reveal that inflation is the most important concern of the people of any society as it badly affects their standard of living. Many people are of the opinion that inflation is enemy number one. A high rate of inflation makes the life of the poor really miserable. It redistributes income and wealth in favour of some and harms others greatly. By making the rich richer and also the poor poorer, it militates against social justice. Moreover, inflation lowers national output and employment and impedes long run economic growth. This is specially true in many developing countries like India.

The effects of inflation can, broadly, be divided into the following categories :

(a) Effects on real income

(b) Effects on distribution of income and wealth

(c) Effects on output

and (d) Effects on long-run economic growth.

Let us now discuss the above-mentioned effects of inflation, in brief :

8.9 Effects on Real Income

Inflation erodes real incomes of the people. The effects of inflation can be examined if the difference between money income or nominal income and real income can, clearly, be understood. It is the change in the general price level that makes the crucial difference between these two concepts. We know that money income means the incme such as wages, interest and rent received in terms of rupees. Real income, on the other hand, implies amount of goods and services which we can buy with the money income. That

means, real income is the purchasing power of a person's money income. If our nominal income or money income increases at a lower rate than the rate of increase in the general price level (that means, the rate of inflation) people will be able to buy less goods and services, that is, people's real income will be lower. Real income will increase only if nominal income rises faster than the rate of inflation. Inflation, obviously, reduces the purchasing power of money and, therefoe, it, adversely, affects real income of the people.

8.10 Effects of inflation on distribution of income and wealth

A vital effect of inflation is that it redistributes income and also wealth in favour of some at the cost of some others. Inflation, adversely, affects those who receive relatively fixed income and goes in favour of businessmen, traders, producers and some others who get flexible incomes. It brings windfall profits for the producers and also traders. Naturally, all people do not lose as a result of inflation; rather some people gain from it. However, unanticipated inflation harms creditors and goes in favour of the debtors. Obviously, inflation, in this way, redistributes income in favour of the latter. Obviously, those who get fixed incomes, stand to lose from the effect of inflation. Workers and salaried people who earn from fixed wages and salaried people who earn from fixed wages and salaries, get hit extremely in consequence of unanticipated inflation. These people very often enter into contract with the employers regarding wages or salaries fixed in nominal terms. Naturally, when inflation occurs, the purchasing power of their nominal income falls to a great extent leading to a high fall in their levels of living. Obviously, when inflation persists for some years, we notice great demands for revision of wages and salaries. Pensioners also fall in this category of the people who get income in fixed nominal terms. Definitely, businessmen, that is, entrepreneurs and traders stand to get a gain by inflation. During periods of inflation the prices of goods produced by entrepreneurs rise relatively faster than the cost of production since wages lag behind the rise in prices of goods. Consequently, inflation leads to an increase in the profits of the businessmen.

Inflation also, adversely, affects wealth holders who hold their wealth in the form of cash, demand deposits, savings and fixed deposits and also interest bearing bonds and debentures. Naturallay, these wealth holders get severely hurt by inflation, as it reduces the real value of their wealth.

Inflation reduces the real rate of interest earned by the depositors. Consequently, it has been observed that during periods of rapid inflation people like to convert their holdings of money into golds and physical property in order to advoiding the loss due to inflation.

8.11 Effects on Output

There is a good deal of uncertainty and at the same time disagreement as to whether inflation will favourably or adversely affect national output. It can be said that its effect on output also depends on whether it has been caused by demand-pull or cost-push factors.

8.12 Effects Of Inflation On Long-run Economic Growth

Many economists argue that inflation of a mild nature has a tonic effect on the long-run economic growth. The driving force in the process of economic growth has been high profit margins created by inflation. They go on arguing that wages lag behind the rise in the general price level and thus create higher profit margins for businessmen and industrialists. This, naturally, tends to increase the profit share in national income. Obviously, savings go up and this, as a consequence, ensures higher rate of investment. With greater rate of inflation more accumulation of capital goods gets made possible. This accumulated capital generates a higher rate of long-run economic growth.

8.13 Bad Effect of Inflation on Economic Growth

However, it is widely recognised that far from encouraging savings and encouraging higher rate of economic growth, inflation slows down the rate of capital accumulation. There are, however, many reasons behind this observation :

(i) When because of rapid inflation value of money gets declined, people will not like to keep money with themselves and will, therefore, be eager to spend it before its value goes down heavily. This leads to an increase in the consumption demand and, therefore, lowers their saving. This discourages them to go on saving.

(ii) Inflation leads to unproductive form of investment in gold, jewellery, real-estate, construction of houses etc.

(iii) A highly considerable consequence of inflation is that it acclerates the problem of poverty in the underdeveloped countries.

(iv) Finally, inflation adversely affects the balance of paymets (BOP) of a country and, thereby, hampers economic growth of the country concerned.

8.14 Phillips Curve

The famous “Phillips Curve” has been discussed in the context of inflation.

The relationship between inflation and unemployment had drawn considerable interest among economists throughout the world over the last seventy years or so. A noted British

economist, A.W. Phillips, published an article in the year 1958. In doing so he used historical data on the rate of wage inflation and unemployment rate of the United Kingdom spanning almost a century from 1861 to 1957. In his study he showed that, in fact, there existed an inverse relationship between the rate of wage inflation and the rate of unemployment. It was also shown that a similar negative relationship holds for rate of change of prices (that is, inflation) and the unemployment level. However, the rate of wage increase had long been taken to be the primary determinant of the rate of price increase (or vice-versa). Hence a Phillips curve can, easily, be restated as a relationship between unemployment and the rate of price inflation. This inverse relationship implies that for reducing unemployment price in the form of a high rate of inflation has to be paid while for reducing the rate of inflation, price in terms of a higher rate of unemployment has to be done.

On graphically fitting a curve to the historical data, a downward sloping curve exhibiting the inverse relationship between the rate of inflation and rate of unemployment was obtained. This trade-off is known to exist in the short run, not in the long run. This has been presented in diagram 8.4 where along the horizontal axis rate of unemployment has been measured while along the vertical axis rate of change in price and rate of change in wage have been indicated.

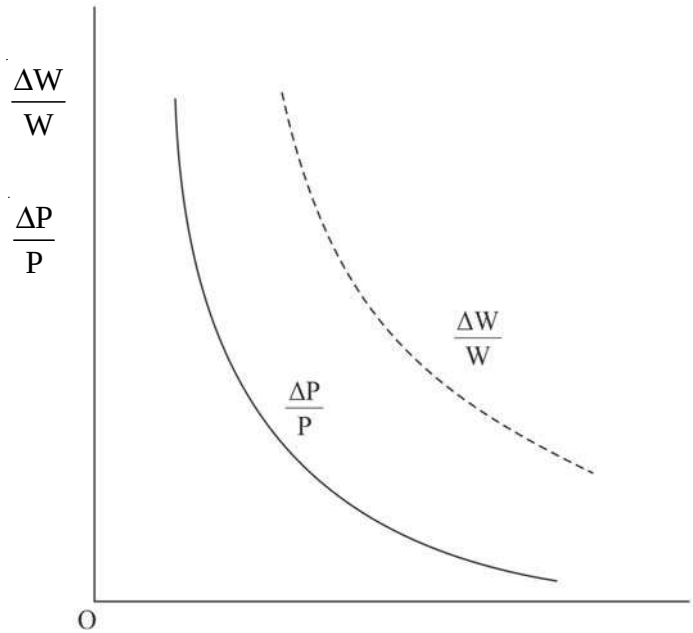


Figure. 8.4 : Rate of Unemployment

The Phillips curve drawn in diagram 8.4 is downward sloping from left to right. This suggests that there is a trade-off between inflation and unemployment.

Originally, the Phillips curve was constructed in terms of the negative relationship between the rate of increase in wage rate, $\left(\frac{\Delta W}{W}\right)$ and unemployment rate ($\dot{U}$).

Now, there is a direct relationship between the inflation rate and wage rate. For illustration, every thing being equal, if initially, money wage rate increases by 7% per year and labour productivity increases by 3% per year, then the price level will increase by 4%. This suggests, immediately, that there will be inflation rate to the tune of 4%. A rise in the money wage rate from 7% to 9% per year with no change in the 3% growth rate of labour productivity will tend to raise the inflation rate from 4% to 6%. Hence a rise in the wage rate without an increase in productivity will have a direct effect on inflation.

Now, if in addition to this direct effect between the rate of wage increase and the price increase, there is also an inverse relationship between the rate of wage increase and rate of unemployment, then it follows that there is an inverse relationship between the inflation rate and the unemployment rate. This analysis of the particular relationship between the rate of increase in price and the unemployment rate has been shown in the short run version of the Phillips curve as has been demonstrated in diagram (8.4).

In the above diagram both the rate of change of wage rate ($\Delta w/w$) and the rate of inflation ($\Delta p/p$) have been measured along the vertical axis while the rate of unemployment has been shown along the horizontal axis. The upper curve relates $\frac{\Delta W}{W}$ to $\dot{U}$ whereas the lower one relates $\frac{\Delta P}{P}$ to $\dot{U}$, which is based on the assumption of 3% increase in labour productivity per year. The relationship between the wage rate and the unemploymet rate assumes that $\frac{\Delta W}{W}$ will be 3% per year if $\dot{U}$ is 9%. Accordingly, the Phillips curve reveals that $\frac{\Delta P}{P}$ will be zero or the price level will be stable if $\dot{U}$ is 9%. Again, if $\frac{\Delta W}{W}$ is 9% with $\dot{U}$ of 6%, $\frac{\Delta P}{P}$ is 6%.

From the above analysis we can, easily, observe that there is a negative relationship between the unemployment, rate and rate of change in wage rate which, in turn, indirectly shows the negative relationship between the rate of change in inflation and unemployment.

8.15 Sample Questions

A. Objective-type Questions :

Choose the correct alternative :

(1) Inflation-unemployment trade off had, first, been shown by : (a) Milton Friedman, (b) J. M. Keynes, (c) J. R. Hicks, (d) A. W. Phillips.

Ans. (d)

(2) Which of the following is the cause of inflation? : (a) Increase in money supply, given full employment, (b) Increase in cost of prodution, (c) Increase in aggregate demand over available output, (d) All of the above.

Ans. (d)

(3) One of the following is not true : (a) Cost-push inflation may appear only at the full-employment stage, (b) Demand-pull inflation appears after full employment, (c) Demand-pull inflation is caused by both monetary and non-monetary factors, (d) Cost-push inflation is also known as wage-push inflation.

Ans. (a)

(4) Phillips curve shows the relationship between inflation and unemployment as : (a) A positive one indicating trade-off between the two, (b) An inverse one indicating trade off between the two, (c) A direct and proportional one, (d) None of the above.

Ans. (b)

(5) One of the following is not true during inflation : (a) Wage-earners gain, (b) Profit-earners gain, (c) Real incomes decline, (d) Speculators gain.

Ans. (a)

(6) When there is inflation the economic condition of the pensioners gets : (a) Improved, (b) Deteriorated, (c) Remains unaltered, (d) Cannot be ascertained.

Ans. (b).

B. Short-type Questions :

1. How is inflation caused by demand-pull factors?
2. What is meant by inflation? It is said that during inflation the rich becomes richer and the poor becomes poorer. Justify the statement.

3. What is a Phillips curve? What does this curve indicate?
4. Distinguish between demand-pull and cost-push inflation. How do you separate them from each other?
5. What is Inflation? Briefly discuss the different types of inflation that are experienced in an economy.

C. Essay-type Questions :

1. What do you mean by inflation? Show its efffects on (i) income distribution and (ii) economic growth.
2. How do you describe demand-pull inflation and cost-push inflation in terms of aggregate demand and supply curves? How is the disticntion between these varieties of inflation made?
3. Explain with the help of a neat diagram the causes of “inflationary gap”.
4. What is a Phillips curve? Draw such a curve and explain its nature.
5. “From the standpoint of reality inflation is a very unpopular matter in an economy”. Do you aggee with this view? If so, substantiate reasons behind your opinion.

Suggested Readings

Ackley, G : Macro Economics

Bilas (1967) : Microeconomic Theory

D. N. Dwivedi : Managerial Economics

Dean, J. (1951) : Managerial Economics (PHI)

Dwivedi, D. N. (2002) : Managerial Economics, Sixth Revised Edition, Vikash Publishing House Pvt. Ltd.

Ferguson (1969) : The Neo-classical Theory of Production and Distribution, Cambridge.

Ferguson and Gould (1989) : Microeconomic Theory, Irwin Publications in Economics.

Heathfield : Production Function

Henderson & Quandt : Microeconomic Theory, A Mathematical Approach

Koutsoyiannis, A (1979) : Modern Microeconomics (Macmillan)

Mithani, D. M. (2000) : Managerial Economics, Himalaya Publishing House, Hyderabad.

Pindyck and Rusinfeld : Microeconomics (Prentice Hall of India)

Ryan and Pearce : Price Theory (Macmillan)

Samvelson, P. A. (1948) : Foundations of Economic Analysis (Harvard University Press).

Sarkhel, J : Macroeconomics, Book Syndicate Private Ltd. College Street, Calcutta-700012.

Sarkhel, Jaydeb (2012) : Microeconomic Theory (Book Syndicate Calcutta, Pvt. Ltd. Calcutta): 700073.

Savage and Small (1967) : Introduction to Managerial Economics, Hutchison University.

Sen, A : Microeconomic Theory and Applications (Oxford University press).

Spencer and Siegelman (1959) : Managerial Economics, Richard D. Irwin.

Vaish, M. C. : Macroeconomics, Himalayan Publishing House.

Varshney and Maheswari : Managerial Economics, Sultan Chand and Sons.

Notes

Notes