

# **Micro Economics**

## **CENTRE FOR DISTANCE & ONLINE EDUCATION**



Accredited with NAAC **A** Grade

**12-B Status from UGC**

**Programme: Master of Economics**

**Course: Micro Economics**

**Course Code: MAECC301**

**Semester-I**

---

## Programme Coordinator

---

***Dr. Vibhor Jain***

Associate Professor, Department of Arts  
Center for Distance and Online Education,  
Teerthanker Mahaveer University, Moradabad

---

## Units Written by

---

**Dr. Aditya Kumar Sharma**

---

## Editor

---

***Professor Vipin Jain***

*Director,*

Center For Distance and Online Education,  
Education, Teerthanker Mahaveer University,  
Moradabad

***Dr. Amit Kansal***

Assistant Director,

Center for Distance and Online  
Teerthanker Mahaveer University,  
Moradabad

---

**Copyright** : **Teerthanker Mahaveer University**

**Edition** : **2023 (Restricted Circulation)**

**Published by** : **Teerthanker Mahaveer University, Moradabad**

---

**SYLLABUS**  
**MICROECONOMICS**  
**MAECC101**

**BLOCK I: BASIC OF MICRO ECONOMICS**

Unit-1: Basic Economic Problem - Choice and Scarcity; Distinction between Micro and Macro Economics.

Unit-2: Methods of Analysis: Deductive and Inductive Methods of Analysis - Economic Statics and Dynamics.

**BLOCK II: CONSUMER BEHAVIOUR AND DEMAND ANALYSIS**

Unit-3: Consumer Behaviour - Cardinal and Ordinal Utility - Law of Diminishing Marginal Utility - Law of Equi - Marginal Utility.

Unit-4: Demand: Demand Schedule and Demand Curve – Determinants of Demand.

Unit-5: Elasticity of Demand: Definition and Types of Elasticity of Demand.

Unit-6: Measurement of Elasticity: Measurement of Price Elasticity of Demand: Determinants of Price Elasticity of Demand - Demand and Supply Equilibrium.

**BLOCK III: PRODUCTION AND COST**

Unit-7: Factors of Production - Production Function - Short Run and Long Run.

Unit-8: Production Function - Law of Variable Proportions - Laws of Returns - Isoquants - Producer's Equilibrium.

Unit-9: Cost Function: Meaning - Types - Cost in Short and Long Run - Importance.

Unit-10: Revenue Function: Meaning - Types - Short and Long Run Revenue - Price - Profit.

**BLOCK IV: DIFFERENT MARKET STRUCTURES**

Unit-11: Market Structure: Characteristics - Market Schedule- Perfect Competition and Imperfect Competition.

Unit-12: Market Pricing: Pricing under Different Market Structure.

Unit-13: Monopoly - Monopolistic Competition - Meaning – Features - Price Discrimination.

Unit-14: Duopoly and Oligopoly - Concepts - Features – Price Determination under Duopoly and Oligopoly.

**BLOCK - I**  
**BASIC OF MICRO ECONOMICS**

---

**UNIT 1 ECONOMY AND ITS BASIC PROBLEMS**

---

*Structure*

- 1.0 Introduction
- 1.1 Objectives
- 1.2 Basic Economic Problem
- 1.3 Problems of Choice and Scarcity
  - 1.3.1 Basic Economic Decisions
  - 1.3.2 How the Market Mechanism Solves the Basic Problems
- 1.4 The Scope of Economics
  - 1.4.1 Distinction between Micro and Macro Economics
- 1.5 Answers to Check Your Progress Questions
- 1.6 Summary
- 1.7 Key Words
- 1.8 Self Assessment Questions and Exercises
- 1.9 Further Readings

---

**1.0 INTRODUCTION**

---

The basic economic problem asserts that an economy's finite resources are insufficient to satisfy all human wants and needs. It assumes that human wants are unlimited, but the means to satisfy human wants are limited. It is often said that the central purpose of economic activity is the production of goods and services to satisfy our changing needs and wants. The basic economic problem is about scarcity and choice. Because of scarcity, choices have to be made by consumers, businesses and governments.

In this unit, you will study about basic economic problem in detail. The unit also distinguishes between micro economics and macroeconomics.

---

**1.1 OBJECTIVES**

---

After going through this unit, you will be able to:

- Explain basic economic problem
- Discuss the problems of choice and scarcity
- Distinguish between micro and macro economics

---

## 1.2 BASIC ECONOMIC PROBLEM

---

A natural curiosity of a student who begins to study a subject or a science is to know the nature and scope of his subject of study. Such as it is, a student of economics would like to know 'what is economics' and 'what is its subject matter'. Surprisingly, there is no precise answer to these questions. Attempts made by economists over the past 300 years to define economics have not yielded a precise and universally acceptable definition of economics. Economists right from Adam Smith—the 'father of economics'—down to modern economists have defined economics differently depending on their own perception of the subject matter of economics of their era. For example, Adam Smith (1776) defined economics as 'an inquiry into the nature and causes of the wealth of the nations'. Nearly one-and-a-half century later, Alfred Marshall, an all time great economist, defined economics differently. According to Alfred Marshall (1922), "Economics is the study of mankind in the ordinary business of life; it examines that part of individual and social action which is most closely connected with the attainment and with the use of the material requisites of well being". Lionel Robbins (1932) has defined it more precisely: "Economics is the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses". One can find a number of other definitions in economics literature. None of the definitions of economics, however, captures the entire subject matter of modern economics, though they do throw some light on what economics is about.

The fact is that economics has not yet been defined precisely and appropriately. The reason is, as Zeuthen has observed, "Economics is an unfinished science" and as Schultz has remarked, "Economics is still a very young science and many problems in it are almost untouched". These observations made half-a-century ago hold still true. It seems that, after Robbins, no serious attempt was made to define economics. Defining economics has been so fruitless an effort that some modern authors of economics texts, including those by reputed economists like Samuelson, Baumol, and Stiglitz avoid the issue of defining economics. For example, William J. Baumol (a Nobel laureate) and Allen S. Blinder write in their own text, "Many definitions of economics have been proposed, but we prefer to avoid any attempt to define the discipline in a single sentence or paragraph", and let "the subject matter speak for itself."

However, the study of economic science, or of any science for that matter, must commence with a working definition of it. In this regard, most modern texts follow Robbins' definition of economics, even though modern economics goes far beyond what Robbins thought to be the subject matter of economics. Let us begin with Robbins' view on the subject matter of economics and then look how far it goes beyond his view.

*Economics* as a social science studies *economic behaviour* of the people and its consequences. What is economic behaviour? *Economic behaviour* is essentially the process of evaluating economic opportunities open to an individual or a society and, given the resources, making choice of the best of the opportunities. The objective behind this economic behaviour is to maximize gains from the available resources and opportunities. In their efforts to maximize their gains from their resources, people have to make a number of choices regarding the use of their resources and spending their earnings. The basic function of economics is to observe, explain and predict how people (individuals, households, firms and the government) as decision-makers make choices about the use of their resources (land, labour, capital, knowledge and skills, technology, time and space, etc.) to maximize their income, and how they as consumers decide how to spend the income to maximize their total utility. Thus, economic is fundamentally the study of choice-making behaviour of the people. The choice-making behaviour of the people is studied in a systematic or scientific manner. This gives economics the status of a social science.

For the purpose of economic analysis, people are classified according to their decision-making capacity as *individuals, households, firms* and the *society*, and according to the nature of their economic activity as *consumers, producers, factor owners* and *economy managers*, i.e., *the government*. As *consumers*, individuals and households, with their given income have to decide 'what to consume and how much to consume'. They have to make these decisions because consumers are, by nature, utility maximizers and consuming any commodity in any quantity does not maximize their gains, the satisfaction. As *producers*, firms, farms, factories, shopkeepers, banks, transporters, etc. have to choose 'what to produce, how much to produce and how to produce' because they too are gain maximizers and producing any commodity in any quantity by any technique will not maximize their gains (profits). As *labour*, they have to choose between alternative occupations and places of work because any occupation at any place will not maximize their earnings. Likewise, the *government* has to choose how to tax, whom to tax, how much to spend and how to spend so that social welfare is maximized at a given social cost. Economics as a social science studies how *people* make their choices.

It is this economic behaviour of the individuals, households, firms, government and the society as a whole which forms the *central theme* of economics as a social science. Thus, *economics is fundamentally the study of how people allocate their limited resources to produce and consume goods and services to satisfy their endless wants with the objective of maximizing their gains.*

---

### 1.3 PROBLEMS OF CHOICE AND SCARCITY

---

The need for making choice arises because of some *basic facts of economic life*. Let us look at the basic facts of human life in some detail and how they create the problem of choice-making.

**1. Human Wants, Desires and Aspirations are Limitless.** The history of human civilization bears evidence to the fact that human desire to consume more and more of better and better goods and services has ever since been increasing. For example, housing need has risen from a hut to luxury palace, and if possible, a house in space; the need for means of transportation has gone up from mule and camel to supersonic jet planes; demand for means of communication has risen from messengers and postal services to cell phones with camera; need for computational facility from manual calculation to superfast computers; and so on. For an individual, only the end of life brings the end to his/her needs. But for homo sapiens, needs and desires continue to grow endlessly.

*Human wants, desires and needs are endless in the sense that they go on increasing with increase in people's ability to satisfy them. The endlessness of human wants can be attributed to (i) people's insatiable desire to raise their standard of living, comforts and efficiency; (ii) human tendency to accumulate things beyond their present need; (iii) increase in knowledge about inventions and innovations of new goods and services with greater convenience, efficiency and serviceability; (iv) multiplicative nature of some want (e.g., buying a car creates want for many other things—petrol, driver, cleaning, parking place, safety locks, spare parts, insurance, etc.); (v) biological needs (e.g., food, water, etc.) are repetitive; (vi) imitative and competitive nature of human beings creating needs due to demonstration and bandwagon effects; and (vii) influence of advertisements in modern times creating new kind of wants. For these reasons, human wants continue to increase endlessly.*

Apart from being *unlimited*, another and an equally important feature of human wants is that *they are gradable*. In simple words, all human wants are not equally urgent and pressing, at a point time or over a period of time. While some wants have to be satisfied as and when they arise (e.g., food, clothes and shelter) and some can be postponed, e.g., purchase of a car. Also, while satisfying some wants gives a greater satisfaction than others. Given their intensity and urgency, human wants can be arranged in the order of their priority. The priority of wants, however, varies from person to person, and from time to time for the same person. Therefore the question arises as to 'which want to satisfy first' and 'which the last'. Thus, the consumer has to make choice 'what to consume' and 'how much to consume'. Economics studies how consumers (individuals and household) make choice between their wants and how they allocate their expenditure between different kinds of goods and services they choose to consume.

**2. Resources are Scarce.** The need for making choice between the various goods that people want to produce and consume arises mainly because *resources that*

are available to the people at any point of time for satisfying their wants are scarce and limited. What are the resources? Conceptually, any thing which is available and can be used to satisfy human wants and desire is a *resource*. In economics, however, *resources* that are available to individuals, households, firms, and societies at any point of time are traditionally classified as follows.

- (i) *natural resources* (including cultivable land surface, space, lakes, rivers, coastal range, minerals, wildlife, forest, climate, rainfall, etc.);
- (ii) *human resources* (including manpower, human energy, talent, professional skill, innovative ability and organizational skill, jointly called *labour*);
- (iii) *man-made resources* (including machinery, equipments, tools, technology and building, called together *capital*); and
- (iv) *Entrepreneurship*, i.e., the ability, knowledge and talent to put *land, labour* and *capital* in the process of production, and ability and willingness to assume *risk* in business.

To these basic resources, economists add other categories of resources, viz., *time, technology* and *information*. All these *resources are scarce*. Resource scarcity is a relative term. It implies that resources are scarce in relation to the demand for resources. The *scarcity of resources is the mother of all economic problems*. If resources were unlimited, like human wants, there would be no economic problem and, perhaps, no economics as a subject of study. *It is the scarcity of resources in relation to human wants that forces people to make choices*.

Furthermore, the problem of making choice arises also because resources have *alternative uses* and alternative uses have *different returns or earnings*. For example, a building can be used to set up a shopping center, business office, a 'public school', a hospital or for residential purpose. But the return on building varies from use to use of the building. Therefore, a return maximizing building owner has to make choice between the alternative uses of the building. If the building is put to a particular use, the landlord has to forego the return expected from its other alternative uses. This is called **opportunity cost**. Economics as a social science analyses how people (individuals and society) make their choices between the economic goals they want to achieve, between the goods and services they want to produce, and between the alternative uses of their resources with the objective of maximizing their gains. The gain maximizers evaluates the costs and benefits of the alternatives while deciding on the final use of the resources. Economics studies the process of making choices between the alternative uses. This is what constitutes, according to Robbins, the subject matter of economics.

**3. People are Gain Maximizers.** Yet another important aspect of human nature that leads to the choice-making behaviour is that most people aim at maximizing their gains from the use of their limited resources. 'Why people want to maximize their gains' is no concern of economics? Traditional economics assumes maximizing



behaviour of the people as a part of their rational economic behaviour. This assumption is based on observed facts. As *consumers*, they want to maximize their utility or satisfaction; as *producers*, they want to maximize their output or profit; and as *factor owners*, they want to maximize their earnings. *People's desire to maximize their gains* is a very important aspect of economic behaviour of the people giving rise to economics. If the people were not to maximize their gains, the problem of choice making would not arise. Consumers would not bother as to 'what to consume' and 'how much to consume'; producers would not bother as to 'what to produce', 'how much to produce' and 'how to produce'; and factor owners would not care as to where and how to use the resources. But, in reality, they do maximize their gains. Economics studies how people maximize their gains. Economics goes far beyond choice-making behaviour

The foregoing description of economics may give the impression that economics ends at the study of choice-making behaviour of the people. Not quite so. Robbins' definition of economics confines subject matter of economics to the study of economic behaviour of the people at its micro level. It may thus be said that Robbins' definition confines economics to what is now called *microeconomics*. But economics goes far beyond the scope of microeconomics. If economics is confined to the study of choice-making behaviour of the individual economic man, many other and more important economic issues that constitute a major part of modern economic science will have to be left out. Look at some of the major national and international economic issues.

- How is the level of output and employment determined in a country?
- Why are some countries very rich and some countries very poor?
- What are the factors that determine the overall economic growth of a country?
- What causes fluctuations in the national output, employment and the general price level?
- How do international trade and international capital flows affect the domestic economy?
- What causes inflation and what are its effects on economy's growth and employment?
- Why is about 35 per cent of India's population still 'below the poverty line' even after five decades of planned development with emphasis on 'removal of poverty'?
- Why is there large scale unemployment in India and why have the efforts to solve the problem of unemployment failed?
- Why has the Government of India been faced with fiscal deficits of a dangerous magnitude over the past two decades and why has it failed to reduce it to a manageable level?

- Why does the government need to intervene with the market system and adopt measures to control and regulate production and consumption, saving and investment, export and imports, wages and prices, and so on?

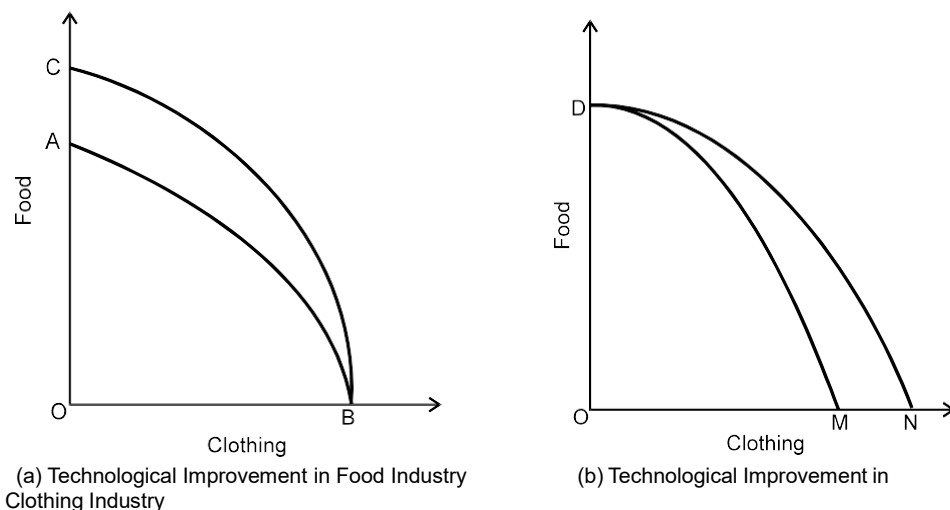
One can point out many other issues which do not fall within the purview of microeconomics. Analysis of and finding answer to such economic problems constitute now a major and also a more important subject matter of economics than the choice-making aspect of it. The study of the issues mentioned above has created a relatively new branch of economics, called *macroeconomics*. As noted above, the *scope of economics* continues to grow and expand in its scope, size and analytical rigour. Boundaries of economic science are not yet precisely marked though economics is claimed to be 'the oldest and best developed of the social sciences'. Let us now have a glance at the *scope of economics* as it is known today and its major branches and specialized areas of economics.

### *1.3.1 Basic Economic decisions*

The major economic problems faced by an economy—whether capitalist, socialist or mixed—may be classified in two broad groups:

- microeconomic problems which are related to the working of the economic system; and
- macroeconomic problems related to the growth, employment, stability, external balance, and macroeconomic policies for the management of the economy as a whole.

We will first discuss the microeconomic problems which are immediately relevant to our simplified economic system. Macroeconomic problems will be taken up in the following sub-section.



**Fig. 1.1** *Technological Improvement and Production Possibility Frontier*

The basic microeconomic problems are:

- (i) What to produce and how much to produce?
- (ii) How to produce?
- (iii) For whom to produce or how to distribute the social output.

These problems assume a macro nature when considered at the economy level. However, we will discuss them first at the micro level because these problems have to be resolved at the micro in the following sub-section.

*(i) What to Produce?*

The problem 'what to produce' is the problem of choice between commodities. This problem arises mainly for two reasons: (i) scarcity of resources does not permit production of all the goods and services that people would like to consume; and (ii) all the goods and services are not equally valued in terms of their *utility* by the consumers. Some commodities yield higher utility than the others. Since all the goods and services cannot be produced for lack of resources, and all that is produced may not be bought by the consumers, the problem of choice between the commodities arises. The problem 'what to produce' is essentially the problem of efficient allocation of scarce resources so that output is maximum and output-mix is optimum. The objective is to satisfy maximum needs of maximum number of people.

The question 'how much to produce' is the problem of determining the quantity of each commodity and service to be produced. This problem too arises due to scarcity of resources. For, surplus production would mean wastage of scarce resources. This problem also implies the allocation of resources between various goods and services to be produced.

The basic economic problem of unlimited wants and limited resources make it necessary for an economic system to devise some method of determining 'what to produce' and 'how much to produce', and ways and means to allocate the available resources for the production of goods and services. In a free enterprise economy, the solution to the problems 'what to produce' and 'how much to produce' is provided by the **price mechanism**.

*(ii) How to Produce?*

The problem 'how to produce' is the problem of **choice of technique**. Here the problem is how to determine an optimum combination of inputs—labour and capital—to be used in the production of goods or services. This problem too arises mainly because of scarcity of resources. If labour and capital were available in unlimited quantities, any amount of labour and capital could be combined to produce a commodity. But, since resources are scarce, it becomes imperative to choose a technology which uses resources most economically.

Another very important factor which gives rise to this problem is that a given quantity of a commodity can be produced with a number of alternative techniques, i.e., alternative input combinations. For example, it is always technically possible to produce a given quantity of wheat with more of labour and less of capital (i.e., with a labour-intensive technology) and with more of capital and less of labour (i.e., with a capital-intensive technology). The same is true of most commodities. In case of some commodities however choices are limited. For example, production of woollen carpets and other items of handicrafts are by nature labour-intensive, while production of cars, TV sets, computers, aircraft, etc. are capital-intensive. In case of most commodities, however, alternative technology may be available. But the alternative techniques of production involve varying costs. Therefore, the problem of choice of technology arises.

In a free market economy, the market system itself provides solution to the problem of choice of technology through price mechanism. The market mechanism yields a pricing system which determines the prices of both labour and capital. Factor prices and factor-quantities determine the cost of production for the business firms. Profit maximising firms find out an input combination which minimise their cost of production. This becomes inevitable for the firms because their resources are limited and, with given resources, they intend to maximise their profits.

The process through which business firms arrive at the optimum input combination and make choices between the alternative techniques of production are the topics in the 'Theory of Production' or 'Theory of Firms Behaviour' discussed later in the book.

### *(iii) For Whom to Produce: How to Distribute Social Output*

In a modern economy, all the goods and services are produced by the business firms. The total output generated by the business firms is known as 'society's total product' or 'national output'. The total output ultimately flows to the households. Here a question arises: how is the national output shared among the households or what determines the share of each household? A possible answer to this question is that, in a free enterprise economy, it is the price-mechanism which determines the distribution pattern of the national output. Price-mechanism determines the price of each factor in the factor market. Once factor price is determined, the income of each household is determined by the *quantity of the factor(s)* which it sells in the factor market. Those who possess a large amount of highly priced resources, are able to earn higher incomes and consume a larger proportion of national output than those who possess a small quantity of low-priced resources.

But the problem does not end here. For, other questions then arises: why some people have a command over larger proportion of resources than the others? Why those who have more, get more and more? Why those who have less, get less and less? In other words, why do rich get richer and poor get poorer? Is this distribution of national production fair? If not, how can disparities in incomes or sources of incomes be removed, or at least, reduced?

Price mechanism of free enterprise system has not been able to provide a solution to these questions. These problems have long been debated inconclusively. They remain alive today as they were during the days of Adam Smith and David Ricardo. These questions are the subject of the 'Theory of Distribution'.

When questions related to production and distribution are looked into from the *efficiency* point of view, the economists address themselves to other questions: How efficient is the society's production and distribution system? How does it affect welfare of the society? How can production and distribution be made more efficient or welfare oriented? Economists' attempt to answer these questions has led to growth of another branch of economics, i.e., *Welfare Economics*.

### Major Macroeconomic Problems

The economic problems discussed above are of micro nature. These problems taken together make the subject matter of Microeconomic Theory or 'Price Theory'. Apart from micro problems, there are certain macroeconomic problems of prime importance confronted by an economy. Following Lipsey, these problems may be specified as follows.

**1. How to Increase Production Capacity of the Economy.** This is essentially the problem related to economic growth of the country. The need for increasing production capacity of the economy arises for at least two reasons. *First*, most economies of the world have realised by experience that their population has grown at a rate much higher than their productive resources. This leads to the poverty especially in the less developed countries. Poverty in itself is a cause of a number of socio-economic problems. Besides, it has frequently jeopardised the sovereignty and integrity of the nations. Colonisation of poor nations by the richer and powerful imperialist nations during pre-twentieth century period is the evidence to this fact. Therefore, growth of economy and sparing resources for defence has become a necessity. *Secondly*, over time some economies have grown faster than others while some economies have remained almost stagnant. The poor nations have been subjected to exploitation and economic discrimination. This has impelled upon the poor nations to make their economies grow, to protect themselves from exploitation and to give their people a respectable status in the international community.

While various economies have been facing the problem of growth, economists have engaged themselves in finding an answer to such questions as: What makes an economy grow? Why some economies grow faster than the others? This has led to the growth of Theories of Economic Growth.

**2. How to Stabilise the Economy.** An important feature of the free enterprise system has been the economic fluctuation of these economies. Though economic ups and downs are not unknown in the controlled economies, free enterprise economies have experienced it more frequently and more severely. Economic fluctuations cause wastage of resources, e.g., idleness of manpower or involuntary

unemployment, idle capital stock, etc., particularly during the periods of depression. Economists have devoted a good deal of attention to explain this phenomenon. This problem is studied under Trade Cycles or Business Cycles.

**3. Other Problems of Macro Nature.** In addition to the macro problems mentioned above, there are many other economic problems of this nature, which economists have studied extensively and intensively. The most important problem of this category are the problems *unemployment* and inflation. While widespread unemployment is the biggest problem confronting the developing economies, inflation is a global problem. The abounding literature on these problems has yet to offer a solution to these problems. Another set of macro problems is associated with international trade. The major questions to which economists have devoted a good deal of their attention are: What is the basis of trade between the nations? How are the gains from trade shared between the nations? Why do deficits and surpluses arise in trade balances? How is an economy affected by deficits or surplus in its balance of payment position? New problems continue to emerge as economy passes through different phases of economic growth.

### *1.3.2 How the Market Mechanism Solves the Basic Problems*

The way basic problems of an economy are solved depends on the nature of its economic system. While in a socialist economy they are solved by the government agencies, like central planning authority, in a free enterprise or mixed capitalist economy, the basic economic problems are resolved by **Price Mechanism** or **Marked Mechanism**. We discuss here only how market mechanism solves the basic economic problems in a free enterprise or a mixed capitalist economy. For other economic systems, a brief answer is provided in the next section.

#### *What is Market Mechanism?*

Market Mechanism refers to a process through which market forces of demand and supply interact to determine the price and output of each good and service. A free market economy functions through the market forces of demand and supply. The demand and supply forces interact to determine the price of each good and service. In the process, a price system is generated. Prices perform two functions in the market system. One, prices serve as signals for the producers to decide 'what to produce' and for the consumers to decide 'what to consume'. Second, prices force the demand and supply conditions to adjust themselves to the prevailing prices. Let us now see how each of the basic problems is solved by the market mechanism or price mechanism.

#### *What to Produce*

The goods and services that are produced in a market economy are determined by the consumers' demand. Only those goods and services which are demanded by the consumers or users are produced by the producers. Each penny a consumer spends on a commodity is treated as a vote for producing that commodity.

Continuing demand is a continuous process of voting. Increasing demand for a good causes increase in its price. Rise in price makes profits to go up. The profit-seeking producers concentrate on the production of this commodity. If they produce a commodity not in demand, it will go waste and their profit motive will be defeated. The consumer is thus 'sovereign' in a free enterprise market economy—consumer determines 'what to produce'.

### *How to Produce*

'How to Produce' is the question of choice of technology. The proportion in which labour and capital are combined to produce a commodity is also determined by the market forces, i.e., the supply of and demand for labour and capital. Firms produce for profit and try to maximise it. It requires, among other things, minimising cost of production. Costs can be minimised by using more of a cheap factor and less of a costly factor. If labour is cheaper than capital then more of labour and less of capital is used to produce a commodity. On the contrary, if capital is cheaper, more of capital and less of labour is used.

What makes a factor cheaper or costlier? It depends on the supply of and demand for that factor. If supply of a factor exceeds its demand, price of that factor will be lower, and the factor will be treated as a cheaper factor. But if demand for a factor exceeds its supply, the price of that factor will be high and the factor will be treated as a costly factor. Given the factor prices, firms combine labour and capital in such proportions that minimize cost of production. This determines the production technology. This is how market forces offer a solution to the problem 'how to produce'.

### *For Whom to Produce*

The problem 'For Whom to Produce' is also solved by the market mechanism. The simple market rule is: produce for those who have ability and willingness to pay. Ability to pay depends on incomes and incomes are determined by employment pattern of factors. Market mechanism determines the pattern of demand for factors of production. Given the supply of factors, market mechanism determines the price of each factor—rent, wages, interest and profits, respectively, for land, labour, capital and organisation. Once factor prices and employment pattern of factors of production (i.e., what factor is employed in *what quantity and at what price*) are determined, the distribution pattern of national income is simultaneously determined. In simple words, employment pattern determines the share of labour, property owners, investors and entrepreneurs in the national income. Once the pattern of income distribution is determined, it determines the demand pattern for the goods and services, for there is a relationship between income and consumption pattern. Thus, in a free enterprise economy, goods and services are produced for those who possess the ability to pay.

---

## 1.4 THE SCOPE OF ECONOMICS

---

The scope of economics is not marked precisely and, as it appears, it cannot be. However, the scope of economics, as it is known today, has expanded vastly in the post-War period. *Modern economics* is now divided into two major branches: *Microeconomics* and *Macroeconomics*. A brief description of the subject matter and approach of microeconomics and macroeconomics follows.

### 1.4.1 Distinction between Micro and Macro Economics

*Microeconomics* is concerned with microscopic study of the various elements of the economic system and not with the system as a whole. As Lerner has put it, “Microeconomics consists of looking at the economy through a microscope, as it were, to see how the million of cells in body economic—the individuals or households as consumers and the individuals or firms as producers—play their part in the working of the whole economic organism”. Thus, micro-economics is the study of the economic behaviour of individual consumer and producer and of individual economic variables, i.e., production and pricing of individual goods and services. Microeconomics studies how consumers and producers make their choices; how their decisions and choices affect the market demand and supply conditions; how consumers and producers interact to settle the prices of goods and services in the market; how prices are determined in different market settings; and how total output is distributed among those who contribute to production, i.e., between landlords, labour, capital supplier, and the entrepreneurs. Briefly speaking, *theory of consumer behaviour, theories of production and cost of production, theory of commodity and factor pricing, efficient allocation of output and factors of production (called welfare economics)* constitute the main themes of *microeconomics*.

### *Macroeconomics*

Macroeconomics is a relatively new branch of economics. It was only after the publication of Keynes’s *The General Theory of Employment, Interest and Money* in 1936, that macroeconomics crystallized as a separate branch of economics. *Macroeconomics* studies the working and performance of the economy as a whole. It analyses behaviour of the national aggregates including national income, aggregate consumption, savings, investment, total employment, the general price level and country’s balance of payments. According to Boulding, “Macroeconomics is the study of the nature, relationship and behaviour of aggregates and averages of economic quantities.” He contrasts macroeconomics with microeconomics in the following words: “Macroeconomics . . . deals not with individual quantities, as such, but aggregates of these quantities—not with individual incomes, but with the national income, not with individual prices but with price levels, not with individual output but with the national output.” More importantly, macroeconomics analyses relationship between the national aggregate variables and how aggregate variables



interact with one another to determine one another. It studies also the impact of public revenue and public expenditure, government's economic activities and policies on the economy. An important aspect of macroeconomics studies is the consequences of international trade and other economic relations between the nations. The study of these aspects of economic phenomena constitutes the major themes of *macroeconomics*.

---

## **1.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS**

---

1. According to Alfred Marshall (1922), "Economics is the study of mankind in the ordinary business of life; it examines that part of individual and social action which is most closely connected with the attainment and with the use of the material requisites of wellbeing".
2. Economic behaviour is essentially the process of evaluating economic opportunities open to an individual or a society and, given the resources, making choice of the best of the opportunities.
3. The problem of 'what to produce' arises mainly for two reasons: (i) scarcity of resources does not permit production of all the goods and services that people would like to consume; and (ii) all the goods and services are not equally valued in terms of their utility by the consumers.
4. Market Mechanism refers to a process through which market forces of demand and supply interact to determine the price and output of each good and service.

---

## **1.6 SUMMARY**

---

- Economists right from Adam Smith—the 'father of economics'—down to modern economists have defined economics differently depending on their own perception of the subject matter of economics of their era.
- According to Alfred Marshall (1922), "Economics is the study of mankind in the ordinary business of life; it examines that part of individual and social action which is most closely connected with the attainment and with the use of the material requisites of wellbeing".
- *Economics as a social science studies economic behaviour of the people and its consequences.*

- *Economic behaviour* is essentially the process of evaluating economic opportunities open to an individual or a society and, given the resources, making choice of the best of the opportunities.
- The objective behind this economic behaviour is to maximize gains from the available resources and opportunities.
- The basic function of economics is to observe, explain and predict how people (individuals, households, firms and the government) as decision-makers make choices about the use of their resources (land, labour, capital, knowledge and skills, technology, time and space, etc.) to maximize their income, and how they as consumers decide how to spend the income to maximize their total utility.
- For the purpose of economic analysis, people are classified according to their decision-making capacity as *individuals, households, firms* and the *society*, and according to the nature of their economic activity as *consumers, producers, factor owners* and *economy managers*, i.e., *the government*.
- *Human wants*, desires and needs are endless in the sense that they go on increasing with increase in people's ability to satisfy them.
- Apart from being *unlimited*, another and an equally important feature of human wants is that *they are gradable*.
- In a free enterprise economy, the solution to the problems 'what to produce' and 'how much to produce' is provided by the price mechanism.
- The goods and services that are produced in a market economy are determined by the consumers' demand.
- *Microeconomics* is concerned with microscopic study of the various elements of the economic system and not with the system as a whole.
- Macroeconomics is a relatively new branch of economics. It was only after the publication of Keynes's *The General Theory of Employment, Interest and Money* in 1936, that macroeconomics crystallized as a separate branch of economics.
- *Macroeconomics* studies the working and performance of the economy as a whole.

---

## 1.7 KEY WORDS

---

- **Production:** It refers to the action of making or manufacturing from components or raw materials, or the process of being so manufactured.
- **Employment:** It refers to an activity or service performed for another especially for compensation or as an occupation.
- **Economy:** It is the state of a country or region in terms of the production and consumption of goods and services and the supply of money.

---

## 1.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

### Short-Answer Questions

1. What are the factors attributing to the endlessness of human wants?
2. How are major economic problems faced by an economy grouped?
3. State the differences between micro and macro economics.

### Long-Answer Questions

1. Economics is a social science. Critically comment and elucidate.
2. What gives rise to the problems of choice and scarcity? Discuss all basic facts as given in the unit.
3. Discuss all microeconomic and macroeconomic problems.
4. Discuss the meaning of market mechanism. How does it solve the basic problems?

---

## 1.9 FURTHER READINGS

---

Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.

Keat Paul G. and K.Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.

Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition. Singapore: Pearson Education, Inc.

Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Publishers.

---

### Structure

- 2.0 Introduction
- 2.1 Objectives
- 2.2 Approaches to Economic Analysis: Micro and Macro Analysis
- 2.3 Deductive and Inductive Methods of Analysis
- 2.4 Economic Statics and Dynamics
- 2.5 Answers to Check Your Progress Questions
- 2.6 Summary
- 2.7 Key Words
- 2.8 Self Assessment Questions and Exercises
- 2.9 Further Readings

---

## 2.0 INTRODUCTION

---

In this unit, you will study the various methods of economic analysis. By definition, economic methodology is the study of methods, especially the scientific method, in relation to economics, including principles underlying economic reasoning. An economic theory derives laws or generalizations through two methods: (1) Deductive method and (2) Inductive method. These methods of analysis have been discussed in detail in this unit.

---

## 2.1 OBJECTIVES

---

After going through this unit, you will be able to:

- Explain micro and macro approaches to economic analysis
- Discuss deductive and inductive methods of analysis
- Describe economic statics and dynamics

---

## 2.2 APPROACHES TO ECONOMIC ANALYSIS: MICRO AND MACRO ANALYSIS

---

There are two approaches to economic analysis—micro and macro approaches—leading to Microeconomics and Macroeconomics which refer to the two branches of economics. Microeconomics deals chiefly with the choice and decision-making behaviour of the individual households, firms and industries, and the relationship between prices and quantities of individual goods and services. On the other hand, macroeconomics is chiefly concerned with the behaviour of and relationships

between the national aggregate such as national income or output, total consumption, saving and investment, total employment, total money supply and general price level, etc.

The two terms 'microeconomics' and 'macroeconomics' were first coined and used by Ragnar Fisch, a German economist in 1933. The prefixes, 'micro' and 'macro' have been derived from the Greek words *mikros* meaning 'small' and *makros* meaning 'large'. The terms 'microeconomics' and 'macroeconomics' are now used to denote the two branches of economics. These terms also connote the two kinds of approaches adopted in economic studies, according to whether a part or the whole of the economic system is converted under economic investigations.

Let us now look at the two approaches of economic analysis in detail.

### 1. Microeconomics Analysis

Microeconomics studies the economic behaviour of individual economic entities and individual economic variables. The economic entities may be individuals or small groups of individuals. The small group of individuals may be household firms and industries consisting of several firms. Thus, microeconomics is the study of such economic units as individual firms and households, individual prices, wages, incomes, individual industries and individual commodities. It is concerned with the microscopic study of various elements of economic systems and not with the system as a whole. As Lerner puts it, 'Microeconomics consists of looking at the economy through a microscope, as it were, to see how the millions of cells in the body economic—the individuals or households as consumers, and the individuals or firms as producers—play their part in the working of the whole economic organism.' The individuals or households as consumers, allocate their given incomes between various goods and services they consume so as to maximize their total satisfaction. The individuals or firms as producers play their part in the working of economic organisms insofar as they allocate their resources to the production of various goods and services they choose to produce, determine their prices and output so as to maximize their profits. Such parts played by individuals or groups of individuals form the subject matter of microeconomics. Microeconomics seeks to explain *how* an individual consumer distributes his disposable income among various goods and services, *how* he attains the level of maximum satisfaction, and *how* he reaches the point of equilibrium. Microeconomics is also concerned with *how* individual firms decide 'what to produce', 'how to produce', and 'at what cost to produce', to minimize the cost of production.

In other words, microeconomics seeks to analyse the mechanism by which various economic units attain the position of equilibrium, the position from which they could not like to deviate, the conditions remaining the same.

Microeconomic analysis assumes that the quantity of national resources, national income, level of aggregate consumption, savings and investment and employment are given. Under the assumption of a given structural format of the economy, microeconomics also examines whether resources are *efficiently* allocated and spells out conditions for Pareto-optimality in allocation of resources between the various economic units so as to maximize output and social welfare. In the opinion of Lerner, economists are concerned in microeconomics, with the avoidance or elimination of waste or with inefficiency arising from the inoptimal allocation of resources. Inefficiency can be removed by reallocating resources between consumers and producers so as to get more of something that is scarce without sacrificing other scarce goods or to replace a commodity by a more preferable one. It also spells out the conditions for the most efficient allocation of resources, i.e., to achieve Pareto-optimality in resource allocation. To be specific, the theory of consumers' behaviour, theory of firms or theory of production, theory of product pricing, theory of factor pricing (or distribution theory) and the theory of economic welfare constitute the body of microeconomics.

### *Limitations of the Microeconomic Approach*

Despite the many uses of microeconomics, it has certain limitations which restrict its practicability. Most limitations of microeconomic theories arise out of the *assumptions* on which they are based. Some major assumptions and the resultant limitations of microeconomic theories are following:

*First*, microeconomic theories assume a given level of national income, employment, saving and investment. In reality, however, these factors are not constant; they are subject to change following changes in their determinants. As such, the validity of microeconomic theories is doubtful.

*Second*, microeconomic theories assume the existence of a free enterprise system in which the 'invisible hands' or market forces are assumed to play their roles freely. It also assumes the absence of any government intervention in the economic activities of the society. In practice, however, government controls and regulations of economic activities are the rules of the day, and are all-pervasive. Therefore, microeconomic propositions have only limited applicability—limited to the conditions assumed for the microeconomic models.

*Third*, another limitation of microeconomics arises out of its very scope of study. It is concerned with the behaviour of individual elements of the economic organism and not with the organism as a whole. Microeconomic theories, therefore, cannot be applied to study the complex economic systems treated as one unit. As Boulding says, 'Description of a large and complex universe of facts like economic system is impossible in terms of individual items' whereas microeconomics is concerned with only 'individual items'.

These limitations of microeconomic theory however do not reduce its importance. There are important and practical reasons for studying and making

use of microeconomic theory. Emphasizing the importance of the 'price theory', the relatively older name of 'microeconomics', Liebhafsky has said, '... *there is a very practical reason for acquiring a knowledge of price theory; the language and concepts of price theory permeate the whole of economics, and in all field of economic analysis they serve the practical purpose of economy of effort and constitute a generally accepted method of regaining and classifying ideas about economic activities and magnitudes.*'

## 2. Macroeconomic Analysis

In contrast to microeconomics, macroeconomics is concerned with the nature, relationships and behaviour of such aggregate quantities and averages as national income, total consumption, savings and investment, total employment, general price level, aggregate expenditure and aggregate supply of goods and services. Clearly, the variables which microeconomics assume to be given, constitute the subject of macroeconomics. Since macroeconomics deals with aggregate quantities of the economy as a whole, it is sometimes also called as *aggregate economics*. As Boulding has defined it, 'Microeconomics is the study of the nature, relationship and behaviour of aggregates and averages of economic quantities.' Thus, 'Macroeconomics...deals not with individual quantities as such, but aggregates of these quantities; not with individual incomes, but with the national income; not with individual prices but with price levels, not with the national income, not with individual prices but with price levels, not with individual output but with the national output.' Ackley has defined it almost identically as 'Macroeconomics concerns itself with such variables as the aggregate volume of the output of an economy, with the extent to which its resources are employed, with the size of national income, with the general price level.'

As Shapiro has pointed out, none of these definitions satisfactorily define macroeconomics. But these and many others put together do give an idea about the nature and scope of macroeconomics. In fact, macroeconomic theories seek to answer such questions as, 'How is the level of national income of a country determined?' 'What determines the levels of overall economic activities in a country?' 'What determines the level of total employment?' 'How is the general level of price determined?' 'What causes fluctuations in national output, employment, and general price level?' 'What determines the level of foreign trade and what factors contribute to the growth of surplus and deficits in the balance of payments position of a country?' The branch of economic science which seeks to answer such questions has been termed as 'macroeconomics'. Briefly speaking, theories of national income, consumption, saving and investment, theory of employment, theories of economic growth, business cycles and stabilization policies, theories of money supply and demand, and theory of foreign trade broadly constitute the subject matter of macroeconomics. Thus, macroeconomics studies the economic system as a whole.

To study the system as a whole and to explain the behaviour of aggregate quantities and the relationship between them requires identifying and measuring the forces which are, by turn, both the cause and effect of the functioning of the economic system. But this is extremely difficult. For, millions of individuals contribute to the national aggregates in thousands of ways and forms and to measure the effects of the activities of each individual is rather impossible; no meaningful generalization regarding the economic aggregates can be made. The macroeconomic approach has made this possible. It ignores the details pertaining to the individual economic agents and quantities, and compresses the unmanageable multitudes of economic facts to a manageable size and makes them capable of interpretation.

Another important justification of macroeconomics lies in its use in formulating public policies. Due to the failure of the market system to provide solution to the economic problems confronting the nations, the economic role of the government and its interference with the economic life of the people has in the recent past increased. The main objective of government activities and programmes is to control and regulate the economic system of the country so as to achieve the desired goals and fulfil the aspirations of the society. For all these, the government has to formulate its economic policies in accordance with the expected behaviour of the economic system as a whole. Macroeconomic theories which are concerned with the economic system as a whole are of immense use in policy formulation. They provide clarity to macroeconomic concepts and quantities, and bring out the relationships between the macro variables of the economy in the form of 'models' or equations.

### *Limitations of Macroeconomic Analysis*

The *first* limitation of macroeconomics arises out of its very nature. The applicability of macroeconomics is limited to only aggregates, as it is the theory of aggregates. It cannot be applied to explain the behaviour of individual components of the economic system and the individual quantities. This is because of the *fallacy of aggregate*, i.e., what is true of aggregate may not be true of individuals which constitute the aggregate. For example, even if the aggregate national output is increasing, the output in some sectors, in some regions, in some individual industries, and factories may be decreasing simultaneously. This is true of consumption, saving, investment and employment.

*Second*, another fallacy of aggregate is that it ignores the structural changes in constituent elements of the aggregate. Hence, conclusions drawn from the analysis of aggregates may involve error of judgement and may be misleading. For example, growth in national income over time might lead to the conclusion that the economy is performing well. But, this growth might be accompanied by such structural changes as excessive substitution of capital for labour and transfer of income from poor to the rich, rise in unemployment, and so on. But growth in national income under such conditions would not be considered a healthy sign for the economy. Similarly, a macro study of aggregate prices may lead to the conclusion that the general price level has remained constant over time, even if agricultural prices



were decreasing and industrial prices were increasing so that on an average, the price level remained constant. But such a structural change in prices needs corrective measures otherwise it may lead to general depression through fall in agricultural incomes and the subsequent fall in purchasing power.

*Third*, macroeconomics deals with aggregates, and ‘aggregates are not reality but a picture or approximation of reality’. Rather, the individual components which constitute the aggregate, are the reality. But individual components are heterogeneous, not homogeneous. Thus, aggregation of heterogeneous components does not make any sense, like adding 5 buckets of milk to 10 yards of cloth, or adding a thousand houses of a city to millions of spoons used by the city dwellers does not make any sense. It does not produce a meaningful quantity. This problem has, however, been solved by reducing all the heterogeneous components to money values by using money as a common denominator in the macroeconomic approach.

---

### 2.3 DEDUCTIVE AND INDUCTIVE METHODS OF ANALYSIS

---

Economic theories which constitute the body of economic science today are the result of scientific investigation into economic facts. The scientific search for economic truths consists of a systematic and logical procedure of arranging and analysing economic facts and establishing the relationship between the facts.

The two kinds of methods which have been adopted at different stages of growth of economic science by different schools of thought are : (a) *Deductive Method* : and (b) *Inductive Method*. Deduction and induction are, in fact, two different forms of logic which are used to draw inferences. In the *deductive method*, reasoning proceeds from general to particular or from universal to individuals. In this method, inferences are drawn from general cases to establish the particular case. On the other hand, in the *inductive method*, reasoning proceeds from particular to the general or from individual to the universal, and a general case is made from the individual cases.

Let us now discuss the two methods in detail.

#### 1. *Deductive Method*

The deductive method is also known as the *analytical method*. In the deductive or analytical method, initially certain assumptions or postulates are made. On the basis of these assumptions, certain logical conclusions are drawn which become the testable hypotheses. The hypotheses are then tested against observed facts. The hypotheses confirmed by the facts are accepted as tentative theories. If a theory so formulated stands the tests, time and again, it becomes a law, e.g., law of demand, law of diminishing marginal utility, etc.

The deductive approach proceeds by the following major steps: (a) selecting the problem for analysis; (b) specifying the assumptions or postulates; (c)

formulating hypotheses on the basis of assumptions; and (d) testing the validity of the hypotheses.

The *first* step in any scientific analysis is *to specify the problem of the study*. The problem chosen for the study generally is, and should be, of practical importance to the society. This, however, is not necessary. Economists may, and in fact they do, select a problem of their own interest which may not serve any immediate social purpose. In fact, most early scientific discoveries have been the result of a scientist's own curiosity rather than the result of research undertaken to serve certain predetermined social ends.

The *second* step is to specify the *assumptions*. Assumptions serve several purposes in scientific analysis. They determine the scope and dimensions of the study and also specify the factors to be taken as constant. Assumptions are essentially used as the premise on which hypotheses are built.

Despite its merits, the deductive method has certain *disadvantages*. *First*, although it is claimed to be a simple method, it is a highly complicated method as it requires great skill and logical acumen to derive conclusions from the postulates. *Second*, since this method often leads to a high degree of abstraction, it involves the risk of yielding theories far from the reality. *Third*, the theories formulated are of limited applicability. That is, they are applicable within the framework of the assumptions which are often unreal. *Finally*, the deductive approach very often turns to be a mere *intellectual exercise* yielding results of little practical use.

## 2. Inductive Method

The inductive approach to formulating economic principles is the reverse of the deductive approach. While the deductive method is a descending process which proceeds from general to particular, the inductive method is an ascending process which proceeds from particular to general. Inductive analysis begins with observed facts regarding the recurrence of an economic event or existence of an economic phenomenon and its causes. It then establishes the cause-and-effect relationship between the events, making a general case. The general case is then used to explain individual economic events. For example, people have observed over centuries that when crops are damaged by flood, drought or inclement weather, agricultural production falls and agricultural prices go up. This makes a general case of price behaviour in response to change in supply. This general case may be applied to explain or predict the price behaviour in case of a particular crop, i.e., how price will behave given the supply position.

The inductive method involves the following steps:

The *first* step in formulating inductive economic laws is the same as in the case of the deductive approach, i.e., the selection of the problem for analysis. It may be any economic problem, such as returns to increasing inputs, unemployment, inflation, industrial unrest, etc.

The *second* step is collection, classification and analysis of data by using appropriate statistical techniques in order to find out the relationship between the variables.

The *final* step is to find out the reasons for the relationship established through statistical analysis and to set the rules for the verification of the principle.

Like the deductive approach, the inductive approach too has its own *merits* and *demerits*. As regards its *merits*, since this approach analyses economic phenomena on the basis of observed facts, it has been claimed to be more close to reality. Besides, inductive method is considered as the most important way of testing or verifying an established economic theory. Further, economic theories based on inductive approach can be a better tool of predicting future course of economic events.

*Demerits* of the inductive method lie in the problems of data and the statistical tools of analysis which are frequently used in this method. Collecting appropriate, requisite data on a particular economic problem is in itself a difficult task, particularly where conceptual problems are involved. Data-related problems arise mainly because experiments in an economic phenomenon is not possible in the same manner as in natural sciences. Besides, different investigators may arrive at different conclusions, from the same data, if their assumptions differ. This makes the conclusions doubtful. Further, in the inductive approach, 'there can be no absolute assurance that the result of the generalisation will actually be attained in a particular case.'

The inductive generalizations, therefore, turn to be merely statements of tendencies, not even testable hypotheses.

---

## 2.4 ECONOMIC STATICS AND DYNAMICS

---

A **static economy** is one in which nothing is changing. It is like studying a "photograph" or movie "still" in which all objects are motionless. In the context of economic analysis, it is assumed that nothing is changing in the economy; the types and number of economic organisations remain the same; people's taste and preferences do not change; all prices are constant; the output and employment are fixed; the quantity of goods and services that are produced and consumed remain unchanged—all at a given point of time. The equilibrium of an economy under these static conditions is called **Static Equilibrium**.

Close to static equilibrium is the concept of **Stationary Equilibrium**. When an economy remains in static equilibrium between two points of time, it is said to be **stationary**. If the economy today is exactly what it was yesterday or over any period in the past, it is in the stationary state "in which nothing ever happens".

A **Dynamic Economy** may be defined as one which is changing continuously. The change may be autonomous or induced. A dynamic economy is

in the process of continuous change. Such an economy passes through the stages of equilibrium and disequilibrium. When a dynamic economy is in the state of equilibrium, it is said to be **Dynamic Equilibrium**.

## 2.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. There are two approaches to economic analysis namely micro and macro approaches.
2. The *first* limitation of macroeconomics arises out of its very nature. The applicability of macroeconomics is limited to only aggregates, as it is the theory of aggregates.
3. The objective of economic analysis is not to formulate exact and precise economic laws but to develop the understanding of the economic system.

## 2.6 SUMMARY

- There are two approaches to economic analysis—micro and macro approaches—leading to Microeconomics and Macroeconomics which refer to the two branches of economics.
- The two terms ‘microeconomics’ and ‘macroeconomics’ were first coined and used by Ragnar Fisch, a German economist in 1933.
- Microeconomics studies the economic behaviour of individual economic entities and individual economic variables. The economic entities may be individuals or small groups of individuals.
- Microeconomic analysis assumes that the quantity of national resources, national income, level of aggregate consumption, savings and investment and employment are given.
- Despite the many uses of microeconomics, it has certain limitations which restrict its practicability. Most limitations of microeconomic theories arise out of the *assumptions* on which they are based.
- In contrast to microeconomics, macroeconomics is concerned with the nature, relationships and behaviour of such aggregate quantities and averages

as national income, total consumption, savings and investment, total employment, general price level, aggregate expenditure and aggregate supply of goods and services.

- Economic theories which constitute the body of economic science today are the result of scientific investigation into economic facts.
- The two kinds of methods which have been adopted at different stages of growth of economic science by different schools of thought are : (a) *Deductive Method* : and (b) *Inductive Method*.
- The deductive method is also known as the *analytical method*. In the deductive or analytical method, initially certain assumptions or postulates are made.
- The inductive approach to formulating economic principles is the reverse of the deductive approach.
- While the deductive method is a descending process which proceeds from general to particular, the inductive method is an ascending process which proceeds from particular to general.
- A **static economy** is one in which nothing is changing. It is like studying a “photograph” or movie “still” in which all objects are motionless.

---

## 2.7 KEY WORDS

---

- **Consumption:** It is the use of goods and services by households.
- **Hypothetical:** It means to be based on or serving as a hypothesis.
- **Equilibrium:** It is a state in which opposing forces or influences are balanced.

---

## 2.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

### *Short-Answer Questions*

1. What are the assumption of microeconomic analysis?
2. What are the limitations of the microeconomic approach?
3. Write a brief note on inductive method.

### *Long-Answer Questions*

1. Distinguish between microeconomic and macroeconomic approaches to economic analysis. Elucidate on differences or similarities.

2. What do you understand by deductive method? What are the steps involved?
3. How is a static economy different from a dynamic economy?

---

## 2.9 FURTHER READINGS

---

- Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.
- Keat Paul G. and K.Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.
- Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition. Singapore: Pearson Education, Inc.
- Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.
- Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Publishers.

---

**BLOCK - II**
**CONSUMER BEHAVIOUR AND DEMAND ANALYSIS**


---

**UNIT 3 CONSUMER BEHAVIOUR**


---

*Structure*

- 3.0 Introduction
- 3.1 Objectives
- 3.2 Cardinal and Ordinal Utility
- 3.3 Cardinal Utility Theory
  - 3.3.1 Law of Diminishing Marginal Utility
- 3.4 Consumer Equilibrium and The Law of Equi-Marginal Utility
- 3.5 Derivation of Demand Curve (Cardinal Utility Approach)
  - 3.5.1 Drawbacks of Cardinal Approach
- 3.6 Ordinal Utility Theory
  - 3.6.1 The Diminishing Marginal Rate of Substitution
- 3.7 Comparison of Cardinal and Ordinal Utility Approach
- 3.8 Answers to Check Your Progress Questions
- 3.9 Summary
- 3.10 Key Words
- 3.11 Self Assessment Questions and Exercises
- 3.12 Further Readings

---

**3.0 INTRODUCTION**


---

The cardinal and ordinal concepts of utility arise out of question whether 'utility is measurable'. Utility is a psychological phenomenon. It is a feeling of satisfaction, pleasure or happiness. Measurability of utility has, however, been a contentious issue. Classical economists, viz., Jeremy Bentham, Leon Walrus, Carl Menger, etc. and neo-classical economists, notably Alfred Marshall, believed that utility is cardinally or quantitatively measurable like height, weight, length, temperature and air pressure. This belief resulted in the *Cardinal Utility* concept. However, modern economists, most notably J.R. Hicks and R.G.D. Allen, hold the view that utility is not quantitatively measurable—it is not measurable in absolute terms. Utility can be expressed only ordinally or in terms of 'less than' or 'more than'. It is, therefore, possible to list the goods and services in order of their preferability or desirability. For example, suppose a person prefers chocolate to ice cream and ice cream to cold drink. He or she can express his/her preference as chocolate > ice cream > cold drink. This is known as the *ordinal* concept of utility.

This unit discusses the origin of the two concepts of utility and their use in the analysis of demand. The law of diminishing marginal utility and the law of equi-marginal utility are also explained in this unit.

---

### 3.1 OBJECTIVES

---

After going through this unit, you will be able to:

- Define cardinal and ordinal utility
- Explain the law of diminishing marginal utility
- Describe the law of equi-marginal utility

---

### 3.2 CARDINAL AND ORDINAL UTILITY

---

Let us discuss the origin of the two concepts of utility and their use in the analysis of demand.

#### *Cardinal Utility*

Some early psychological experiments on an individual's responses to various stimuli led neo-classical economists to believe that utility is measurable and cardinally quantifiable. According to neo-classical economists, utility can be measured in terms of money. That is, utility of a unit of a commodity for a person is equal to the amount of money he is willing to pay for it. This belief gave rise to the concept of cardinal utility. It implies that utility can be assigned a cardinal number like 1, 2, 3, etc. Neo-classical economists built up the theory of consumption on the assumption that utility is cardinally measurable. They coined and used a term 'util' meaning 'units of utility'. In their measure of utility, they assumed (i) that one 'util' equals one unit of money, and (ii) that utility of money remains constant.

It has, however, been realized over time that *absolute* or cardinal measurement of utility is not possible. Difficulties in measuring utility have proved to be insurmountable. Neither economists nor scientists have succeeded in devising a technique or an instrument for measuring the feeling of satisfaction, i.e., utility. Nor could an appropriate measure of unit be devised. Numerous factors affect the state of consumer's mood, which are impossible to determine and quantify. *Utility is therefore not measurable in cardinal terms.* Nevertheless, cardinal utility concept continues to remain the basis of the analysis of consumer behaviour.

#### *Ordinal Utility*

The modern economists have discarded the concept of *cardinal utility* and have instead employed the concept of *ordinal utility* for analyzing consumer behaviour. The concept of *ordinal utility* is based on the fact that it may not be possible for consumers to express the utility of a commodity in absolute or quantitative terms, but it is always possible for a consumer to tell introspectively whether a commodity is more or less or equally useful compared to another. For example, a consumer may not be able to say that ice cream gives 5 utils and chocolate gives 10 utils. But



he or she can always specify whether chocolate gives more or less utility than ice cream. This assumption forms the basis of the ordinal theory of consumer behaviour.

While neo-classical economists maintained that cardinal measurement of utility is practically possible and is meaningful in consumer analysis, modern economists maintain that utility being a psychological phenomenon is inherently immeasurable, theoretically, conceptually as well as quantitatively. They also maintain that the concept of ordinal utility is a feasible concept and it meets the conceptual requirement of analyzing consumer behaviour in the absence of any cardinal measures of utility.

### *Two approaches to the consumer demand analysis*

Based on cardinal and ordinal concepts of utility, there are two approaches to the analysis of consumer behaviour.

- (i) **Cardinal Utility Approach**, attributed to Alfred Marshall and his followers, is also called the Neo-classical Approach.
- (ii) **Ordinal Utility Approach**, pioneered by J.R. Hicks, a Nobel laureate, and R.G.D. Allen, is also called the Indifference Curve Analysis.

The two approaches are not in conflict with one another. In fact, they represent two levels of sophistication in the analysis of consumer behaviour. Both the approaches are important for assessing and analyzing consumer demand for a commodity—be it for theoretical purpose or for business decision-making, depending on the level of sophistication required.

It is important to note in this regard that in spite of tremendous developments in consumption theory based on ordinal utility, the classical demand theory based on cardinal utility has retained its appeal and applicability to the analysis of consumer behaviour. Besides, the study of classical demand theory serves as a foundation for understanding the advanced theories of consumer behaviour. The study of classical theory of demand is of particular importance and contributes a great deal in managerial decisions.

---

## **3.3 CARDINAL UTILITY THEORY**

---

The Cardinal Utility Theory was developed by classical economists, viz., Gossen (1854) of Germany, William Stanley Jevons (1871) of England, Leon Walras (1874) of France, Karl Menger (1840–1921) of Austria. Neo-classical economists, particularly Alfred Marshall (1890) made significant refinements in the Cardinal Utility Theory. This led the Cardinal Utility Theory to be known as ‘Neo-classical Utility Theory’ or ‘Marshallian Utility Theory’ of demand.

Before we proceed to describe the Cardinal Utility Theory, let us first explain the basic concepts and axioms used in this theory.

### (a) The Meaning of Utility

The notion of “Utility” was introduced to social thought by the British philosopher, Jeremy Bentham, in the 18th century and to economics by William Stanley Jevons in the 19th century. In its economic meaning, the term “utility” is synonymous with “pleasure”, “satisfaction” and a sense of fulfilment by desire. A person consumes a commodity because he or she derives pleasure out of it. In other words, he derives utility from the consumption of the goods and services.

In abstract sense, the term “*utility*” refers to the power or property of a commodity to satisfy human needs. For example, bread has the power to satisfy hunger; water quenches our thirst; books fulfill our desire for knowledge; and postal stamps take our letters to their destination, and so on. All the goods that people hold or consume possess utility. Utility can also be defined as the “want-satisfying power” of a commodity. But it is not absolute—it is relative. It is relative to a person’s need. In other words, whether a commodity possesses utility depends on whether a person needs that commodity. All the persons need not derive utility from all the commodities. For example, non-smokers do not derive any utility from cigarettes; strict vegetarians do not derive any utility from meat and chicken; a book on economics has no utility for those who are not students of economics, and so on. The utility derived by a person from a commodity depends on his or her intensity of desire for that commodity: *the greater the need, the greater the utility*.

Besides, utility of some commodities depends on the availability of complementary goods. For example, electricity operated gadgets (e.g., TV, VCR, computers, refrigerators, etc.) yield utility only where electricity is available and petrol has utility only for those who possess an automobile.

Furthermore, the concept of utility is “**ethically neutral**”. It is neutral between good and bad and between useful and harmful. For example, some drugs are bad and harmful, for every body but they yield utility to the drug-addicts. Utility is free from moral values. It is not subject to social desirability of consuming a commodity. Eating beef may be immoral or socially undesirable for Hindus, but if a Hindu takes it, it satisfies his hunger.

### (b) Measurability of Utility

Measurability of utility has been and remains a debatable issue. Essentially, utility is a psychological phenomenon—it is a feeling of pleasure or a feeling of satisfaction and achievement. Can utility be measured in **absolute terms**? As mentioned above, the early and the modern economists have different answers to this question.

The classical and neo-classical economists held the view that **utility is quantitatively or cardinally measurable**. It can be measured like height, weight, length and temperature. Their method of measuring utility can be described as follows:

- (i) Walras, a classical economist, used the term 'util' meaning 'units of utility'. The term was used as an accounting unit like kilogram, meter, etc.
- (ii) The classical economists used 'util' as the measure of utility under the assumption that one unit of money equals one 'util'. It implies that price that a consumer pays for a commodity equals the utility derived from the commodity.
- (iii) They assumed that marginal utility of money remains constant, i.e., the utility one derives from each successive unit of money income remains constant whatever the stock of money one holds.

This method of measuring utility has been rejected by the modern economists. For, it was realised over time that absolute or cardinal measurement of utility is not possible. The difficulties in measuring utility proved insurmountable. Money was not found to be a reliable measure of utility because the utility of money itself changes with change in its stock. Neither economists nor psychologists nor other scientists could devise a reliable technique or instrument for measuring the feeling of satisfaction or utility. The modern economists have therefore discarded the concept of cardinal utility.

Notwithstanding the problems in quantitative measurement of utility, the consumption theory based on cardinal utility concept provides deep insight into the consumer psychology and consumer behaviour and remains an indispensable element of consumption theory. In fact, it serves as a starting point in the study of further advances in the theory of consumer behaviour. In this chapter, we discuss the theory of consumer behaviour based on cardinal utility concept.

### Total and Marginal Utility

The concept of cardinal utility, makes it possible to define the **Total and Marginal Utility** in quantitative terms. Total utility ( $TU$ ), with reference to a single commodity, may be defined as the sum of the utility derived from all the units consumed of the commodity. For example, if a consumer consumes 4 units of a commodity and derives  $U_1, U_2, U_3$  and  $U_4$  utils from the successive units consumed, then

$$TU = U_1 + U_2 + U_3 + U_4$$

If he consumes  $n$  units, the total utility ( $TU$ ) from  $n$  units can be expressed as

$$TU_n = U_1 + U_2 + U_3 + \dots + U_n$$

In case the number of commodities consumed and their units are greater than one, then

$$TU = TU_x + TU_y + TU_z + \dots + TU_n$$

where subscripts  $x, y, z$  and  $n$  denote commodities.

The **Marginal Utility** may be defined in three ways. **One**, the marginal utility is the utility derived from the marginal or the last unit consumed.

**Second**, the marginal utility is the addition to the total utility—the utility derived from the consumption or acquisition of one additional unit. Or, Marginal Utility ( $MU$ ) is the change in the total utility resulting from the change in the consumption. That is,

$$MU = \frac{\Delta TU}{\Delta Q}$$

where  $\Delta TU$  = change in total utility, and  $\Delta Q$  = change in quantity consumed of a commodity.

**Three**, marginal utility ( $MU$ ) may also be expressed as

$$MU = TU_n - TU_{n-1}$$

### 3.3.1 Law of Diminishing Marginal Utility

The law of diminishing marginal utility is central to the cardinal utility analysis of the consumer behaviour. This law states that as the quantity consumed of a commodity increases per unit of time, the utility derived by the consumer from the successive units goes on decreasing, provided the consumption of all other goods remains constant. This law stems from the facts (i) that the utility derived from a commodity depends on the intensity or urgency of the need for that commodity, and (ii) that as more and more quantity of a commodity is consumed, the intensity of desire decreases and therefore the utility derived from the marginal unit decreases. For example, suppose you are very hungry and you are offered sandwiches to eat. The satisfaction which you derive from the first piece of sandwich would be the maximum because intensity of your hunger was the highest. When you eat the second piece, you derive a lower satisfaction because intensity of your hunger is reduced due to consumption of one piece of sandwich. As you go on eating more and more sandwiches, the intensity of your hunger goes on decreasing and therefore the satisfaction which you derive from the successive units goes on decreasing. This phenomenon is generalized in the form of a theory called the **Law of Diminishing Marginal Utility**.

**Numerical Example.** Table 3.1 present a numerical illustration of the law of diminishing marginal utility. As the table shows, total utility ( $TU$ ) increases with increase in consumption of sandwiches, but at a decreasing rate. It means that  $MU$  decreases with increase in consumption. This is shown in the last column of the table.

**Table 3.1** Total and Marginal Utility

Sandwiches	Total Utility ( $TU$ )	Marginal Utility
0	0	0 – 0 = 0
1	30	30 – 0 = 30
2	50	50 – 30 = 20
3	60	60 – 50 = 10
4	65	65 – 60 = 5
5	66	66 – 65 = 1
6	60	60 – 66 = – 6

It may be seen in the table that the total utility reaches its maximum at 66 utils when 5 sandwiches are consumed. Here,  $MU = 1$ . Consumption of the 6th sandwich yields negative utility of 6 and therefore total utility starts declining.

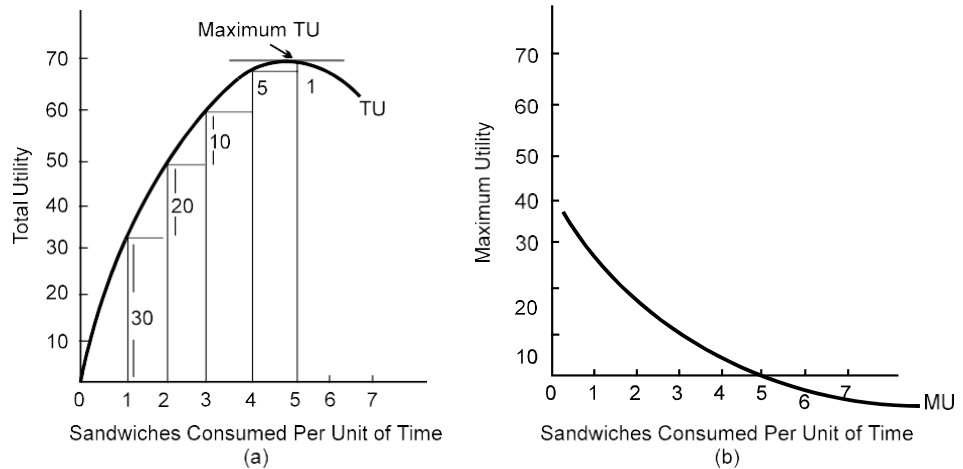
**Graphical Illustration.** The law of diminishing marginal utility is graphically illustrated in Fig. 3.1 (a) and (b). The total utility ( $TU$ ) and marginal utility ( $MU$ ) curves have been obtained by plotting the data given in Table 3.1. The total utility curve goes on rising till the 5th sandwich is consumed [Fig. 3.1 (a)]. Note that the  $TU$  curve is rising but at a diminishing rate. It shows decrease in the  $MU$ , i.e., the utility added to the total. The diminishing  $MU$  is shown by the  $MU$  curve in Fig. 3.1 (b). Beyond 5 sandwiches consumed, the marginal utility turns negative. It means that additional consumption of sandwiches yields disutility in the form of discomfort or displeasure.

### Assumptions

The law of diminishing marginal utility holds only under certain given conditions. These conditions are often referred to as the *assumptions* of the law.

**First**, the unit of the consumer goods must be standard, e.g., a cup of tea, a bottle of cold drink, a pair of shoes or trousers, etc. If the units are excessively small or large, the law may not hold.

**Second**, consumer's taste or preference must remain unchanged during the period of consumption.



**Fig. 3.1** Total and Diminishing Marginal Utility: Cardinal Approach

**Third**, there must be continuity in consumption and where break in continuity is necessary, it must be appropriately short.

**Fourth**, the mental condition of the consumer must remain normal during the period of consumption of a commodity. If a person is eating and also drinking (alcohol) the utility pattern will not be certain.

Given these conditions, the law of diminishing marginal utility holds universally. In some cases, e.g., accumulation of money, collection of hobby items like stamps, old coins, rare paintings and books, and melodious songs, marginal utility may initially increase rather than decrease, but it does decrease eventually. It may thus be stated that the Law of Diminishing Marginal Utility generally operates universally.

---

### 3.4 CONSUMER EQUILIBRIUM AND THE LAW OF EQUI-MARGINAL UTILITY

---

As mentioned earlier, a consumer is assumed to be a utility maximizer. A consumer reaches equilibrium position, when he maximises his total utility given his income and prices of commodities he consumes. Analysing consumer's equilibrium requires answering the question as to how a consumer allocates his money income between the various goods and services he consumes to maximize his total utility.

Before we proceed, let us describe the **assumptions** of the Marshallian approach to the determination of consumer's equilibrium.

#### *Assumptions*

1. **Rationality.** It is assumed that the consumer is a rational being in the sense that he satisfies his wants in the order of their merit. It means that he consumes first a commodity which yields the highest utility and the last which gives the least.
2. **Limited Money Income.** The consumer has a limited money income to spend on the goods and services he consumes.
3. **Maximisation of Satisfaction.** Every rational consumer intends to maximise his satisfaction from his given money income.
4. **Utility is Cardinal Measurable.** The cardinalists assume that utility is cardinal measurable, i.e., it can be measured in absolute terms. For them, utility of one unit of a commodity equals the amount of money paid for it.
5. **Diminishing Marginal Utility.** The utility gained from successive units of a commodity consumed decreases as a consumer consumes a larger quantity of the commodity.
6. **Constant Utility of Money.** The marginal utility of money remains constant and each unit of money has utility equal to 1.

7. **Utility is Additive.** Cardinalists maintain that utility is not only cardinally measurable but also utility derived from various goods and services consumed by a consumer can be added together to obtain the total utility. The additivity of the utility can be expressed through a utility function. Suppose that the basket of goods and services consumed by a consumer contains  $n$  items, and their quantities expressed as  $q_1, q_2, q_3, \dots, q_n$ . The total utility function ( $TU$ ) of the consumer is expressed as

$$U = f(q_1, q_2, q_3, \dots, q_n)$$

Given the utility function, the total utility gained from  $n$  items is expressed as

$$TU_n = U_1(q_1) + U_2(q_2) + U_3(q_3) \dots + U_n(q_n)$$

*Consumer equilibrium: one commodity case*

A consumer consumes a large number of goods and services. Let us however begin our analysis of consumer's equilibrium with a simple case of a consumer consuming only one commodity. Although unrealistic, this case provides an insight for analysing a general case of consumer behaviour.

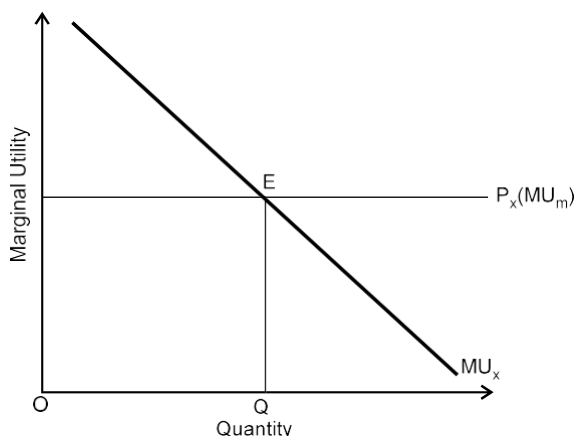
To illustrate consumer's equilibrium, let us suppose that a consumer with certain money income consumes only one commodity,  $X$ . Since both his money income and commodity  $X$  have utility for him, he can either spend his money income on commodity  $X$  or retain it with himself. If the consumer holds his total income, the marginal utility of commodity  $X$  (i.e.,  $MU_x$ ) is bound to be greater than marginal utility of money income ( $MU_m$ ). In that case, total utility can be increased by exchanging money for the commodity. Therefore, a utility maximising consumer exchanges his money income for the commodity so long as  $MU_x > MU_m$ . As assumed above, marginal utility of commodity of  $X$  is subject to diminishing returns whereas marginal utility of money income ( $MU_m$ ) remains constant. Therefore, the consumer will exchange his money income for commodity  $X$  so long as  $MU_x > P_x (MU_m)$ . The utility maximising consumer reaches his equilibrium—the level of his maximum satisfaction—where

$$MU_x = P_x(MU_m) \tag{3.1}$$

Equation (3.1) states the necessary condition for utility maximization. Alternatively, the consumer reaches equilibrium where

$$\frac{MU_x}{P_x(MU_m)} = 1 \tag{3.2}$$

Consumer's equilibrium in a *single-commodity case* is illustrated graphically in Fig. 3.2. The horizontal line  $P_x(MU_m)$  shows the constant utility of money weighed by  $P_x$ , the price of commodity  $X$  and  $MU_x$  curve represents the diminishing marginal utility of commodity  $X$ .



**Fig. 3.2** Consumer's Equilibrium: Cardinal Approach

As Figure 3.2 shows, the  $P_x(MU_m)$  line and  $MU_x$  curve intersect at point  $E$ , where  $MU_x = P_x(MU_m)$ . Therefore, consumer is in equilibrium at point  $E$ . At any point above  $E$ ,  $MU_x > P_x(MU_m)$ . Therefore, if consumer exchanges his money income for commodity  $X$ , he increases his satisfaction. At any point below  $E$ ,  $MU_x < P_x(MU_m)$ , the consumer can therefore increase his satisfaction by reducing his consumption. That is, at any point other than  $E$ , consumer gets satisfaction less than maximum. Therefore, point  $E$  is the point of equilibrium.

#### *Consumer equilibrium: the general case—the law of equi-marginal utility*

We have explained above consumer's equilibrium in a single commodity case. In reality, however, a consumer consumes a large number of goods. The  $MU$  schedules of different commodities is not the same. Some commodities yield a higher  $MU$  schedule and some lower.  $MU$  of some goods decreases more rapidly than that of others. A rational and utility maximising consumer consumes commodities in the order of their utilities. He picks up first the commodity which yields the highest utility and then the commodity yielding the second highest utility and so on. He switches his expenditure from one commodity to another in order of their marginal utility. He continues to switch his expenditure from one commodity to the other till he reaches a stage where  $MU$  of each commodity is the same *per unit of expenditure*.

Let us now analyse a simple two-commodity case. We assume that a consumer consumes only two commodities  $X$  and  $Y$ , their prices being  $P_x$  and  $P_y$ , respectively. Following the equilibrium rule of single commodity case, the consumer distributes his income between commodities  $X$  and  $Y$ , so that

$$MU_x = P_x(MU_m)$$

and 
$$MU_y = P_y(MU_m)$$

or alternatively, in terms of Eq. (3.2) consumer is in equilibrium where

$$\frac{MU_x}{P_x(MU_m)} = 1 \quad \dots(3.3)$$



$$\text{and } \frac{MU_y}{P_y(MU_m)} = 1 \quad \dots(3.4)$$

Equations (3.3) and (3.4) may be written together and equilibrium condition for two-commodity case can be expressed as

$$\frac{MU_y}{P_y(MU_m)} = 1 = \frac{MU_y}{P_y(MU_m)}$$

$$\text{or } \frac{MU_x}{MU_y} = \frac{P_x(MU_m)}{P_y(MU_m)} \quad \dots(3.5)$$

Since, according to assumption 5,  $MU_m = 1$ , Eq. (3.5) may be rewritten as

$$\frac{MU_x}{MU_y} = \frac{P_x}{P_y} \quad \dots(3.6)$$

$$\text{or } \frac{MU_x}{P_x} = \frac{MU_y}{P_y} \quad \dots(3.7)$$

Equation (3.7) leads to the conclusion that the consumer reaches his equilibrium when the marginal utility derived from each rupee spent on the two commodities X and Y is the same.

### *The General Case*

The two-commodity case provides the basis for generalising the consumer's equilibrium. In fact, as mentioned above, a consumer consumes a large number of goods and services with his given income and at given prices. Supposing a consumer consumes A to Z goods and services, his equilibrium condition may be expressed as

$$\frac{MU_A}{P_A} = \frac{MU_B}{P_B} = \frac{MU_C}{P_C} = L = \frac{MU_Z}{P_Z} \quad \dots(3.8)$$

As Eq. 3.8 shows, what a utility maximising consumer intends to equalise is not the marginal utility of each commodity he consumes, but the marginal utility of each unit of his expenditure on various goods and services.

To sum up, the equi-marginal utility rule of consumer equilibrium based on cardinal utility approach may be stated as a consumer maximises his satisfaction by equalising *MU* of each rupee spent on all goods and services he consumes.

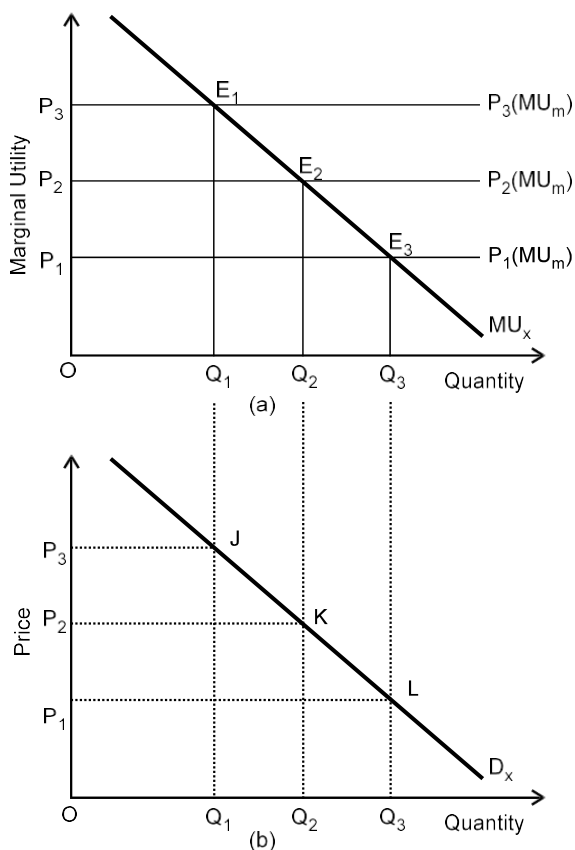
---

## **3.5 DERIVATION OF DEMAND CURVE (CARDINAL UTILITY APPROACH)**

---

We now turn, in this section, to derive the demand curve following the cardinal utility approach. For the derivation of the demand curve, we consider a single-

commodity ( $X$ ) case. According to cardinal utility approach, a consumer reaches his equilibrium where  $MU_x = P_x(MU_m)$ . This logic of consumer's equilibrium provides a convenient basis for the derivation of individual demand curve for a commodity. Marshall was the first economist to explicitly derive the demand curve from consumer's utility function. Marshall gave the equilibrium condition, as noted above, for a single commodity, say  $X$ , as  $MU_x = P_x(MU_m)$ . The same equilibrium condition can be used to derive consumer's demand curve for commodity  $X$  as shown in Fig. 3.3(a).



**Fig. 3.3** Derivation of Demand Curve

Suppose that the consumer is in equilibrium at point  $E_1$ , where  $MU_x = P_3(MU_m)$ . Here equilibrium quantity is  $OQ_1$ . Now if the price of the commodity  $X$  falls to  $P_2$ , the equilibrium condition will be disturbed making  $MU_x > P_2(MU_m)$ . Since  $MU_m$  is constant, the only way to restore the equilibrium condition is to reduce  $MU_x$ . This can be done only by consuming more of commodity  $X$ .

Thus, by consuming  $Q_1Q_2$  additional units of  $X$  he reduces his  $MU_x$  to  $E_2Q_2$  and, thereby, restores equilibrium condition, i.e.,  $MU_x = P_x(MU_m)$ . The consumer reaches a new equilibrium point at  $E_2$  where  $MU_x = P_2(MU_m)$ . Similarly, if price falls further, he buys and consumes more to maximise his satisfaction.

Figure 3.3 (a) reveals that when price is  $P_3$ , equilibrium quantity is  $OQ_1$ . When price decreases to  $P_2$  equilibrium point shifts downward to point  $E_2$  where equilibrium quantity is  $OQ_2$ . Similarly, when price decreases to  $P_1$  and  $P_x(MU_m)$  line shifts downward, equilibrium point shift to  $E_1$  where equilibrium quantity is  $OQ_3$ . It may be inferred from the above analysis that as price decreases, the equilibrium quantity consumed increases as  $OQ_1 < OQ_2 < OQ_3$ .

This price and equilibrium quantity relationship is shown in part (b) of Fig. 3.3. The price-quantity combination corresponding to equilibrium point  $E_3$  is shown at point  $J$ . Similarly, the price-quantity combinations corresponding to equilibrium points,  $E_2$  and  $E_1$  are shown by points  $K$  and  $L$ , respectively. By joining point  $J$ ,  $K$  and  $L$  we get the demand curve for commodity  $X$ . The demand curve,  $D_x$ , is the usual downward sloping Marshallian demand curve.

Before we close the discussion on the cardinal utility theory, let us point out the major deficiencies of this approach.

### 3.5.1 Drawbacks of Cardinal Approach

The following are the basic weaknesses of the Cardinal Utility Theory as pointed out by the economists.

**First**, the very first assumption of cardinal approach that utility is cardinally (or objectively) measurable is untenable. Utility is a subjective concept which cannot be measured objectively.

**Second**, cardinal utility approach assumes that marginal utility of money remains constant and serves as a measure of utility. This assumption is unrealistic because marginal utility of money, like that of all other goods, is subject to change. And, therefore, it cannot serve as a measure of utility derived from goods and services.

**Third**, the psychological Law of Diminishing Marginal Utility has been established from introspection. This law is accepted as an axiom without empirical verification.

**Fourth**, the cardinal utility approach and derivation of demand curve on the basis of this approach are based on the *ceteris paribus* assumption which is unrealistic. It is for this reason that this theory ignores the **substitution** and **income effects** which might operate simultaneously.

**Finally**, the cardinal approach considers that the effect of price changes on the demand curve is exclusively price effect. This assumption is also unrealistic because price effect may include income and substitution effects.

---

## 3.6 ORDINAL UTILITY THEORY

---

In the preceding section, we have explained the cardinal utility theory of consumer behaviour. In this section, we will discuss the ordinal utility theory of consumer behaviour. The technique economists use under ordinal utility approach is called

“indifference curve”. Therefore the ordinal utility approach is also known as indifference curve analysis of consumer behaviour.

The indifference technique was invented and used by Francis Y. Edgeworth (1881) to show the possibility of exchange of commodities between two individuals. About a decade later, Irving Fisher (1892) used indifference curve to explain consumer’s equilibrium. Both Edgeworth and Fisher, however, believed in cardinal measurability of utility. It was Vilfred Pareto who introduced, in 1906, the ordinal utility hypothesis to the indifference curve analysis. In the subsequent decades, many significant contributions were made by Eugen E. Slutsky, W.E. Johnson, and A.L. Bowley. Yet, indifference curve technique could not gain much ground in the analysis of consumer behaviour till early 1930s. In 1934, John R. Hicks and R.G.D. Allen developed systematically the ordinal utility theory as a powerful analytical tool of consumer analysis. Later, Hicks provided a complete exposition of indifference curve technique in his *Value and Capital*. Though in his later work, *A Revision of Demand Theory*, he has dropped some of his earlier assumptions, indifference analysis is regarded as the most powerful tool of consumer analysis.

The fundamental departure that indifference curve analysis makes from the Marshallian marginal utility analysis is the hypothesis that utility can be measured only ordinally, not cardinally. Recall that ‘cardinalists’ assumed that utility is cardinally measurable, and that utility of one commodity is independent of other commodities. In contrast, the ‘Ordinalists’ believe that *cardinal measurement of utility is neither feasible nor necessary to analyse consumer’s behaviour*. According to ordinalists, all that is required to analyse consumer’s behaviour is that the consumer should be able to order his preferences. In fact, the consumer is able to express his preference for the quantity of a commodity to that of others. For example, a consumer can always say that he prefers to 10 kg of wheat to 5 kg of rice.

### *Assumptions of Ordinal Utility Theory*

The indifference curve analysis of consumer’s behaviour makes, at least implicitly, the following assumptions:

1. **Rationality.** The consumer is a rational being. He aims at maximising his total satisfaction, given his income and prices of goods and services he consumes. Furthermore, he has full knowledge of his circumstances.
2. **Ordinal Utility.** Indifference curve analysis assumes that utility can be expressed only ordinally. That is, the consumer is able to tell only the order of his preferences.
3. **Transitivity and Consistency of Choice.** Consumer’s choices are *transitive*. Transitivity of choice means that if a consumer prefers *A* to *B* and *B* to *C*, he must prefer *A* to *C*. Or, if he treats  $A = B$  and  $B = C$ , he must treat  $A = C$ . Consistency of choice means that, if he prefers

$A$  to  $B$  in one period, he will not prefer  $B$  to  $A$  in another period or treat them as equal. The transitivity and consistency in consumer's choices may be symbolically expressed as follows.

**Transitivity.** If  $A > B$ , and  $B > C$ , then  $A > C$ , and

**Consistency.** If  $A > B$ , in one period, then  $B \not> A$  or  $B \not\equiv A$  in another.

4. **Nonsatiety.** It is also assumed that the consumer is not oversupplied with goods in question and that he has not reached the point of saturation in case of any commodity. Therefore, a consumer always prefers a larger quantity of all the goods.
5. **Diminishing Marginal Rate of Substitution.** The marginal rate of substitution means the rate at which a consumer is willing to substitute one commodity ( $X$ ) for another ( $Y$ ), i.e., the units of  $Y$  he is willing to give up for one unit of  $X$  so that his total satisfaction remains the same. This rate is given by  $\frac{\Delta Y}{\Delta X}$ . The assumption is that  $\frac{\Delta Y}{\Delta X}$  goes on decreasing, when a consumer continues to substitute  $X$  for  $Y$ . (We shall know more about marginal rate of substitution in the subsequent sections).

#### *The meaning and nature of indifference curve*

An indifference curve may be defined as *the locus of points, each representing a different combination of two goods but yielding the same level of utility or satisfaction. Since each combination of two goods yields the same level of utility, the consumer is indifferent between any two combinations of goods when it comes to making a choice between them.* A consumer is very often confronted with such a situation in real life. Such a situation arises because he consumes a large number of goods and services, and often he finds that one commodity serves as a substitute for another. It gives him an opportunity to substitute one commodity for another, and to make various combinations of two substitutable goods. It may not be possible for him to tell *how much* utility a particular combination gives, but it is always possible for him to tell which one between any two combinations is preferable to him. It is also possible for him to tell which combinations give him equal satisfaction. If a consumer is faced with equally good combinations, he would be indifferent between the combinations. When such combinations are plotted graphically, the resulting curve is known as **indifference curve**. Indifference curve is also called **Iso-utility Curve** and **Equal Utility Curve**.

For example, let us suppose that a consumer consumes only two commodities  $X$  and  $Y$  and he makes five combinations which he calls  $a$ ,  $b$ ,  $c$ ,  $d$  and  $e$ . All these combinations yield him equal utility. Therefore, he is indifferent between the combinations  $a$ ,  $b$ ,  $c$ ,  $d$ , and  $e$  of two commodities,  $X$  and  $Y$ . His combinations are presented in Table 3.2, which may be called as **indifference schedule**—a schedule of various combinations of two goods, between which a consumer is indifferent. The last column of the table shows an undefined utility ( $u$ ) derived from each combination of  $X$  and  $Y$ . If combinations  $a$ ,  $b$ ,  $c$ ,  $d$ ,

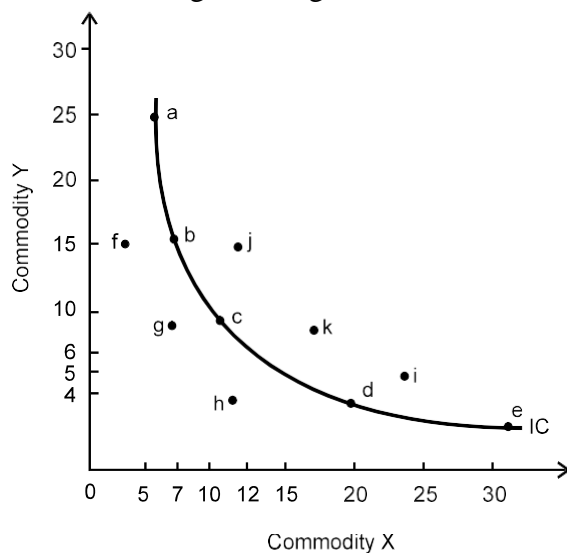
and  $e$  given in Table 3.2 are plotted and joined to form a smooth curve (as shown in Fig. 3.4), the resulting curve is known as **indifference curve**. On this curve, one can locate many other points showing many other combinations of  $X$  and  $Y$  which yield the same satisfaction. Therefore, the consumer is indifferent between the combinations which may be located on the indifference curve.

**Table 3.2** Indifference Schedule of Commodities  $X$  and  $Y$

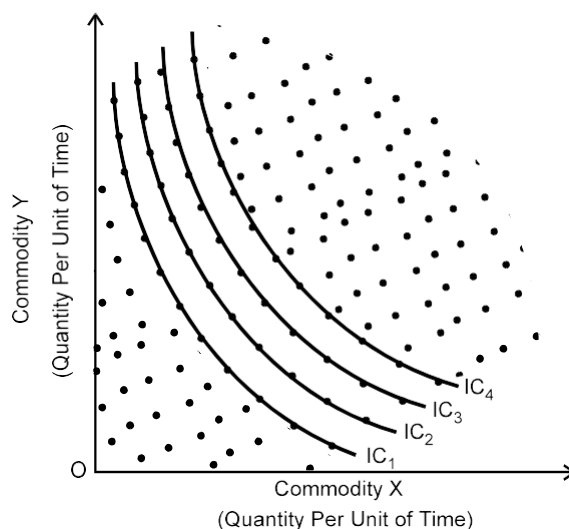
Combination	Commodity $X$	Commodity $Y$	Utility
$a$	25	5	$u$
$b$	15	7	$u$
$c$	10	12	$u$
$d$	6	20	$u$
$e$	4	30	$u$

### The Indifference Map

We have drawn a single indifference curve in Fig. 3.4 on the basis of an indifference schedule given in Table 3.2. The combinations of the two commodities,  $X$  and  $Y$ , given in the indifference schedule or those indicated by the indifference curve are by no means the only combinations of the two commodities. The consumer may be faced with many other combinations with *less* of one or both the goods—each combination yielding the same level of satisfaction but less than the level of satisfaction indicated by the indifference curve  $IC$  in Fig. 3.4. Therefore, another indifference curve can be drawn, say, through points  $f$ ,  $g$  and  $h$ . Note that this indifference curve falls below the curve  $IC$  given in Fig. 3.4. Similarly, he may be faced with many other combinations with *more* of one or both the goods—each combination yielding the same satisfaction—but greater than the satisfaction indicated by the lower indifference curves. Thus, another indifference curve can be drawn above the  $IC$  given in Fig. 3.4, say, through points  $j$ ,  $k$ , and  $l$ . This exercise may be repeated as many times as one wants, each time generating a new indifference curve.



**Fig. 3.4** Indifference Curve



**Fig. 3.5** *The Indifference Map*

In fact the area between  $X$  and  $Y$  axes is known as *indifference plane* or *commodity space*. This plane is full of finite points and each point on the plane indicates a different combination of goods  $X$  and  $Y$ . Intuitively, it is always possible to locate any two or more points indicating different combinations of goods  $X$  and  $Y$  yielding the same satisfaction. It is thus possible to draw a number of indifference curves without intersecting or touching the other, as shown in Fig. 3.5. The set of indifference curves,  $IC_1$ ,  $IC_2$ ,  $IC_3$  and  $IC_4$  drawn in this manner make the *indifference map*. In fact, an indifference map may contain any number of indifference curves, ranked in the order of consumer's preferences.

### 3.6.1 *The Diminishing Marginal Rate of Substitution*

When a consumer makes different combination of two goods, yielding the same level of satisfaction, he substitutes one good for another. The rate at which he substitutes one good for the other is called the 'Marginal Rate of Substitution ( $MRS$ )'. One of the basic postulates of *indifference curve analysis* is that ( $MRS$ ) diminishes. The axiomatic assumption of ordinal utility theory is analogous to the assumption of 'Diminishing Marginal Utility' in *cardinal utility theory*. The *postulate of diminishing marginal rate of substitution states* an observed behavioural rule that when a consumer substitutes one commodity (say  $X$ ) for another (say  $Y$ ), the 'Marginal Rate of Substitution' ( $MRS$ ) decreases as the stock of  $X$  increases and that of  $Y$  decreases.

#### *Measuring Marginal Rate of Substitution ( $MRS$ )*

Conceptually, the  $MRS$  is the rate at which one commodity can be substituted for another, the level of satisfaction remaining the same. The  $MRS$  between two

commodities,  $X$  and  $Y$ , can also be defined as the number of units of  $X$  which are required to replace one unit of  $Y$  (or number of units of  $Y$  that are required to replace one unit of  $X$ ), in the combination of the two goods so that the total utility remains the same. It implies that the utility of units of  $X$  (or  $Y$ ) given up is equal to the utility of additional units of  $Y$  (or  $X$ ) added to the basket.

To explain symbolically the concept of  $MRS$ , let us suppose that the utility function of a consumer is given as

$$U = f(X, Y)$$

Let us now suppose that the consumer substitutes  $X$  for  $Y$ . When the consumer foregoes some units of  $Y$ , his stock of  $Y$  decreases by  $-\Delta Y$ . His loss of utility may be expressed as

$$-\Delta Y \cdot MU_y$$

On the other hand, as a result of substitution, his stock of  $X$  increases by  $\Delta X$ . His utility from  $\Delta X$  equals

$$+\Delta X \cdot MU_x$$

For the total utility  $U$  to remain the same,  $-\Delta Y \cdot MU_y$  must be equal to  $\Delta X \cdot MU_x$ . That is,

$$-\Delta Y \cdot MU_y + \Delta X \cdot MU_x = 0 \quad \dots(3.9)$$

Rearranging the terms in Eq. (3.9), we get  $MRS$  of  $X$  for  $Y$  as

$$-\frac{\Delta Y}{\Delta X} = \frac{MU_x}{MU_y} \quad \dots(3.10)$$

Here,  $\Delta Y/\Delta X$  is simply the *slope* of the indifference curve, which gives the  $MRS_{x,y}$  when  $X$  is substituted for  $Y$ . Similarly,  $\Delta X/\Delta Y$  gives  $MRS_{y,x}$  when  $Y$  is substituted for  $X$ .

$$\begin{aligned} MRS_{x,y} &= \frac{\Delta Y}{\Delta X} = \frac{MU_x}{MU_y} \\ \text{and } MRS_{y,x} &= \frac{\Delta X}{\Delta Y} = \frac{MU_y}{MU_x} = \text{Slope of indifference curve} \end{aligned}$$

### The Diminishing $MRS$

As mentioned basic postulate of ordinal utility theory is that the  $MRS_{x,y}$  (or  $MRS_{y,x}$ ) decreases. That is, the number of units of a commodity that a consumer is willing to sacrifice for an additional unit of another goes on decreasing when he goes on substituting one commodity for another. The diminishing  $MRS_{x,y}$  which can be obtained from Table 3.4, are presented in Table 3.3.



**Table 3.3** The Diminishing MRS between Commodities X and Y

Movements on IC	Change in Y (-ΔY)	Change in X (ΔX)	MRS <sub>y,x</sub> (ΔY/ΔX)
From point a to b	- 10	2	- 5.0
From point b to c	- 5	5	- 1.0
From point c to d	- 4	8	- 0.5
From point d to e	- 2	10	- 0.2

As Table 3.3 shows, when the consumer moves from point a to b on the indifference curve (Fig. 3.4) he gives up 10 units of commodity Y and gets only 2 units of commodity X, so that

$$MRS_{y,x} = \frac{\Delta Y}{\Delta X} = \frac{10}{2} = 5$$

As he moves down from points b to c, he loses 5 units of Y and gains 5 units of X, giving

$$MRS_{x,y} = \frac{\Delta Y}{\Delta X} = \frac{5}{5} = 1$$

The  $MRS_{x,y}$  goes on decreasing as the consumer moves further down along the indifference curve. The diminishing marginal rate of substitution causes the indifference curves to be convex to the origin.

The diminishing marginal rate of substitution can also be illustrated graphically, as shown in Fig. 3.6.

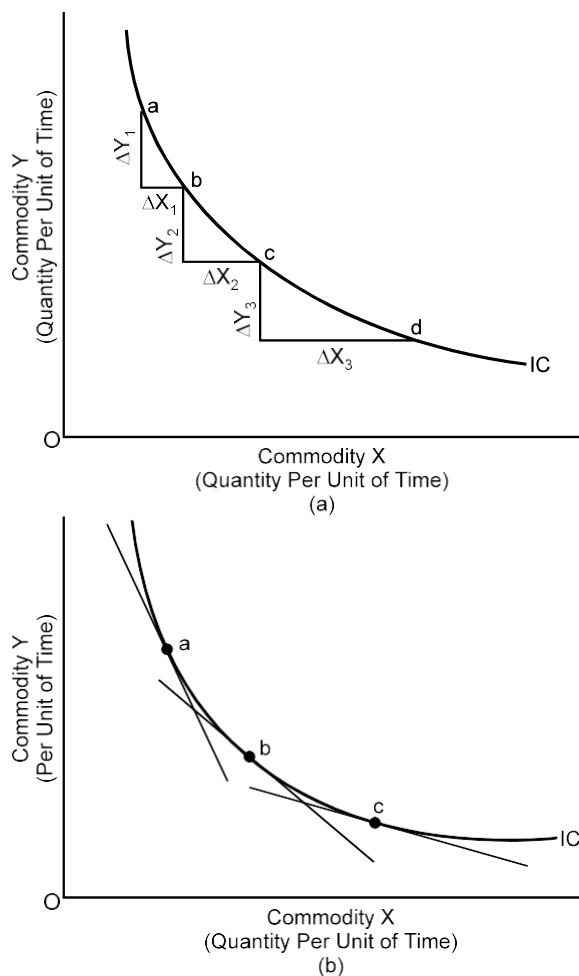
As the consumer moves from point a to b, to c, and to d, he gives up a constant quantity of Y, (i.e.,  $\Delta Y_1 = \Delta Y_2 = \Delta Y_3$ ). To substitute a constant quantity of ΔY, he requires an increasing quantity of X (i.e.,  $\Delta X_1 < \Delta X_2 < \Delta X_3$ ). Since

MRS is given by the slope of the indifference curve, (i.e., ΔY/ΔX), arranging the slopes between points a and b, b and c, and c and d, in descending order, we get

$$\frac{\Delta Y_1}{\Delta X_1} > \frac{\Delta Y_2}{\Delta X_2} > \frac{\Delta Y_3}{\Delta X_3}$$

These inequalities show that MRS (=ΔY/ΔX) goes on decreasing as the consumer moves from point a towards point d.

The diminishing MRS is geometrically illustrated in Fig. 3.6 (b). The lines tangent to the indifference curve at points a, b and c measure the slope of the curve at these points. It can be seen from the figure, that as the consumer moves from point a towards d, the tangential lines become flatter indicating decrease in the slope of the indifference curve. This also proves the decrease in MRS all along the indifference curve.



**Fig. 3.6** Diminishing Marginal Rate of Substitution

### Why Does MRS Diminish?

The negative slope of the indifference curve implies that two commodities are not perfect substitutes for each other. In case they are perfect substitutes, the indifference curve will be a straight line with a negative slope. Since goods are not perfect substitutes for each other, the *subjective value attached to the additional quantity (i.e., MU) of a commodity decreases fast in relation to the other commodity whose total quantity is decreasing*. Therefore, when the quantity of one commodity (say,  $X$ ) increases and that of other (say,  $Y$ ) decreases, it becomes increasingly difficult for the consumer to sacrifice more and more units of commodity  $Y$  for one unit of  $X$ . But, if he is required to sacrifice additional units of  $Y$ , he will demand increasing units of  $X$  to maintain the level of his satisfaction. As a result, the  $MRS$  decreases.

Furthermore, when the combination of two goods at a point of indifference curve is such that it includes a large quantity of one commodity, (say,  $Y$ ) and a small quantity of the other (commodity  $X$ ), then consumer's capacity to sacrifice  $Y$

is greater than to sacrifice  $X$ . Therefore, he can sacrifice a large quantity of  $Y$  in favour of a smaller quantity of  $X$ . This is an observed behavioural rule that the consumer's willingness and capacity to sacrifice a commodity is greater when its stock is greater and it is lower when the stock of a commodity is smaller. These are the reasons why  $MRS$  decreases all along on the indifference curve.

---

### 3.7 COMPARISON OF CARDINAL AND ORDINAL UTILITY APPROACH

---

Having outlined the indifference curve technique of deriving Marshallian demand curve, let us now compare the cardinal and ordinal utility approaches to consumer's analysis and look into the relative merits of the two approaches.

#### *Similarity between the two approaches*

Some of the assumptions made under the two approaches are the same. For example, both cardinal and ordinal utility approaches assume rationality on the part of the consumer. Both the approaches assume that the consumer aims at maximising his total utility given his income and market prices. The *diminishing marginal utility* assumption of the cardinal utility approach is implicit in the *diminishing marginal rate of substitution* assumption of the ordinal utility approach.

#### *Equilibrium Conditions are Identical*

Both cardinal and ordinal utility approaches use identical equilibrium condition. Recall that the necessary (or the first order) equilibrium condition of cardinal utility approach is given as

$$\frac{MU_x}{MU_y} = \frac{P_x}{P_y} \quad \dots(3.11)$$

and the same under ordinal utility approach is given as

$$MRS_{x,y} = \frac{P_x}{P_y} \quad \dots(3.12)$$

$MRS_{x,y}$  in Eq. (3.12) is the same as  $MU_x/MU_y$  in Eq. (3.11). Therefore, Eqs. (3.11) and (3.12) are the same. Thus, in spite of the fact that cardinal and ordinal approaches are based on different assumptions regarding measurability of utility, both arrive at the same conclusion.

The second order equilibrium condition under cardinal utility approach is that the total expenditure must not exceed consumer's total income. For example, if a consumer having money income  $M$ , consumes only two goods,  $X$  and  $Y$ , given their prices as  $P_x$  and  $P_y$ , then

$$Q_x \cdot P_x + Q_y \cdot P_y = M$$

where  $Q_x$  = quantity of  $X$  and  $Q_y$  = quantity of  $Y$ . This is similar to the second order condition of the ordinal utility approach, i.e., the first order equilibrium condition must be fulfilled at the highest possible indifferent curve. It implies the same thing, i.e., at the equilibrium,

$$Q_x \cdot P_x + Q_y \cdot P_y = M$$

### Superiority of indifference curve approach

In spite of their similarity in some respects, Hicksian indifference curve analysis is superior in many respects to the Marshallian ordinal utility approach. The indifference curve analysis has proved helpful in making some major advances in the theory of consumer behaviour at least in following respects.

**First**, while cardinal utility approach assumes cardinal measurability of utility, ordinal utility approach assumes only ordinal expression of utility. The assumptions of indifference curve approach are less stringent or restrictive than those of cardinal utility approach. Besides, the ordinal utility approach does not assume constancy of utility of money. Marshallian assumption of constancy of marginal utility of money is incompatible with demand functions involving more than one good.

**Secondly**, indifference curve approach provides a better criterion for the classification of goods into substitutes and complements. This is considered as one of the most important contributions of ordinal utility approach. The cardinal utility approach uses the sign of cross-elasticity for the purpose of classifying goods into substitutes and complements. The cross-elasticity between two goods,  $X$  and  $Y$  is given by

$$e_{y,x} = \frac{\frac{\partial Q_y}{\partial P_x}}{\frac{Q_y}{P_x}} \cdot \frac{P_x}{P_y}$$

If cross-elasticity has a positive sign, it means  $X$  and  $Y$  are substitutes for each other and if cross-elasticity has a negative sign, it means they are complements. This method of classifying goods into substitutes and complements is absurd and misleading. For, the measure of cross-elasticity uses the total effect of a price change ( $\partial P$ ) on quantity demanded ( $\partial Q$ ) without compensating for the change in real income caused by the change in price of a commodity (i.e.,  $\partial P$ ). On the contrary, according to ordinal utility approach, two goods  $X$  and  $Y$  are substitutes for each other only if cross-elasticity measured after eliminating the income effect is positive.

Although Hicksian criterion for classifying goods into substitutes and complements is theoretically superior to simple cross-elasticity (unadjusted for real income-effect), economists consider it *impracticable*. For, estimating income and substitution effects of a price-change is an extremely difficult task. On the other hand, the usual cross-elasticity method is feasible because it requires only the knowledge of market demand function which is empirically estimable.

**Thirdly**, indifference curve analysis provides a more realistic measure of consumer's surplus compared to one provided by Marshall. Marshallian concept of 'Consumer's surplus' is based on the assumptions that utility is cardinally

measurable in terms of money and that utility of money remains constant. Neither of the two assumptions is realistic. Indifference curves analysis, on the other hand, measures consumer's surplus in terms of ordinal utility. Hicksian measure of consumer's surplus is of great importance in welfare economics and in the formulation and assessment of government policy.

---

### 3.8 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

---

1. The concept of *ordinal utility* is based on the fact that it may not be possible for consumers to express the utility of a commodity in absolute or quantitative terms, but it is always possible for a consumer to tell introspectively whether a commodity is more or less or equally useful compared to another.
2. The term 'utility' refers to the power or property of a commodity to satisfy human needs.
3. The indifference technique was invented and used by Francis Y. Edgeworth (1881) to show the possibility of exchange of commodities between two individuals.
4. The marginal rate of substitution means the rate at which a consumer is willing to substitute one commodity ( $X$ ) for another ( $Y$ ), i.e., the units of  $Y$  he is willing to give up for one unit of  $X$  so that his total satisfaction remains the same.

---

### 3.9 SUMMARY

---

- Utility is a psychological phenomenon. It is a feeling of satisfaction, pleasure or happiness.
- Classical economists, viz., Jeremy Bentham, Leon Walrus, Carl Menger, etc. and neo-classical economists, notably Alfred Marshall, believed that utility is cardinally or quantitatively measurable like height, weight, length, temperature and air pressure. This belief resulted in the Cardinal Utility concept.
- However, modern economists, most notably J.R. Hicks and R.G.D. Allen, hold the view that utility is not quantitatively measurable—it is not measurable in absolute terms.
- Utility can be expressed only ordinally or in terms of 'less than' or 'more than'.

- Some early psychological experiments on an individual's responses to various stimuli led neo-classical economists to believe that utility is measurable and cardinally quantifiable.
- According to neo-classical economists, utility can be measured in terms of money. That is, utility of a unit of a commodity for a person is equal to the amount of money he is willing to pay for it.
- The modern economists have discarded the concept of cardinal utility and have instead employed the concept of ordinal utility for analyzing consumer behaviour.
- The concept of ordinal utility is based on the fact that it may not be possible for consumers to express the utility of a commodity in absolute or quantitative terms, but it is always possible for a consumer to tell introspectively whether a commodity is more or less or equally useful compared to another.
- Cardinal Utility Approach, attributed to Alfred Marshall and his followers, is also called the Neo-classical Approach.
- Ordinal Utility Approach, pioneered by J.R. Hicks, a Nobel laureate, and R.G.D. Allen, is also called the Indifference Curve Analysis.
- The notion of "Utility" was introduced to social thought by the British philosopher, Jeremy Bentham, in the 18th century and to economics by William Stanley Jevons in the 19th century.
- In abstract sense, the term "utility" refers to the power or property of a commodity to satisfy human needs.
- The concept of cardinal utility, makes it possible to define the Total and Marginal Utility in quantitative terms.
- Total utility (TU), with reference to a single commodity, may be defined as the sum of the utility derived from all the units consumed of the commodity.
- The law of diminishing marginal utility is central to the cardinal utility analysis of the consumer behaviour. This law states that as the quantity consumed of a commodity increases per unit of time, the utility derived by the consumer from the successive units goes on decreasing, provided the consumption of all other goods remains constant.
- The equi-marginal utility rule of consumer equilibrium based on cardinal utility approach may be stated as a consumer maximises his satisfaction by equalising *MU* of each rupee spent on all goods and services he consumes.

---

### 3.10 KEY WORDS

---

- **Utility:** It is the capacity of a commodity to satisfy human wants.
- **Marginal:** It is used to indicate the change in some benefit or cost, when an additional unit is produced.

- **Equilibrium:** It is a condition or state in which economic forces are balanced.

---

### 3.11 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

#### Short-Answer Questions

1. Compare and contrast cardinal and ordinal utility.
2. Write a short note on total and marginal utility.
3. Derive the demand curve following the cardinal utility approach.
4. Mention the drawbacks of cardinal approach.
5. Briefly discuss the meaning and nature of indifference curve.

#### Long-Answer Questions

1. Discuss in detail the meaning and concept of cardinal utility theory.
2. What is law of diminishing marginal utility? What are the assumptions under which the law functions.
3. Describe the assumptions of the Marshallian approach to the determination of consumer's equilibrium. Explain consumer equilibrium with one commodity case.
4. What is the law of equi-marginal utility? Explain in terms of consumer equilibrium.
5. Discuss in detail the ordinal utility theory.

---

### 3.12 FURTHER READINGS

---

Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.

Keat Paul G. and K.Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.

Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition. Singapore: Pearson Education, Inc.

Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Publishers.

---

*Structure*

- 4.0 Introduction
- 4.1 Objectives
- 4.2 The Law of Demand
  - 4.2.1 Demand Curve and Demand Schedule
  - 4.2.2 Derivation of Individual Demand Curve (Utility Analysis)
  - 4.2.3 Reasons and Exceptions to the Law of Demand
- 4.3 Determinants of Market Demand
- 4.4 Answers to Check Your Progress Questions
- 4.5 Summary
- 4.6 Key Words
- 4.7 Self Assessment Questions and Exercises
- 4.8 Further Readings

---

**4.0 INTRODUCTION**

---

Supply and demand are the economic forces of the free market that control what suppliers are willing to produce and what consumers are willing and able to purchase. This willingness or want creates demand which becomes the basis of supply.

Demand refers to how much of a commodity, or service consumers are willing to buy at a particular price. Demand, which does influence supply, plays a key role in setting the price of a particular product in the market economy. Since demands of buyers are endless, not all that is demanded can be supplied due to scarcity of resources.

In this unit, you will study about market demand, its meaning and determinants of demand in detail.

---

**4.1 OBJECTIVES**

---

After going through this unit, you will be able to:

- Explain the law of demand
- Describe demand curve and demand schedule
- Assess the determinants of demand
- Discuss the demand function



---

## 4.2 THE LAW OF DEMAND

---

The term 'demand' refers to the *quantity demanded* of a commodity *per unit of time* at a given price. It implies also a *desire* backed by ability and willingness to pay. A mere desire of a person to purchase a commodity is not his demand. He must possess adequate resources and must be willing to spend his resources to buy the commodity. Besides, the *quantity demanded* has always a reference to 'a price' and 'a unity of time'. The quantity demanded referred to 'per unit of time' makes it a *flow* concept. Apparently there may be some problems in applying this flow concept to the demand for durable consumer goods like house, car, refrigerators, etc. But this apparent difficulty may be resolved by considering the fact that the total service of a durable good is not consumed at one point of time and its utility is not exhausted in a single use. The service of a durable good is consumed over time. At a time, only a part of its service is consumed. Therefore, the demand for the services of durable consumer goods may also be visualised as a demand per unit of time. However, this problem does not arise when the concept of demand is applied to total demand for a consumer durable. Thus, the demand for consumer goods also is a flow concept.

### *The law of demand*

The quantity of a commodity that an individual or a household consumes per unit of time is determined by a number of factors including price of the commodity, price of its substitutes and complements, consumer's income, his/her wealthholding, taste and preference, expectations about future income and price, demonstration effect, etc. In the short run, however, all factors other than price of the commodity are assumed to remain constant. In the **short run**, therefore, the law of demand is linked to the price of the commodity.

Let it be emphasized that price of a commodity is the most important determinant of its demand, and the only determinant in the short run when all other determinants of demand are assumed to remain constant. The relationship between price and demand is expressed by the *law of demand*. The **law of demand** states that *quantity of a product demanded per unit of time increases when its price falls, and decreases when its price increases, other factors remaining constant*. The assumption 'other factors remaining constant' implies that income of the consumers, prices of the substitutes and complementary goods, consumers' taste and preference, and number of consumers, remain unchanged.

The **law of demand** can be illustrated through a demand schedule. A demand schedule is a series of quantities which consumers would like to buy per unit of time at different prices. It shows seven alternative prices and the corresponding

quantities (number of cups of tea) demand per day. Each price has a unique quantity demanded, associated with it. As price per cup of tea decreases, daily demand for tea increases, in accordance with the law of demand.

#### 4.2.1 Demand Curve and Demand Schedule

The law of demand can also be presented through a curve called *demand curve*. A demand curve is a locus of points showing various alternative price-quantity combinations. It shows the quantities of a commodity that consumers or users would buy at different prices per unit of time under the assumptions of the law of demand. An individual's demand curve for tea as given in Fig. 4.1 can be obtained by plotting the data given in Table 4.1.

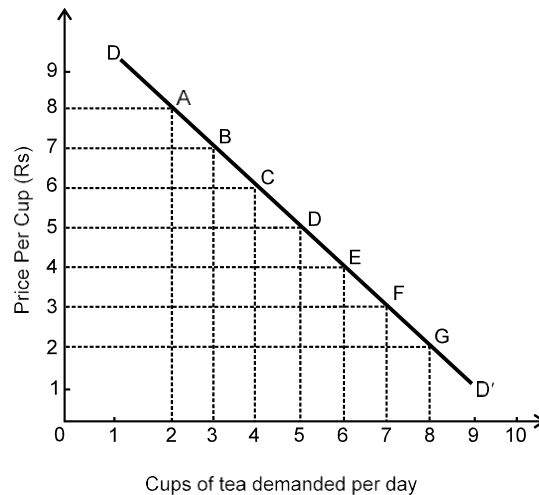
**Table 4.1** Demand Schedule for Tea

Price per cup of tea ( ₹ )	No. of cups of tea demand per consumer per day	Symbols representing per price-quantity combination
8	2	A
7	3	B
6	4	C
5	5	D
4	6	E
3	7	F
2	8	G

In Fig. 4.1, the curve from point A to point G passing through points B, C, D, E and F is the demand curve DD. Each point on the demand curve DD shows a unique price-quantity combination. The combinations read in alphabetical order show decreasing price of tea and increasing number of cups of tea demanded per day. Price-quantity combinations in reverse order of alphabets show increasing price of tea per cup and decreasing number of cups of tea per day consumed by an individual. The whole demand curve shows a functional relationship between the alternative price of a commodity and its corresponding quantities which a consumer would like to buy during a specific period of item—per day, per week, per month, per season, or per year. *The demand curve shows an inverse relationship between price and quantity demanded.* This inverse relationship between price and quantity demanded makes demand curve slope downward to the right.

#### Why Demand Curve Slopes Downward to the Right

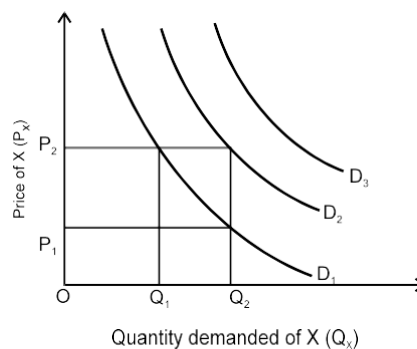
As Fig. 4.1 shows, demand curve slopes downward to the right. The downward slope of the demand curve reads the law of demand, i.e., the quantity of a commodity demanded per unit of time increases as its price falls, and *vice versa*.



**Fig. 4.1** The Demand Curve

### Shift in Demand Curve

When demand curve changes its position retaining its shape (though not necessarily), the change is known as *shift in demand curve*. Let us suppose that the demand curve,  $D_2$  in Fig. 4.2 is the original demand curve for commodity  $X$ . As shown in the figure, at price  $OP_2$  consumer buys  $OQ_2$  units of  $X$ , other factors remaining constant. Now if any of the other factors (e.g., consumer's income) changes, it will change the consumer's ability and willingness to buy commodity  $X$ . For example, if consumer's disposable income decreases, say, due to increase in income tax, he may be able to buy only  $OQ_1$  units of  $X$  instead of  $OQ_2$  at price  $OP_2$ . This is true for the whole range of price of  $X$ —the consumers would be able to buy less of commodity  $X$  at all other prices. This will cause a *downward shift* in demand curve from  $D_2$  to  $D_1$ . Similarly, increase in disposable income of the consumer due to reduction in taxes may cause an *upward shift* from  $D_2$  to  $D_3$ . Such changes in the position of the demand curve are known as *shifts in demand curve*.



**Fig. 4.2** Shift in Demand Curves

### Reasons for Shift in Demand Curve

Shifts in a price-demand curve may take place owing to the change in one or more of other determinants of demand. Consider, for example, the decrease in demand

for commodity  $X$  by  $OQ_1Q_2$  in Fig 4.2. Given the price  $OP_1$ , the demand for  $X$  might have fallen from  $OQ_2$  to  $OQ_1$  (i.e., by  $Q_1Q_2$ ) for any of the following reasons:

- (i) fall in the consumer's income so that he can buy only  $OQ_1$  of  $X$  at price  $OP_2$ —it is **income effect**;
- (ii) price of  $X$ 's substitute falls so that the consumers find it beneficial to substitute  $Q_1Q_2$  of  $X$  with its substitute—it is **substitution effect**;
- (iii) advertisement made by the producer of the substitute, changes consumer's taste or preference against commodity  $X$  so much that they replace  $Q_1Q_2$  of  $X$  with its substitute, again a **substitution effect**;
- (iv) price of complement of  $X$  increases so much that they can now afford only  $OQ_1$  of  $X$ ; and
- (v) also for such reasons as commodity  $X$  is going out of fashion; its quality has deteriorated; consumer's technology has so changed that only  $OQ_1$  of  $X$  can be used, and due to change in season if commodity  $X$  has only seasonal use.

#### 4.2.2 Derivation of Individual Demand Curve (Utility Analysis)

Deriving demand curve is the ultimate aim of the entire utility analysis—cardinal or ordinal. The basic purpose of the entire exercise in indifference curve technique so far is to construct the individual demand curve for a commodity. As stated earlier, an individual demand curve shows the relationship between the quantity demanded of a commodity (say,  $X$ ) and its price ( $P_x$ ), under *ceteris paribus* assumption. Thus, to draw an individual demand curve, we need different levels of  $P_x$  and the corresponding quantities demanded of  $X$ . This information can be obtained from the price-consumption curve. The price-consumption curve (*PCC*) shown in Fig. 4.3(a), contains the information required for constructing the individual demand curve for  $X$ . The derivation of the demand curve is illustrated in Fig. 4.3(b).

In Fig. 4.3(a), quantity of commodity  $X$  is measured on X-axis and consumer's money income ( $M$ ) on Y-axis. Note that, on Y-axis, commodity  $X$  has been replaced by money income ( $M$ ) only for the sake of simplicity. As Fig. 4.3(a) shows, with  $P_x$  decreasing from  $P_1$  to  $P_2$  and then to  $P_3$ , the budget line rotates to the right, from  $MN_1$  to  $MN_2$  and then to  $MN_3$ . As a result, the consumer moves from equilibrium point  $E_1$  to  $E_2$  and finally to point  $E_3$  on the *PCC*.

The information provided by the *PCC* is given in the following table. As Fig. 4.3(a) shows,  $M$  is constant at  $OM$  and quantity of  $X$  goes on increasing as price ( $P_x$ ) decreases, i.e.,  $N_1 < N_2 < N_3$ . Therefore,  $M/N_1 > M/N_2 > M/N_3$ , i.e.,  $P_1 > P_2 > P_3$ . The data provided by the *PCC* is summarised in the table. The demand curve may be constructed by plotting the price-quantity data given in the table.

<i>Price = M/N</i>	<i>Equilibrium</i>	<i>Quality demanded of X</i>
$M/N_1 = P_1$	$E_1$	$OX_1$
$M/N_2 = P_2$	$E_2$	$OX_2$
$M/N_3 = P_3$	$E_3$	$OX_3$

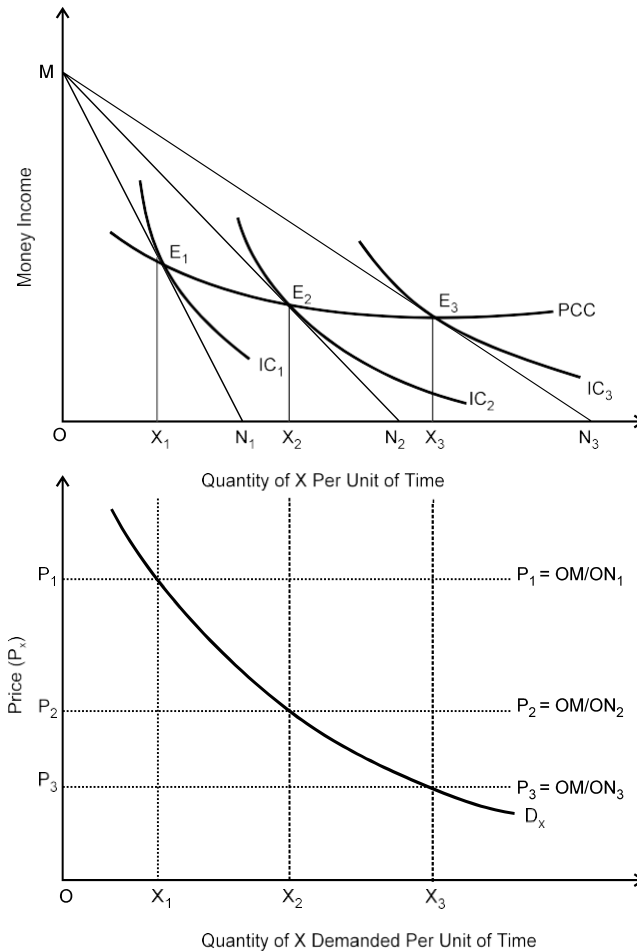
$M$  = money income;  $N$  = quantity purchased

The demand curve may be constructed directly from Fig. 4.3(a). This is shown in Fig. 4.3(b). Part (b) of the figure is derived from part (a). In Fig. 4.3(b), the horizontal axis measures the quantity of  $X$ . Quantities  $OX_1, OX_2$ , and  $OX_3$ , are the same equilibrium quantities of  $X$  as in Fig. 4.3(a). Price of  $X$  is measured on the vertical axis. Prices  $P_1, P_2$  and  $P_3$  are derived from information contained in

1 2 3

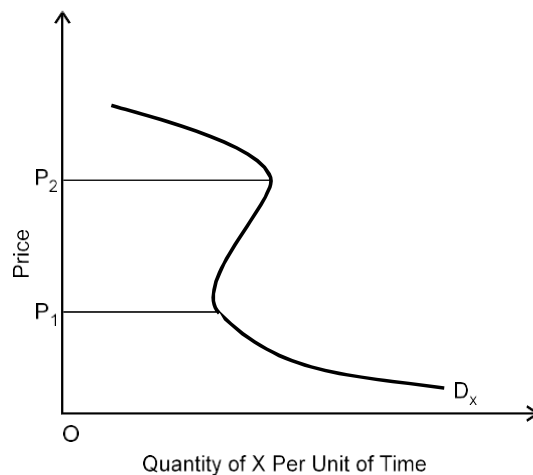
Part (b) of the figure. The relationship between prices  $P_1, P_2$  and  $P_3$  and the respective quantities  $OX_1, OX_2$  and  $OX_3$  shown by the curve  $D_x$  is the demand curve for commodity  $X$ . The demand curve  $D_x$  is a usual downward sloping demand curve.

The precise shape and slope of the demand curve depends, however, on the direction in which income and substitution effects work as a result of a fall in the price of a commodity. In fact, the substitution effect is always negative but income effect is uncertain. Therefore, given the negative substitution effect, the shape and nature of demand curve depends on the direction and magnitude of the income effect. The four possible combinations of substitution and income effects of a fall in price of a commodity and the corresponding nature of the demand curve may be summarised as follows.



**Fig. 4.3** Derivation of individual Demand Curve

1. When substitution effect is negative and income effect is positive, quantity demanded of  $X$  increases as  $P_x$  decreases. The demand curve therefore slopes downward to the right. This is a case of 'normal' goods.
2. If income effect is negative but less than the (negative) substitution effect (as it happens in case of inferior goods) the demand curve slopes downward to the right more steeply than usual.
3. If income effect is zero, the demand curve follows the substitution effect, i.e., as price decreases, demand increases. The demand curve has a negative slope, but is relatively flatter.
4. If income effect is negative and is more powerful than the substitution effect (as it happens in the case of Giffen goods) demand curve bends backward as in the case of Engel curve. But it is most unlikely that any demand curve will slope downward to the left throughout its whole length. It will be so only over that range of price change over which negative income effect is stronger than the substitution effect. Therefore, the most likely shape of the demand curve for a Giffen good is one shown in Fig. 4.4. The demand curve for a Giffen good, may begin as an usual downward sloping one, till price falls to  $P_2$ . But if price falls further, the income-effect may become negative and so powerful that it outweighs the substitution effect. Then the demand curve for a Giffen good becomes a backward sloping one. If price continues to fall, say below  $P_1$ , the demand may once again increase for the Giffen good. This seems to be most likely shape of the demand curve for a Giffen good.



**Fig. 4.4** Demand Curve for a Giffen Good

#### 4.2.3 Reasons and Exceptions to the Law of Demand

The reasons behind the law of demand, i.e., **inverse relationship between price and quantity demanded** are following.

**(i) Substitution Effect**

When the price of a commodity falls it becomes relatively cheaper if price of all other related goods, particularly of substitutes, remain constant or, in other words, substitute goods become relatively costlier. Since consumers substitute cheaper goods for costlier ones, demand for the relatively cheaper commodity increases. The increase in demand on account of this factor is known as *substitution effect*.

**(ii) Income Effect**

As a result of fall in the price of a commodity, the real income of its consumer increase at least in terms of this commodity. In other words, his/her purchasing power increases since he is required to pay less for the same quantity. The increase in real income (or purchasing power) encourages demand for the commodity with reduced price. The increase in demand on account of increase in real income is known as *income effect*.

It should however be noted that the *income effect* is negative in case of inferior goods. In case price of an inferior good accounting for a considerable proportion of the total consumption expenditure falls substantially, consumers' real income increases: they become relatively richer. Consequently, they substitute the superior good for the inferior ones, i.e., they reduce the consumption of inferior goods. Thus, the *income effect* on the demand for inferior goods becomes negative.

**(iii) Diminishing Marginal Utility**

*Diminishing marginal utility* is also responsible for increase in demand for a commodity when its price falls. When a person buys a commodity, he exchanges his money income with the commodity in order to maximise his satisfaction. He continues to buy goods and services so long as marginal utility of money ( $MU_m$ ) is less than marginal utility of the commodity ( $MU_c$ ). Given the price of a commodity, he adjusts his purchase so that  $MU_c = MU_m$ . This proposition holds good under both Marshallian assumption of constant  $MU_m$  and Hicksian assumption of diminishing  $MU_m$ . Under Marshallian approach,  $MU_m$  remaining constant,  $MU_c = P_c$  and a utility maximising consumer reaches his equilibrium where

$$MU_m = P_c = MU_c$$

When price falls,  $(MU_m = P_c) < MU_c$ . Thus, equilibrium condition is disturbed. To regain his equilibrium condition, i.e.,  $MU_m = P_c = MU_c$ , he purchases more of the commodity. For, when the stock of a commodity increases, its  $MU$  decreases and once again  $MU_m = MU_c$ . That is why demand for a commodity increases when its price decreases.

The law of demand does not apply to the following cases:

- (a) **Expectations regarding future price.** When consumers expect a continuous increase in the price of a durable commodity, they buy more of it despite increase in its price. They do so with a view to avoiding the pinch of still higher price in future. Similarly, when consumers anticipate a considerable decrease in the price in future, they postpone their purchases and wait for the price to fall to the expected level rather than buy the commodity when its price initially falls. Such decisions of the consumers are contrary to the law of demand.
- (b) **Status goods.** The law does not apply to the commodities which serve as a 'status symbol', enhance social prestige or display wealth and richness, e.g., gold, precious stones, rare paintings and antiques, etc. Rich people buy such goods mainly because their prices are high.
- (c) **Giffen goods.** An exception to this law is also the classic case of Giffen goods named after Robert Giffen (1837–1910). Giffen goods does not mean any specific commodity. It may be any inferior commodity much cheaper than its superior substitutes, consumed mostly by the poor households as an essential consumer good. If price of such goods increases (price of its substitute remaining constant), its demand increases instead of decreasing. For instance, let us suppose that the monthly minimum consumption of foodgrains by a poor household is 30 kgs including 20 kgs of bajra (an inferior good) at the rate of Rs 10 per kg and 10 kgs of wheat (a superior good) at Rs. 20 per kg. It spends a fixed amount of Rs. 400 on these items. Now, if price of bajra increases to Rs. 12 per kg the household will be forced to reduce the consumption of wheat by 5 kgs and increase that of bajra by the same quantity in order to meet its minimum monthly consumption requirement within Rs. 400. Obviously, household's demand for bajra increases from 20 to 25 kgs when its price increases.

---

### 4.3 DETERMINANTS OF MARKET DEMAND

---

The market demand for a product is determined by a number of factors, viz., price of the product, price and availability of the substitutes, consumer's income, his own preference for a commodity, utility derived from the commodity, 'demonstration effect', advertisement, credit facility by the sellers and banks, off season discounts, number of the uses of the commodity, population of the country, consumer's expectations regarding the future trend in the price of the product, consumers' wealth, past levels of demand, past levels of income, government policy, etc. But all these factors are not equally important. Besides, some of these



factors are not quantifiable, e.g., consumer's preferences, utility, demonstration effect, expectations, etc. and hence are not usable in the demand estimation. Nevertheless, we will discuss here how some important quantifiable and non-quantifiable determinants determine the market demand for a product.

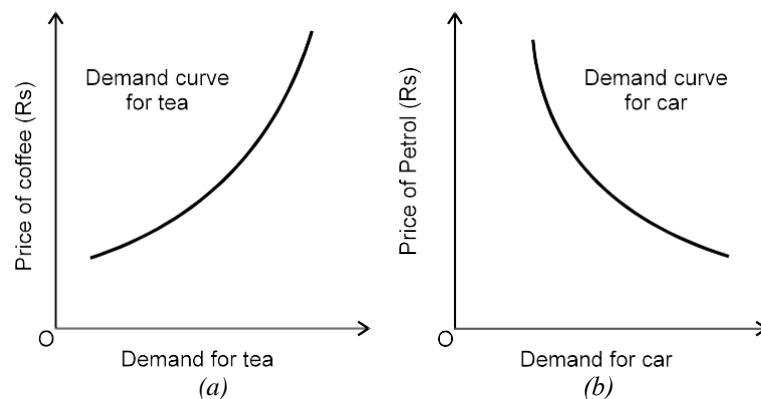
### 1. Price of the Commodity

As stated above, price is the most important determinant of the quantity demanded of a commodity. The price-quantity relationship is the central theme of demand theory. The nature of relationship between price of a commodity and its quantity demanded has already been discussed under the 'Law of Demand'. It may be added that the price-quantity relationship shows the movement along the *same* demand curve whereas other determinants of demand (discussed below) cause shift in the demand curve.

### 2. Price of Substitutes and Complementary Goods

The demand for a commodity depends also on the prices of *its substitutes and complementary goods*. Two commodities are deemed to be *substitutes* for each other if change in the price of one affects the demand for the other in the same direction. For instance, commodities *X* and *Y* are, in economic sense, *substitutes* for each other if a rise in the price of *X* increase the demand for *Y*, and *vice versa*. Tea and coffee, hamburger and hot-dog, alcohol and drugs are some common examples of substitutes.

By definition, the relationship between demand of a product (say, tea) and the price of its substitute (say, coffee) is positive in nature. When price of the substitute (coffee) of a product (tea) falls (or increases), demand for the product falls (or increases). The relationship of this nature is given in Fig. 4.5(a).



**Fig. 4.5** Demand for Substitutes and Complements

A commodity is deemed to be a *complement* of another when it complements the use of the other. In other words, when the use of any two goods goes together so that their demand changes (increases or decreases) simultaneously,

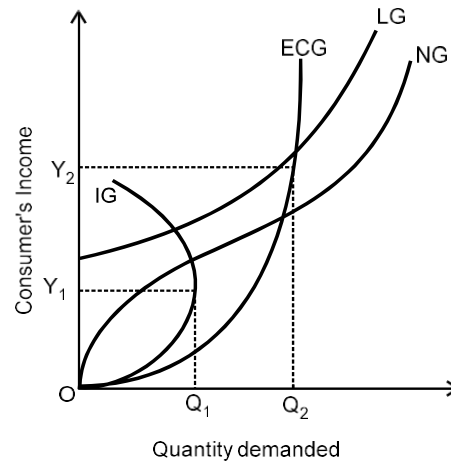
they are treated as complements. For example, petrol is a complement to motor vehicles; butter and jam are complements to bread; milk and sugar are complements to tea and coffee. Technically, two goods are complements to one another if an increase in the price of one causes a decrease in the demand for another. By definition, there is an inverse relationship between the demand for a good and the price of its complement. For instance, an increase (or decrease) in the price of petrol causes a decrease (or an increase) in the demand for car, other things remaining the same. The nature of relationship between the demand for a product and the price of its complement is given in Fig. 4.5(b).

### 3. Consumer's Income

Income is the basic determinant of the quantity demanded of a product as it determines the purchasing power of the consumer. That is why the people with higher current disposable income spend a larger amount on normal goods and services than those with lower incomes. Income-demand relationship is of a more varied nature than that between demand and its other determinants.

For the purpose of income-demand analysis, goods and services may be grouped under four broad categories, viz., (a) essential consumer goods; (b) inferior goods; (c) normal goods, and (d) prestige or luxury goods. The relationship between income and the different kinds of consumer goods is presented through the Engel Curves in Fig. 4.6.

- (a) **Essential Consumer Goods (ECG).** The goods and services which fall in this category are consumed, as a matter of necessity, by almost all persons of a society, e.g., foodgrains, salt, vegetable oils, matches, cooking fuel, a minimum clothing and housing, etc. Quantity demanded of such goods increases with increase in consumer's income only upto a certain limit, other factors remaining the same. The relation between demand of this category and consumer's income is shown by curve *ECG* in Fig. 4.6. As the curve shows, consumer's demand for essential goods increase until his income rises to  $OY_2$  and beyond this level of income, it does not.
- (b) **Inferior Goods.** Inferior and superior goods are generally known to the consumers by and large. For instance, every consumer knows that *bajra* is inferior to wheat and rice; *bidi* (an indigenous cigarette) is inferior to cigarette, cars without AC are inferior to AC cars, kerosene-stove is inferior to gas-stove; travelling by bus is inferior to travelling by taxi, and so on. In economic terminology, however, a commodity is deemed to be inferior if its demand decreases with the increase in consumers' income. The relation between income and demand for an inferior good is shown by curve *IG* in Fig. 4.6 under the assumption that other determinants of demand remain the same. Demand for such goods may initially increase with increase in income (say, upto  $Y_1$ ) but it decreases when income increases beyond a certain level.



**Fig. 4.6** Income-Demand Curves

- (c) **Normal Goods.** Technically, normal goods are those which are demanded in increasing quantities as consumers' income rises. Clothing is the most important example of this category of goods. The nature of relation between income and demand for the normal goods is shown by curve *NG* in Fig. 4.6. As the curve shows, demand for such goods increases with the increase in income of the consumer, but at different rates at different levels of income. Demand for normal goods initially increases rapidly, and later, at a lower rate. With the increase in the consumers' income, its income-elasticity decreases.

It may be noted from Fig. 4.6 that upto a certain level of income ( $Y_1$ ) the relation between income and demand for all types of goods is positive. While demand for some *NG*<sub>s</sub> increases at a rate, for some at a low rate. The difference is of degree only. The income-demand relationship becomes distinctly different beyond the level of income  $Y_1$ .

- (d) **Prestige or Luxury Goods.** Prestige goods are those which are consumed mostly by the rich section of the society, e.g., luxury cars, stone studded jewellery, costly cosmetics, decoration items (like antiques), etc. Demand for such goods arises only beyond a certain level of consumer's income. The income-demand relationship of this category of goods is shown by the curve *LG* in Fig. 4.6.

#### 4. Consumers' Taste and Preference

Consumers' taste and preferences play an important role in determining the demand for a product. Taste and preferences depend, generally, on the social customs, religious values attached to a commodity, habits of the people, the general life-style of the society, and also the age and sex of the consumers. Changes in these factors change consumers' taste and preferences. As a result, consumers reduce or give up the consumption of some goods and include some others in their consumption basket. Generally, if consumers' liking, taste and preference for certain

goods and services change following the change in fashion, people switch their consumption pattern from cheaper, old fashioned goods over to costlier 'mod' goods, so long as price differentials commensurate with their preference. Consumers are prepared to pay higher prices for 'mod' goods even if their virtual utility is the same as that of old-fashioned goods. This fact reveals that tastes and preferences also influence demand for goods and services.

### 5. *Expected Utility at Equilibrium*

Most consumers have limited income to satisfy their unlimited wants. They spend their income on the various goods they consume in such a manner that the total satisfaction derived from the limited income is maximum. A consumer maximises his total satisfaction or his total utility when marginal utility (per unit of expenditure) derived from each commodity is the same. For example, let us suppose that a consumer has to spend his limited income on bread ( $b$ ), shirts ( $s$ ), and cinema shows ( $c$ ). Given their respective price as  $P_b$ ,  $P_s$  and  $P_c$ , he would spend his income on these items so that marginal utility ( $MU$ ) per unit of expenditure from each of these goods is the same. That is, the consumer spends his income in such a way that

$$\frac{MU_b}{P_b} = \frac{MU_s}{P_s} = \frac{MU_c}{P_c}$$

where  $MU_b$ ,  $MU_s$  and  $MU_c$  denote the  $MU$  of bread, shirts and cinema shows, respectively. This is a necessary condition of consumer's equilibrium. Since  $MU$  schedule for each of these goods is bound to be different, a rational consumer buys different quantities of these goods with a view to equalising their  $MU$  per unit of expenditure. The equilibrium condition itself determines the quantity of each goods (given their  $MU$  schedule) which a utility maximising consumer would like to buy. Although, in practice, a consumer may not be able to achieve the theoretical precision of his equilibrium, his pattern of expenditure and the quantity of each commodity that he buys would approximate to the equilibrium condition stated above.

### 6. *Consumers' Expectations*

Consumers' expectations regarding the future course of economic events, particularly regarding changes in prices, income, and supply position of goods, play an important role in determining the demand for goods and service in the short-run. As mentioned above, if consumers expect a rise in the price of a commodity, they tend to buy more of it at its current price with a view to avoiding the pinch of price-rise in future. For example, when the automobile owners expect or Government of India announces rise in petrol and diesel prices from a future date, automobile owners buy more of petrol and diesel at their current prices. On the contrary, if consumers expect a fall in the price of certain goods, they postpone their purchases of such goods with a view to taking advantage of lower prices in

future, mainly in case of non-essential goods. This behaviour of consumers reduces the current demand for the goods whose prices are expected to decrease in future.

Similarly, an expected increase in income on account of announcement of revision of pay-scales, dearness allowance, bonus, etc., induces increase in current purchase, and *vice versa*. Besides, if consumers or users expect scarcity of certain goods in future on account of reported fall in future production, labour strikes on a large scale, diversion of civil supplies towards the military use, etc., the current demand for such goods would increase, more so if their prices show an upward trend. Consumers demand more for future consumption; profiteers demand more to make money out of expected scarcity. In simple words, expectation regarding the shortage of a commodity in future increases its current demand at the prevailing price.

### *7. Demonstration Effect*

When new commodities or new models of existing ones appear in the market, rich people buy them first. Some people buy new goods or new model of goods because they have genuine need for them while others buy because they want to exhibit their affluence. Fashion goods make the most common case for this category of goods. But once new commodities come in vogue, many households buy them not because they have a genuine need for them but because others or neighbours have bought these goods. The purchases by the latter category of the buyers are made out of such feelings as jealousy, competition, equality in the peer group, social inferiority and the desire to raise their social status. Purchases made on account of these factors are the result of 'Demonstration Effect' or the 'Band-wagon Effect'. These effects have a positive effect on the demand. On the contrary, when a commodity becomes the thing of common use, some people, mostly rich, decrease or give up the consumption of such goods. This is known as 'Snob Effect'. It has a negative effect on the demand for the related goods.

### *8. Consumer-Credit Facility*

Availability of credit to the consumers from the sellers, banks, relations and friends or from any other source encourages the consumers to buy more than what they would buy in the absence of credit facility. That is why the consumers who can borrow more can consume more than those who can borrow less. Credit facility affects mostly the demand for consumer durables, particularly those which require bulk payment at the time of purchase.

### *9. Population of the Country*

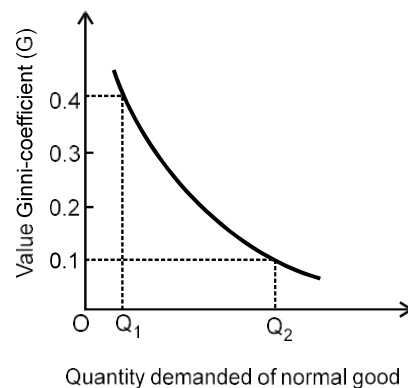
The total domestic demand for a product depends also on the size of the population. Given the price, per capita income, taste and preferences, etc. the larger the

population, the larger the demand for a product. With an increase (or decrease) in the size of population, employment percentage remaining the same, demand for the product will increase (or decrease). The relation between market demand for essential and normal goods and the size of population is similar to the income-demand relation.

### 10. Distribution of National Income

Apart from the level of individual incomes, the distribution pattern of national income also affects the demand for a commodity. If national income is evenly distributed, market demand for normal goods will be the largest. If national income is unevenly distributed, i.e., if majority of population belongs to the lower income groups, market demand for essential goods (including inferior ones) will be the largest whereas the same for other kinds of goods will be relatively low. Furthermore, given a distribution of national income and a market demand for various types of goods, if national income gets distributed in favour of rich so that this section becomes smaller, the demand for essential goods will increase and the same for other kinds of goods will decrease and *vice versa*, as far as Engel's Law holds.

The relationship between market demand for a normal good and national income distribution is illustrated in Fig. 4.7. In the figure, vertical axis measures the Gini coefficient (a measure of national income distribution- $G$ ) and the horizontal axis measures the quantity demanded of a normal good. As Fig. 4.7 shows, as  $G$  decreases from 0.4 to 0.1, (i.e., income distribution becomes more and more even) quantity demanded of a normal good increases from  $Q_1$  towards  $Q_2$ .



**Fig. 4.7** Gini-coefficient and Demand

---

#### 4.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

---

1. A demand curve is a locus of points showing various alternative price-quantity combinations.
2. The law of demand does not apply to the commodities which serve as a 'status symbol', enhance social prestige or display wealth and richness, e.g., gold, precious stones, rare paintings and antiques, etc. Rich people buy such goods mainly because their prices are high.
3. The total quantity which all the consumers of a commodity are willing to buy at a given price per time unit, other things remaining the same, is known as market demand for the commodity.

---

#### 4.5 SUMMARY

---

- The term 'demand' refers to the quantity demanded of a commodity per unit of time at a given price. It implies also a desire backed by ability and willingness to pay. A mere desire of a person to purchase a commodity is not his demand.
- The law of demand states that quantity of a product demanded per unit of time increases when its price falls, and decreases when its price increases, other factors remaining constant.
- The assumption 'other factors remaining constant' implies that income of the consumers, prices of the substitutes and complementary goods, consumers' taste and preference, and number of consumers, remain unchanged.
- The law of demand can also be presented through a curve called demand curve. A demand curve is a locus of points showing various alternative price-quantity combinations.
- When demand curve changes its position retaining its shape (though not necessarily), the change is known as shift in demand curve.
- Deriving demand curve is the ultimate aim of the entire utility analysis—cardinal or ordinal. The basic purpose of the entire exercise in indifference curve technique so far is to construct the individual demand curve for a commodity.
- When the price of a commodity falls it becomes relatively cheaper if price of all other related goods, particularly of substitutes, remain constant or, in

other words, substitute goods become relatively costlier. Since consumers substitute cheaper goods for costlier ones, demand for the relatively cheaper commodity increases. The increase in demand on account of this factor is known as substitution effect.

- Diminishing marginal utility is also responsible for increase in demand for a commodity when its price falls.
- The quantity of a commodity which an individual is willing to buy at a particular price of the commodity during a specific time period, given his money income, his taste, and prices of substitutes and complements, is known as individual demand for a commodity.
- Demand function states the relationship between demand for a product (the dependent variable) and its determinants (the independent variables).

---

#### 4.6 KEY WORDS

---

- **Function:** In mathematical language, a function is a symbolic statement of relationship between a dependent and the independent variables.
- **Giffen good:** In economic theory, it is a good that is in greater demand as its price increases.
- **Disposable income:** It refers to income remaining after deduction of taxes and social security charges, available to be spent or saved as one wishes.

---

#### 4.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

##### *Short-Answer Questions*

1. Write a short note on demand curve and demand schedule.
2. Briefly discuss the reasons for shift in demand curve.
3. What do you understand by income effect?
4. Explain demonstration effect.

##### *Long-Answer Questions*

1. Define the law of demand. Discuss the reasons and exceptions to the law of demand.
2. Give the derivation of individual demand curve.
3. Give a detailed explanation of determinants of market demand.



---

## 4.8 FURTHER READINGS

---

Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.

Keat Paul G. and K.Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.

Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition. Singapore: Pearson Education, Inc.

Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Publishers.

---

### Structure

- 5.0 Introduction
- 5.1 Objectives
- 5.2 Definition of Elasticity of Demand
  - 5.2.1 The Uses of Elasticity
- 5.3 Types of Elasticity of Demand
  - 5.3.1 Price Elasticity of Demand
  - 5.3.2 Cross Elasticity of Demand
  - 5.3.3 Income Elasticity of Demand
  - 5.3.4 The Elasticity of Price Expectations
- 5.4 Answers to Check Your Progress Questions
- 5.5 Summary
- 5.6 Key Words
- 5.7 Self Assessment Questions and Exercises
- 5.8 Further Readings

---

## 5.0 INTRODUCTION

---

Elasticity refers to the degree of responsiveness in supply or demand in relation to changes in price. If a curve is more elastic, then small changes in price will cause large changes in quantity consumed. If a curve is less elastic, then it will take large changes in price to reflect a change in quantity consumed.

This unit sheds light on the concept of elasticity of demand, its definition and types. Concepts such as elasticity of price expectations, income elasticity of demand and uses of cross-elasticity are also discussed in this unit.

---

## 5.1 OBJECTIVES

---

After going through this unit, you will be able to:

- Define elasticity of demand
- Discuss the uses and types of elasticity of demand
- Describe cross elasticity and income elasticity of demand
- Explain elasticity of price expectations

---

## 5.2 DEFINITION OF ELASTICITY OF DEMAND

---

The theory of demand, discussed in the preceding unit, states the *direction of change in demand* due to change in its determinants. For example, according to the theory of demand, all other factors remaining constant, when price of a product

increases, demand for the product decreases and *vice versa*. The theory of demand does not tell ‘demand decreases by how much due to a certain increase in price’. However, the knowledge of direction or kind of change in demand due to change in its determinant is not sufficient from the angle of price management. What is more important is the *extent of relationship* between the demand for a product and its determinants. For example, suppose a firm reduces the price of its product from ₹10 to ₹8 with the objective of increasing demand for its product and its total revenue. In effect, however, total revenue decreases from ₹1000 to ₹880. The reason is that demand increased by a lower percentage (10%) than the percentage of decrease in price (20%).

Look at another case of pricing decision issue. When price of the substitute of a firm’s product increases, the question arises: Should the firm change its own price or keep the price unchanged? Similarly, if a firm plans to advertise its product, a question arises: Will the sales promotion yield sales revenue in excess of advertisement cost? The theory of demand, as such, does not provide answer to these questions. Answer to “these questions lies in the *degree of responsiveness* of demand to a change in its determinants. The degree of responsiveness of demand for product to change in its price is called the ***elasticity of demand***.

This unit presents a detailed discussion on the concept of the elasticity of demand, the various methods of measuring demand, and its application to business decisions. The concepts of elasticity of demand generally used in business decisions are:

- (i) Price elasticity of demand,
- (ii) Cross elasticity of demand,
- (iii) Income elasticity of demand,
- (iv) Advertisement elasticity of demand, and
- (v) Elasticity of price expectations.

Before we discuss these elasticities of demand, let us understand the concept of demand elasticity and the method of its measurement.

#### *Definition and measurement of elasticity concept*

In general terms, the *elasticity of demand* is defined as *the degree of responsiveness* of demand for a product to change in its determinants. The measure of the degree of responsiveness of demand to change in its determinants gives the measure of the *extent of relationship* between the demand for a product and any of its determinants. In technical terms, the measure of elasticity of demand is called *elasticity coefficient* measured by the following formula:

$$E_d = \frac{\text{Percentage Change in Quantity Demanded of Product } X}{\text{Percentage Change in Demand Determinant Factor } Y}$$

For instance, suppose a determinant of demand for a product changes by 10 percent and, as a result, demand changes by 15 percent. In that case, the elasticity coefficient equals  $15/10 = 1.5$ .

The general formula for measuring the elasticity of demand can be expressed as follows.

$$E_d = \frac{\Delta Q/Q}{\Delta Y/Y} = \frac{\Delta Q}{\Delta Y} \times \frac{Y}{Q} \quad \dots(5.1)$$

where  $Q$  = quantity demanded initially;  $\Delta Q$  = change in demand,  $Y$  = the original value of demand determinant factor; and  $\Delta Y$  = change in the determinant factor.

### 5.2.1 The Uses of Elasticity

Although Samuelson condemned the concept of elasticity as an ‘essentially arbitrary’ and a more or less useless concept, it has many important uses in both economic analysis and formulation of economic policies. Some important uses of elasticity concept are as follows.

**First**, the concept of elasticity of demand plays a crucial role in business decisions regarding manoeuvring of prices with a view to making larger profits. For instance, when cost of production is increasing, the firm would like to pass incremental cost on to the consumer by raising the price. Firms may decide to change the price even without change in cost of production. But, whether this action (raising the price following the rise in cost or otherwise) will prove beneficial or not depends on (a) the price-elasticity of demand for the products; and (b) its cross-elasticity because when the price of a product increases, its substitutes become automatically cheaper even if their prices remain unchanged. Raising price will be beneficial only if (i) demand of a product has an elasticity less than one, and (ii) demand for its substitute is much less elastic. Similarly, a firm not finding it feasible to increase the price during a period of growth in consumer’s income would like to increase the production. The additional production can be determined only through the income elasticity of demand for the product, other factors remaining the same. Although most businessmen intuitively are aware of the elasticity of demand of the goods, they do make use of precise estimates of elasticity of demand as it adds precision to the business decisions.

**Second**, the elasticity concept can be used also in formulating government policies, particularly in its taxation policy meant to raise revenue or to control prices; in granting subsidies to the industries; in determining prices for public utilities; in fixing the prices of essential goods; and in determining export and import duties and the rate of currency devaluation. To consider an example, suppose government wants to impose sales tax on a particular commodity with the sole objective of raising revenue. Whether adequate revenue can be raised or not will depend on the price elasticity of that commodity. If demand is highly elastic, the revenue yield will be much less than expected. The sales tax will rather cause price distortion and affect production adversely.

**Third**, the concept of elasticity is useful in economic analysis, at least for specifying the relationship between the dependent and independent variables. Besides, the elasticity concept is used in specifying and estimating demand functions.

The most commonly used form of demand function in applied research is the “constant elasticity demand function” of the form.

$$Q_X = A P_X^B Y^C P^D E^{FT}$$

in which  $P_X$ ,  $Y$ ,  $P^D$  and  $E^{FT}$  represent, respectively, price of  $X$ , consumer’s income, price of other goods and a trend factor of ‘taste’, and if arc elasticity coefficients, and  $A$  is a constant.

To sum up, elasticity concept is undoubtedly a useful concept and has a wide application to economic analysis and policy.

### 5.3 TYPES OF ELASTICITY OF DEMAND

In this section, we will study the different types of elasticity of demand.

#### 5.3.1 Price Elasticity of Demand

*Price elasticity of demand is generally defined as the responsiveness or sensitiveness of demand for a commodity to the changes in its price. More precisely, elasticity of demand is the percentage change in demand due to one per cent change in the price of the commodity. A formal definition of price elasticity of demand ( $e_p$ ) is given as*

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

A general formula for calculating coefficient of price elasticity, is given as follows:

$$\begin{aligned} e_p &= \frac{\frac{\Delta Q}{Q}}{\frac{\Delta P}{P}} = \frac{\Delta Q}{Q} \times \frac{P}{\Delta P} \\ &= \frac{\Delta Q}{\Delta P} \times \frac{P}{Q} \end{aligned} \quad \dots(5.2)$$

where  $Q$  = original quantity demanded,  $P$  = original price,  $\Delta Q$  = change in quantity demanded and  $\Delta P$  = change in price.

It is important to note here that a minus sign (–) is generally inserted in the formula before the fraction in order to make the elasticity coefficient a non-negative value.

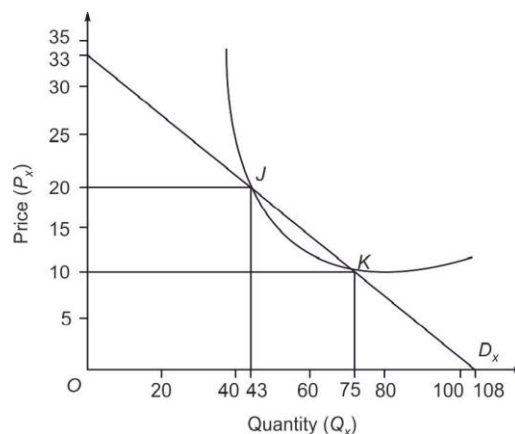
The price elasticity can be measured between any two points on a demand curve (called *arc elasticity*) or at a point (called *point elasticity*). The measurement of the two kinds of price elasticity are discussed below in detail.

### Arc elasticity

The concept of arc elasticity of demand refers to the measurement of demand elasticity for a significant change in price and consequent change in demand. When there is a substantial or a big change in price, then the demand – price point shifts from one point to another on the demand curve. Thus, the measure of elasticity of demand between any two finite points on a demand curve is known as **arc elasticity**. For example, measure of elasticity between points *J* and *K* (Fig. 5.1) is the measure of arc elasticity. The movement from point *J* to *K* on the demand curve ( $D_x$ ) shows a big (50%) fall in the price from 20 to 10 so that  $\Delta P = 20 - 10 = 10$ . The big fall in price causes a large increase in demand from 43 units to 75 units so that  $\Delta Q = 75 - 43 = 32$ . The arc elasticity between points *J* and *K* (moving from *J* to *K*) can be calculated by substituting these values into the elasticity formula as follows:

$$e_p = - \frac{-\Delta Q}{\Delta P} \cdot \frac{P}{Q} \quad (\text{with minus sign})$$

$$= - \frac{-32}{10} \cdot \frac{20}{43} = 1.49 \quad \dots(5.3)$$



**Fig. 5.1** Linear Demand Curve

This means that a one per cent decrease in price of commodity *X* results in a 1.49 per cent increase in demand for it.

**Problem in Using Arc Elasticity** The arc elasticity should be measured and used carefully, otherwise it may lead to wrong decisions. Arc elasticity coefficients differ between the same two finite points on a demand curve if *direction* of change in price is reversed. For instance, as estimated in Eq. (5.3), the elasticity between points *J* and *K*—moving from *J* to *K* equals 1.49. It may be wrongly interpreted that the elasticity of demand for commodity *X* between points *J* and *K* equals 1.49

irrespective of the direction of price change. But it is not true. A reverse movement in the price, i.e., the movement from point  $K$  to  $J$  implies a different elasticity coefficient (0.43). Movement from point  $K$  to  $J$  gives  $P = 10$ ,  $\Delta P = 10 - 20 = -10$ ,  $Q = 75$  and  $\Delta Q = 75 - 43 = 32$ . By substituting these values into the elasticity formula, we get

$$e_p = - \frac{32}{-10} \cdot \frac{10}{75} = 0.43 \quad \dots(5.4)$$

The measure of elasticity coefficient in Eq. (5.4) for the reverse movement in price is obviously different from one given by Eq. (5.3). It means that *the elasticity depends also on the direction of change* in price. Therefore, while measuring price elasticity, the direction of price change should be carefully noted. Otherwise, it will lead to a wrong decision regarding the change in price. For instance, if price elasticity between points  $J$  and  $K$  is taken to be the same whether price increases or decreases, it leads to the conclusion that total sales revenue will remain the same whether price increases or decreases. But, this is a wrong conclusion. The movement from point  $J$  to  $K$  yields a sales revenue of ₹  $10 \times 75 = ₹750$ . But movement from point  $K$  to  $J$  yields a sales revenue of ₹  $20 \times 43 = ₹860$ . It means increasing price is beneficial and decreasing price is harmful.

**Some Modifications** Some modifications have been suggested in economic literature to resolve the problems associated with arc elasticity.

**First**, the problem arising due to the change in the direction of price change may be avoided by using the lower values of  $P$  and  $Q$  in the elasticity formula. In that case,

$$e_p = \frac{\Delta Q}{\Delta P} \cdot \frac{P_l}{Q_l}$$

where  $P_l = 10$  (the lower of the two prices) and  $Q_l = 43$  (the lower of the two quantities). Thus,

$$e_p = - \frac{32}{-10} \cdot \frac{10}{43} = 0.74 \quad \dots(5.5)$$

This method is however devoid of the logic of calculating percentage change in price and demand because the choice of lower values of  $P$  and  $Q$  is arbitrary—it is not in accordance with the rule of calculating percentage change.

**Second**, another method suggested to resolve this problem is to use the average of upper and lower values of  $P$  and  $Q$  in fraction  $P/Q$ . In that case the formula is

$$e_p = \frac{\Delta Q}{\Delta P} \cdot \frac{(P_1 + P_2)/2}{(Q_1 + Q_2)/2}$$

or

$$e_p = \frac{Q_2 - Q_1}{P_2 - P_1} \cdot \frac{(P_1 + P_2)/2}{(Q_1 + Q_2)/2} \quad \dots(5.6)$$

where subscripts 1 and 2 denote lower and upper values of prices and quantities.

By substituting the values from our example, we get,

$$e_p = - \frac{75 - 43}{10 - 20} \cdot \frac{(20 + 10)}{(43 + 75)} = 0.81$$

This method too has its own drawbacks as the elasticity co-efficient calculated through this formula refers to the elasticity mid-way between  $P_1 P_2$  and  $Q_1 Q_2$ . The elasticity co-efficient (0.81) is not applicable for the whole range of price-quantity combinations at different points between  $J$  and  $K$  on the demand curve (Fig. 5.1)—it only gives a mean of the elasticities between the two points.

### Point elasticity

**Point elasticity on a linear demand curve.** Point elasticity is also a way to resolve the problem in measuring the elasticity. The concept of point elasticity is used for measuring price elasticity where change in price is infinitesimally small.

*Point elasticity* is the elasticity of demand at a finite point on a demand curve, e.g., at point  $P$  or at point  $B$  on the linear demand curve  $MN$  in Fig. 5.2. This is in contrast to the arc elasticity between points  $P$  and  $B$ . A movement from point  $B$  towards point  $P$  implies change in price ( $\cdot P$ ) becoming smaller and smaller, such that point  $P$  is almost reached. Here the change in price is infinitesimally small. Measuring elasticity for an infinitesimally small change in price is the same as measuring elasticity at a point. The formula for measuring point elasticity is given below.

$$\text{Point elasticity } (e_p) = \frac{\partial Q}{\partial P} \cdot \frac{P}{Q} \quad \dots(5.7)$$

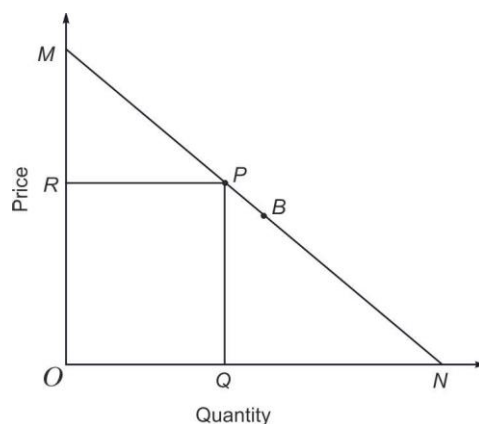


Fig. 5.2 Point Elasticity

Note that  $\frac{\partial Q}{\partial P}$  has been substituted for  $\frac{\Delta Q}{\Delta P}$  in the formula for arc elasticity.

The derivative  $\frac{\partial Q}{\partial P}$  is reciprocal of the slope of the demand curve  $MN$ . Point elasticity is thus the product of price-quantity ratio at a particular point on the



demand curve and the reciprocal of the slope of the demand line. The reciprocal of the slope of the straight line  $MN$  at point  $P$  is geometrically given by  $QN/PQ$ . Therefore,

$$\frac{\Delta Q}{\Delta P} = \frac{QN}{PQ}$$

Note that at point  $P$ , price  $P = PQ$  and quantity demand ( $Q$ ) =  $OQ$ . By substituting these values in Eq. (5.7), we get

$$e_p = \frac{PQ}{OQ} \cdot \frac{QN}{PQ} = \frac{QN}{OQ}$$

Given the numerical values for  $QN$  and  $OQ$ , elasticity at point  $P$  can be easily obtained. We may compare here the arc elasticity between points  $J$  and  $K$  and point elasticity at point  $J$  in

Fig. 5.1. At point  $J$ ,

$$e_p = \frac{QN}{OQ} = \frac{108 - 43}{43} = 1.51$$

Note that point elasticity,  $e_p = 1.51$ , is different from various measures of arc elasticities (i.e.,  $e_p = 1.49$ ,  $e_p = 0.43$ ,  $e_p = 0.7$ , and  $e_p = 0.81$ ).

As has been proved below, geometrically,  $QN/OQ = PN/PM$ . Therefore, elasticity of demand at point  $P$  (Fig. 5.2) may be expressed as

$$e = \frac{PN}{PM}$$

**Proof.** The fact that  $e_p = QN/OQ = PN/PM$  can be proved as follows. Note that in Fig. 5.2, there are three triangles— $\triangle MON$ ,  $\triangle MRP$  and  $\triangle PQN$ —and  $\angle MON$ ,  $\angle MRP$  and  $\angle PQN$  are right angles. Therefore, the other corresponding angles of the three triangles will always be equal and hence,  $\triangle MON$ ,  $\triangle MRP$  and  $\triangle PQN$  are similar.

According to geometrical properties of similar triangles, the ratio of any two sides of a triangles are always equal to the ratio of the corresponding sides of the other triangles. By this rule, between  $\triangle PQN$  and  $\triangle MRP$ ,

$$\frac{QN}{PN} = \frac{RP}{PM} \quad \dots(5.8)$$

Since  $RP = OQ$ , by substituting  $OQ$  for  $RP$ , Eq. (5.8) can be expressed as

$$\frac{QN}{PN} = \frac{OQ}{PM}$$

It follows that

$$\frac{QN}{OQ} = \frac{PN}{PM}$$

It means that price elasticity of demand at point  $P$  in Fig. 5.2 is given by

$$e_p = \frac{PN}{PM}$$

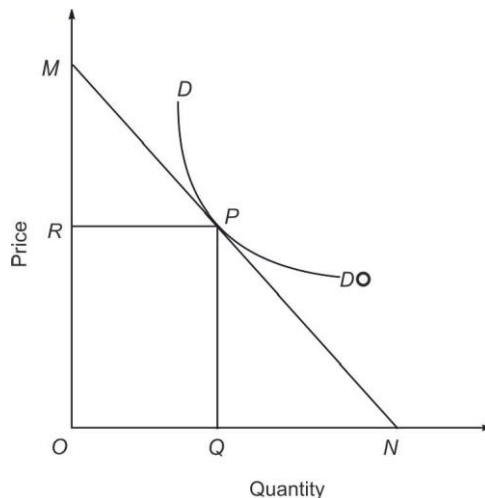
It may thus be concluded that the price elasticity of demand at any point on a linear demand curve is equal to the ratio of lower segment to the upper segments of the line, i.e.,

$$e_p = \frac{\text{Lower segment of Demand Curve}}{\text{Upper segment of Demand Curve}}$$

**Point elasticity on a non-linear demand curve.** The ratio  $\cdot Q / \cdot P$  in respect of a non-linear demand curve is different at each point. Therefore, the method used to measure point elasticity on a linear demand curve cannot be applied straightaway to measure elasticity on a curvilinear demand curve. A simple modification in technique is required. In order to measure point elasticity on a non-linear demand curve, the chosen point is first brought on a linear demand curve. This is done by drawing a tangent through the chosen point. For example, suppose we want to measure elasticity on a non-linear demand curve,  $DD'$  (Fig. 5.3) at point  $P$ . For this purpose, a tangent  $MN$  is drawn through point  $P$ . Since demand curve  $DD'$  and the line  $MN$  pass through the same point ( $P$ ), the slope of the demand curve and that of the line at this point is the same. Therefore, the elasticity of demand curve at point  $P$  will be equal to that of the line at this point. Elasticity of the line at point  $P$  can be measured as

$$e_p = \frac{P}{Q} \cdot \frac{\partial P}{\partial P} = \frac{PQ}{OQ} \cdot \frac{QN}{PQ} = \frac{QN}{OQ}$$

As proved above, geometrically  $= \frac{QN}{OQ} = \frac{PN}{PM}$



**Fig. 5.3** Non-linear Demand Curve

To conclude, at midpoint of a linear demand curve,  $e_p = 1$ . Note that in Fig. 5.4, point  $P$  falls on the mid point of demand curve  $MN$ . At point  $P$ , therefore,  $e = 1$ . It follows that at any point above the point  $P$ ,  $e_p > 1$ , and at any point below the point  $P$ ,  $e_p < 1$ . According to this formula, at the extreme point  $N$ ,  $e_p = 0$ , and at extreme point  $M$ ,  $e_p$  is undefined because division by zero is undefined. It must be noted here that these results are relevant between points  $M$  and  $N$ .

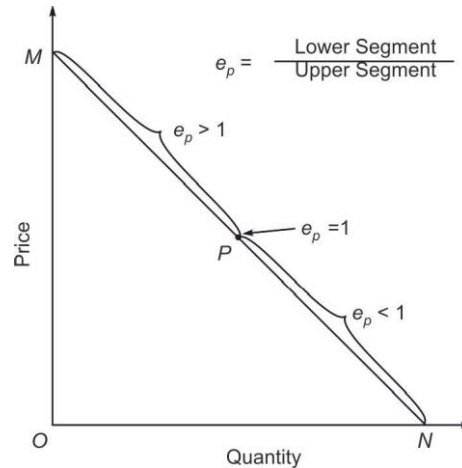


Fig. 5.4 Point Elasticities of Demand

### 5.3.2 Cross Elasticity of Demand

The cross-elasticity is the measure of responsiveness of demand for a commodity to the changes in the price of its substitutes and complementary goods. For instance, cross-elasticity of demand for tea is the percentage change in its quantity demanded due to the change in the price of its substitute, coffee. The formula for measuring cross-elasticity of demand is the same as that of the price elasticity with a difference. For example, cross-elasticity of demand for tea ( $e_{t,c}$ ) can be measured by the formula given below.

$$e_{t,c} = \frac{\text{Percentage change in demand for tea } (Q_t)}{\text{Percentage change in price of coffee } (P_c)}$$

The cross-elasticity of demand for tea with respect price of coffee can be expressed technically as follows.

$$= \frac{P_c}{Q_t} \cdot \frac{\Delta Q_t}{\Delta P_c} \quad \dots(5.9)$$

Similarly, cross-elasticity of demand for coffee with respect to change in the price of tea is measured as follows.

$$e_{c,t} = \frac{P_t}{Q_c} \cdot \frac{\Delta Q_c}{\Delta P_t} \quad \dots(5.10)$$

The same formula is used to measure the cross-elasticity of demand for a good with respect to a change in the price of its complementary goods. Electricity

to electrical gadgets, petrol to automobiles, butter to bread, sugar to tea and coffee, are the examples of complementary goods.

It is important to note that when two goods are substitutes for one another, their demand has positive cross-elasticity because increase in the price of one good increases the demand for its substitute. And, the demand for complementary goods has negative cross-elasticity, because increase in the price of a complementary good decreases the demand for the main *good*.

### Uses of Cross-Elasticity

The concept of cross-elasticity has both theoretical and practical uses.

**Theoretically**, an *important use* of cross-elasticity is to define substitute goods. If cross-elasticity between any two goods is positive, the two goods may be considered as substitutes of one another. Also, the greater the cross-elasticity, the closer the substitute. Similarly, if cross-elasticity of demand for two related goods is negative, the two may be considered as complementary of one another: the higher the negative cross-elasticity, the higher the degree of complementarity.

**Practically**, the concept of cross-elasticity is of vital importance in pricing decisions, i.e., in changing prices of products having substitutes and complementary goods. If cross-elasticity in response to the price of substitutes is greater than one, it would be inadvisable to increase the price; rather, reducing the price may prove beneficial. In case of complementary goods also, reducing the price may be helpful in maintaining the demand in case the price of the complementary good is rising. Besides, if accurate measures of cross-elasticities are available, the firm can forecast the demand for its product and can adopt necessary safeguards against fluctuating prices of substitutes and complements.

The application of cross elasticity of demand, especially with regard to substitute good, can be explained clearly with an example. Suppose two firms, *A* and *B*, produce two goods *X* and *Y*, respectively, both the goods being substitutes for one another. The cross elasticity of demand for good *X* has been estimated at 1.5 and for good *Y* at 1.25. Given the cross elasticities, suppose firms intend to go for price competition. A question arises here: Is it advantageous for the firms to go for price competition?

Answer to this question can be found by assessing the effect of price change on the demand for their products. Suppose firm *A* cuts down its price by 10 per cent. As a result, the demand for *B*'s product *Y* decreases by  $10 \times 1.5 = 15$  per cent, as its demand get shifted to good *X*. Now, let the firm *B* react and cut down its price by 10 per cent. As a result, demand for good *X* decreases by  $10 \times 1.25 = 12.5$  percent as the substitution effect. It means firm *B* regains 12.5 per cent of its lost market (15 per cent). But firm *A* has still a gain of  $15\% - 12.5\% = 2.5$  percent. The final conclusions that emerge from this analysis are (i) firm *A* may initiate the price competition but not the firm *B*, and (ii) if firm *A* reduces its price, firm *B* has to react by reducing its own price, by a higher rate, if necessary.

### 5.3.3 Income Elasticity of Demand

Apart from price of a product and its substitutes, another important determinant of demand for a product is consumer's income. As noted earlier, the relationship between demand for normal and luxury goods and consumer's income is of positive nature, unlike the negative price-demand relationship. That is, the demand for normal goods and services increases with increase in consumer's income and *vice versa*. The responsiveness of demand to the change in consumer's income is known as **income elasticity** of demand.

Income elasticity of demand for a product, say  $X$  (i.e.,  $e_y$ ) is defined as

$$e_y = \frac{\Delta X_q / X_q}{\Delta Y / Y} = \frac{Y}{X_q} \cdot \frac{\Delta X_q}{\Delta Y} \quad \dots(5.11)$$

where  $X_q$  = quantity of  $X$  demanded;  $Y$  = disposable income;  $\Delta X_q$  = change in quantity demanded of  $X$ ; and  $\Delta Y$  = change in income.

Unlike price elasticity of demand (which is negative except in case of Giffen goods), income elasticity of demand is positive because of a positive relationship between income and demand for a product. There is an exception to this rule. Income elasticity of demand for an *inferior good* is negative, because of negative income-effect. The demand for inferior goods decreases with increase in consumer's income and *vice versa*. When income increases, consumers switch over to the consumption of superior commodities. That is, they substitute superior goods for inferior ones. For instance, when income rises, people prefer to buy more of rice and wheat and less of inferior foodgrains like *bajra*, *ragi*, and use more of taxi and less of bus service and so on.

#### *Nature of Commodity and Income Elasticity*

For all normal goods, income elasticity is positive though the degree of elasticity varies in accordance with the nature of commodities. Consumer goods are generally grouped under three categories, viz., necessities (essential consumer goods), comforts, and luxuries. The general pattern of income elasticities for goods of different categories for increase in income and their impact on sales are given in Table 5.1.

The income elasticity of demand for different categories of goods may however vary from household to household and from time to time, depending on choice, taste and preference of the consumers, levels of their 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 income increases regularly and frequently, income-elasticities as given in Table 5.1 will conform to the general pattern.

Some **important uses** of income elasticity are as follows:

**First**, the concept of income elasticity can be used to estimate the future demand provided the rate of increase in income and income elasticity of demand for the products are known. The knowledge of income elasticity can thus be useful in forecasting demand, when changes in personal incomes are expected, other things remaining the same.

**Table 5.1** *Income Elasticity of Different Consumer Goods*

Commodities	Coefficient of income elasticity	Impact on expenditure
1. Necessities	Less than unity ( $e_y < 1$ )	Less than proportionate change in income
2. Comforts	Almost equal to unity ( $e_y \approx 1$ )	Almost proportionate change in income
3. Luxuries	Greater than unity ( $e_y > 1$ )	More than proportionate increase in income

**Second**, the concept of income elasticity may also be used to define the ‘normal’ and ‘inferior’ goods. The goods whose income elasticity is positive for all levels of income are termed as ‘normal goods’. On the other hand, the goods for which income elasticities are negative, beyond a certain level of income, are termed as ‘inferior goods’.

#### 5.3.4 The Elasticity of Price Expectations

Sometimes, mainly during the period of price fluctuations, consumer’s price expectations play a much more important role in determining demand for a commodity than any other factor. The concept of price expectation elasticity refers to the expected change in price in future as a result of change in current prices of a product. The elasticity of price expectation is defined and measured through the following formula:

$$\begin{aligned} \eta_e &= \frac{\frac{\Delta P_f}{P_f}}{\frac{\Delta P_c}{P_c}} \\ &= \frac{\Delta P_f}{\Delta P_c} \cdot \frac{P_c}{P_f} \end{aligned} \quad \dots(5.12)$$

The coefficient,  $\eta_e$ , gives the measure of expected percentage change in future price ( $P_f$ ) as a result of 1 per cent change in current price ( $P_c$ ). If  $\eta_e > 1$ , it

indicates that future change in price will be greater than the present changes in price, and *vice versa*. If  $\eta_e = 1$ , it indicates that the future change in price will be

equal to the change in the current price. If  $\eta_e = 0$ , it indicates no change in future price as a result of change in current price. The concept of elasticity of price-expectation is very useful in formulating future pricing policy. For example if  $\eta_e > 1$ , it indicates sellers will be able to sell more in the future at higher prices. Thus, businessmen may accordingly determine their future pricing policy.

---

## 5.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

---

1. The measure of the degree of responsiveness of demand to change in its determinants is used for measuring the extent of relationship between the demand for a product and any of its determinants.
2. The concept of elasticity of demand plays a crucial role in business decisions regarding manoeuvring of prices with a view to making larger profits.
3. Price elasticity of demand is generally defined as the responsiveness or sensitiveness of demand for a commodity to the changes in its price.
4. The concept of arc elasticity of demand refers to the measurement of demand elasticity for a significant change in price and consequent change in demand.

---

## 5.5 SUMMARY

---

- The theory of demand, discussed in two preceding chapters, states the direction of change in demand due to change in its determinants. For example, according to the theory of demand, all other factors remaining constant, when price of a product increases, demand for the product decreases and vice versa.
- The degree of responsiveness of demand for product to change in its price is called the elasticity of demand.
- In general terms, the elasticity of demand is defined as the degree of responsiveness of demand for a product to change in its determinants.
- The measure of the degree of responsiveness of demand to change in its determinants gives the measure of the extent of relationship between the demand for a product and any of its determinants.
- Price elasticity of demand is generally defined as the responsiveness or sensitiveness of demand for a commodity to the changes in its price.
- The concept of arc elasticity of demand refers to the measurement of demand elasticity for a significant change in price and consequent change in demand.

- Point elasticity on a linear demand curve. Point elasticity is also a way to resolve the problem in measuring the elasticity. The concept of point elasticity is used for measuring price elasticity where change in price is infinitesimally small.
- The cross-elasticity is the measure of responsiveness of demand for a commodity to the changes in the price of its substitutes and complementary goods.
- The responsiveness of demand to the change in consumer's income is known as income elasticity of demand.
- Sometimes, mainly during the period of price fluctuations, consumer's price expectations play a much more important role in determining demand for a commodity than any other factor.
- The concept of price expectation elasticity refers to the expected change in price in future as a result of change in current prices of a product.

---

## 5.6 KEY WORDS

---

- **Substitute good:** It is a product or service that a consumer sees as the same or similar to another product.
- **Commodity:** It is defined as a tangible good that can be bought and sold or exchanged for products of similar value.
- **Elasticity coefficient:** It is used to quantify the concept of elasticity, including price elasticity of demand, price elasticity of supply, income elasticity of demand, and cross elasticity of demand.

---

## 5.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

### *Short-Answer Questions*

1. Briefly mention the uses of elasticity.
2. What is cross-elasticity of demand?
3. Write a note explaining elasticity of price expectations.

### *Long-Answer Questions*

1. Discuss in detail the concepts of elasticity of demand generally used in business decisions.
2. Explain price elasticity of demand. Discuss the measurement of the two kinds of price elasticity in detail.
3. What is income elasticity of demand?



---

## 5.8 FURTHER READINGS

---

Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.

Keat Paul G. and K.Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.

Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition. Singapore: Pearson Education, Inc.

Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Publishers.

## UNIT 6 MEASUREMENT OF ELASTICITY

---

### Structure

- 6.0 Introduction
- 6.1 Objectives
- 6.2 Measurement of Price Elasticity of Demand
- 6.3 Determinants of Price Elasticity of Demand
- 6.4 Demand and Supply Equilibrium
- 6.5 Answers to Check Your Progress Questions
- 6.6 Summary
- 6.7 Key Words
- 6.8 Self Assessment Questions and Exercises
- 6.9 Further Readings

---

### 6.0 INTRODUCTION

---

The price elasticity of demand measures the sensitivity of the quantity demanded to change in the price. If demand does not respond much to price changes, it is said to be inelastic, however, it is said to be elastic if demand changes a lot when the price changes. In other words, demand is elastic when there are close substitutes.

The renowned economist Alfred Marshall, devised the point method of measuring price elasticity of demand. This method is used to measure the price elasticity of demand at any given point in the curve. According to this method, elasticity of demand will be different on each point of a demand curve. Thus, when there is small change in price and quantity demanded of the commodity, the point method is applied.

This unit describes the measurement of price elasticity of demand and discusses the determinants of price elasticity of demand. One of the main concepts of economics, i.e. equilibrium, is also discussed in this unit.

---

### 6.1 OBJECTIVES

---

After going through this unit, you will be able to:

- Discuss the process of measurement of price elasticity of demand
- Describe the determinants of price elasticity of demand
- Explain demand and supply equilibrium

---

## 6.2 MEASUREMENT OF PRICE ELASTICITY OF DEMAND

---

The price elasticity of demand for a product can be measured directly from the demand function. In this section, we describe the method of measuring price elasticity of demand for a product from the demand function—both linear and non-linear. It may be noted here that if a demand function is given, arc elasticity can be measured simply by assuming two prices and working out  $\Delta P$  and  $\Delta Q$ . We will, therefore, confine ourselves here to point elasticity of demand with respect to price.

### *Measuring Price Elasticity from a Linear Demand Function*

Suppose that a linear demand function is given as

$$Q = 100 - 5P$$

Given the demand function, point elasticity can be measured for any price. For example, suppose one has to measure elasticity at  $P = 10$ . Point elasticity is measured as

$$e_p = \frac{\delta Q}{\delta P} \cdot \frac{P}{Q}$$

The term  $\frac{\delta Q}{\delta P} \cdot P$  in the elasticity formula gives the slope of the demand curve. The slope of the demand curve can be found by differentiating the demand function. Thus

$$\frac{\delta Q}{\delta P} = \frac{\delta(100 - 5P)}{\delta P} = -5$$

Having obtained the slope of the demand curve as  $\frac{\delta Q}{\delta P} = -5$ ,  $e_p$  at  $P =$

10 can be calculated as follows. At price  $P = 10$ ,  $Q = 100 - 5(10) = 50$ . By substituting these values into the elasticity formula,

$$e_p = \frac{\delta Q}{\delta P} \cdot \frac{P}{Q}$$

we get,

$$e_p = (-5) \frac{10}{50} = -1$$

Similarly, at

$$P = 8, Q = 100 - 5(8) = 60 \text{ and}$$

$$e_p = -5 (8/60) = -40/60 = -0.67$$

And at

$$P = 15, Q = 100 - 5(15) = 25, \text{ and}$$

$$e_p = -5(15/25) = -75/25 = -3$$

Suppose a non-linear demand function of multiplicative form is given as follows

$$Q = aP^{-b}$$

and we want to compute the price elasticity of demand. The formula for computing the price elasticity is the same, i.e.,

$$e_p = \frac{\frac{\delta Q}{\delta P} \cdot P}{Q} \dots(6.1)$$

What one needs to compute the price-elasticity coefficient is to find first the value of the first term,  $\frac{\delta Q}{\delta P} \cdot P$ , i.e., the slope of the demand curve. The slope can be obtained by differentiating the demand function, Thus,

$$\text{slope of demand curve} = \frac{\delta Q}{\delta P} = -baP^{-b-1} \dots(6.2)$$

By substituting Eq. (6.2) in Eq. (6.1),  $e_p$  can be expressed as

$$e_p = \frac{-baP^{-b-1} \left(\frac{P}{Q}\right)}{Q} \dots(6.3)$$

Since  $Q = aP^{-b}$ , by substitution, we get

$$e_p = \frac{-baP^{-b}}{aP^{-b}} = -b \dots(6.4)$$

Equation (6.4) shows that when a demand function is of a multiplicative or power form, price elasticity coefficient equals the power of the variable  $P$ . This means that price elasticity in the case of a multiplicative demand function remains constant all along the demand curve regardless of a change in price.

### Price Elasticity and Total Revenue

A firm aiming at enhancing its total revenue would like to know whether increasing or decreasing the price would achieve its goal. The price elasticity coefficient of demand for its product at different levels of its price provides the answer to this question. The simple answer is that if  $e_p > 1$ , then decreasing price will increase total revenue and if  $e_p < 1$ , then increasing price will increase total revenue. To prove this point, we need to know the total revenue ( $TR$ ) and the marginal revenue ( $MR$ ) functions and measures of price-elasticity are required. Since  $TR = Q \cdot P$ , we need to know  $P$  and  $Q$ . This information can be obtained through the demand function. Let us recall our earlier demand function given as

$$Q = 100 - 5P$$

Price function ( $P$ ) can be derived from the demand function as

$$P = 20 - 0.2Q \quad \dots(6.5)$$

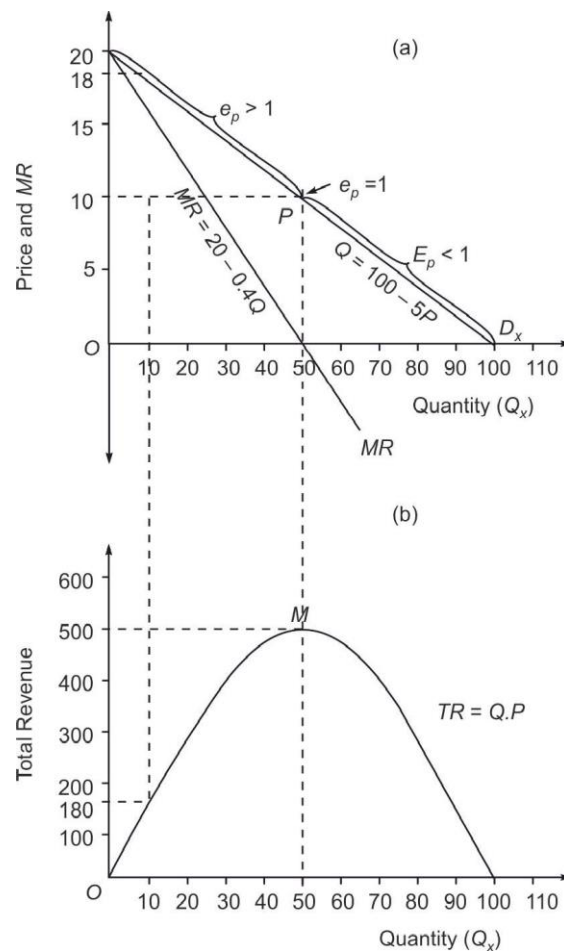
Given the price function,  $TR$  can be obtained as

$$\begin{aligned} TR &= P \cdot Q = (20 - 0.2Q)Q \\ &= 20Q - 0.2Q^2 \end{aligned} \quad \dots(6.6)$$

From this  $TR$ -function, the  $MR$ -function can be derived as

$$MR = \frac{\partial TR}{\partial Q} = 20 - 0.4Q \quad \dots(6.7)$$

The demand function and  $MR$ -function (6.7) are presented graphically in panel (a) and  $TR$ -function (6.6) in panel (b) of Fig. 6.1. As the figure shows, at point  $P$  on the demand curve,  $e = 1$  where output,  $Q = 50$ . Below point  $P$ ,  $e < 1$  and above point  $P$ ,  $e > 1$ . It can be seen in panel (a) of Fig. 6.1 that  $TR$  increases so long as  $e > 1$ ;  $TR$  reaches its maximum level where  $e = 1$ ; and it decreases when  $e < 1$ .



**Fig. 6.1** Price Elasticity and Total Revenue

The relationship between price-elasticity and  $TR$  is summed up in Table 6.1. As the table shows, when demand is *perfectly inelastic* (i.e.,  $e_p = 0$  as is the case of a vertical demand line) there is no decrease in quantity demanded when price is raised and *vice versa*. Therefore, a rise in price increases the total revenue and *vice versa*. As shown in panel (a), over the range of demand curve  $e > 1$ , quantity demanded increases by more than the proportionate decrease in price and hence the total revenue increase when price falls. The total revenue increases till price decreases till  $e = 1$ . If demand for a product is *unit elastic* ( $e_p = 1$ ) quantity demanded increases (or decreases) in the proportion of decrease (or increase) in the price. It implies that a small change in price leaves total revenue remains unchanged. Therefore, total revenue remains unaffected.

If demand for a commodity has  $e_p < 1$ , change in quantity demanded is greater than the proportionate change in price. Therefore, total revenue decreases when price falls and *vice versa*.

The case of *infinitely elastic* demand represented by a horizontal straight line is rare. Such a demand line implies that a consumer has the opportunity to buy any quantity of a commodity and the seller can sell any quantity of a commodity, at a given price. It is the case of a commodity being bought and sold in a perfectly competitive market. A seller, therefore, cannot charge a higher or a lower price.

**Table 6.1** Elasticity, Price-change and Change in  $TR$

Elasticity Co-efficient	Change in Price	Change in $TR$
$e = 0$	Increase	Increase
	Decrease	Decrease
$e > 1$	Increase	Decrease
	Decrease	Increase
$e = 1$	Increase	No change
	Decrease	No change
$e < 1$	Increase	Increase
	Decrease	Decrease
$e = \infty$	Increase	Decrease to zero
	Decrease	Infinite increase*

\*Subject to the size of the market.

### Price Elasticity and Marginal Revenue

The relationship between price-elasticity and the total revenue ( $TR$ ) can be known more precisely by finding the relationship between price-elasticity and marginal revenue ( $MR$ ).  $MR$  is the first derivative of  $TR$ -function and  $TR = P \cdot Q$  (where

$P$  = price, and  $Q$  = quantity sold). The relationship between price-elasticity,  $MR$  and  $TR$  is shown below.

Since  $TR = P \cdot Q$ ,

$$\begin{aligned} MR &= \frac{\partial(P \cdot Q)}{\partial Q} = P \left( \frac{\partial Q}{\partial Q} \right) + Q \left( \frac{\partial P}{\partial Q} \right) = P + Q \frac{\partial P}{\partial Q} \\ &= P \left( 1 + \frac{Q}{P} \cdot \frac{\partial P}{\partial Q} \right) \end{aligned} \quad \dots(6.8)$$

Note that  $\frac{Q}{P} \cdot \frac{\partial P}{\partial Q}$  in Eq. (6.8) gives the reciprocal of elasticity. That is,

$$\frac{Q}{P} \cdot \frac{\partial P}{\partial Q} = \frac{1}{e}$$

Remember that  $e_p$  carries a 'minus' sign.

By substituting  $\frac{1}{e}$  for  $\frac{Q}{P} \cdot \frac{\partial P}{\partial Q}$  in Eq. (6.8), we get

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

or 
$$MR = P [1 + (1/e_p)] \quad \dots(6.9)$$

Given this relationship between  $MR$  and price-elasticity of demand, the decision-makers can easily know whether it is beneficial to change the price. If  $e = 1$ ,  $MR = 0$ . Therefore, change in price will not cause any change in  $TR$ . In case  $e < 1$ ,  $MR < 0$ ,  $TR$  decreases when price decreases and  $TR$  increases when price increases. And, if  $e > 1$ ,  $MR > 0$ ,  $TR$  increases if price decreases and *vice versa*.

**Price Elasticity,  $AR$  and  $MR$**  Given the Eq. (6.9), the formula for price elasticity ( $e_p$ ) can be expressed in terms of  $AR$  and  $MR$ . We know that  $P = AR$ . So Eq. (6.9) can be written as

$$\begin{aligned} MR &= AR \left( 1 + \frac{1}{AR} \right) \\ MR &= AR + \frac{AR}{e_p} \end{aligned}$$

By rearranging the terms, we get

$$MR - AR = \frac{AR}{e_p}$$

$$\frac{MR - AR}{AR} = \frac{1}{e_p}$$

or

The reciprocal of this equation gives the measure of the price elasticity ( $e_p$ ) of demand which can be expressed as

$$\frac{MR - AR}{AR} = \frac{1}{e_p} \quad \text{or} \quad e_p = \frac{AR}{MR - AR}$$

---

### 6.3 DETERMINANTS OF PRICE ELASTICITY OF DEMAND

---

We have noted above that price-elasticity of demand for a product may vary between zero and infinity. However, price-elasticity of demand, at a given price, varies from product to product depending on the following factors.

**1. Availability of Substitutes.** One of the most important determinants of elasticity of demand for a commodity is the availability of its close substitutes. The higher the degree of closeness of the substitutes, the greater the elasticity of demand for the commodity. For instance, coffee and tea, rice and wheat, and petrol and diesel may be considered as close substitutes for one another. If price of one of these goods increases, the other commodity becomes relatively cheaper. Therefore, consumers buy more of the relatively cheaper good and less of the costlier one, all other things remaining the same. The elasticity of demand for the substitute goods will be higher. Besides, the wider the range of the substitutes, the greater the elasticity. For instance, soaps, toothpastes, cigarettes, etc., are available in different brands, each brand being a close substitute for the other. Therefore, the price-elasticity of demand for each brand is much greater than that for the generic commodity. On the other hand, sugar and salt do not have close substitutes and hence their price-elasticity is lower.

**2. Nature of Commodity.** *The nature of a commodity* also affects the price-elasticity of its demand. Commodities can be grouped as luxuries, comforts, and necessities. Demand for luxury goods (e.g., high-price refrigerators, TV sets, cars, decoration items, etc.) is more elastic than the demand for necessities and comforts because consumption of luxury goods can be dispensed with or postponed when their prices rise. On the other hand, consumption of necessary goods, (e.g., sugar, clothes, vegetables) cannot be postponed and hence their demand is inelastic. Comforts have more elastic demand than necessities and less elastic than luxuries. Commodities are also categorized as durable goods and perishable or non-durable goods. Demand for durable goods is more elastic than that for non-durable goods, because when the price of the former increases, people either get the old one repaired instead of replacing it or buy a 'second hand'.

**3. Weightage in the Total Consumption.** Another factor that influences the elasticity of demand is the proportion of income which consumers spend on a particular commodity. If proportion of income spent on a commodity is large, its demand will be more elastic. On the contrary, if the proportion of income spent on



a commodity is small, its demand is less price-elastic. Classic examples of such commodities are salt, matches, books, pens, toothpastes, etc. These goods claim a very small proportion of income. Demand for these goods is generally inelastic because increase in the price of such goods does not substantially affect the consumer's budget. Therefore, people continue to purchase almost the same quantity even when their prices increase.

**4. Time Factor in Adjustment of Consumption Pattern.** Price-elasticity of demand depends also on *the time consumers need to adjust their consumption pattern to a new price*: the longer the time available, the greater the price-elasticity. The reason is that over a period of time, consumers are able to adjust their expenditure pattern to price changes. For instance, if the price of high quality cell phones is decreased, demand will not increase immediately unless people possess excess purchasing power. But over time, people may be able to adjust their expenditure pattern so that they can buy a laptop at a lower (new) price. Consider another example. If price of petrol is reduced, the demand for petrol does not increase immediately and significantly. Over time, however, people get incentive from low petrol prices to buy automobiles resulting in a significant rise in demand for petrol.

**5. Range of Commodity Use.** *The range of uses of a commodity* also influences the price-elasticity of its demand. The wider the range of the uses of a product, the higher the elasticity of demand for the decrease in price. As the price of a multi-use commodity decreases, people extend their consumption to its other uses. Therefore, the demand for such a commodity generally increases more than the proportionate increase in its price. For instance, milk can be taken as it is and in the form of curd, cheese, ghee and butter-milk. The demand for milk will therefore be highly elastic for decrease in price. Similarly, electricity can be used for lighting, cooking, heating and for industrial purposes. Therefore, with decrease in its price, demand for electricity has a greater elasticity. However, for the increase in price, such commodities have a lower price-elasticity because the consumption of a normal good cannot be cut down substantially beyond a point when the price of the commodity increases.

**6. Proportion of Market Supplied.** The elasticity of market demand also depends on the *proportion of the market supplied at the ruling price*. If less than half of the market is supplied at the ruling price, price-elasticity of demand will be higher than 1 and if more than half of the market is supplied,  $e < 1$ .

#### *Application of Price Elasticity*

Having explained the concept and measurement of elasticity, we discuss now the application of price elasticity of demand with respect to (a) manoeuvring price to maximize sales revenue and (b) determination of optimal price for profit maximization.

**1. Manoeuvring of Price.** Price manoeuvring means changing price of the product to achieve business objective. The concept of elasticity of demand plays a crucial role in business-decisions regarding manoeuvring of prices for the benefit of the firm. For instance, when cost of production is increasing, the firm would want to pass the rising cost on to the consumer by raising the price. Firms may decide to change the price even without any change in the cost of production. But, whether raising price following the rise in cost or otherwise proves beneficial or not depends on at least two factors:

- (a) The price-elasticity of demand for the product, i.e., how high or low is the proportionate change in its demand in response to a certain percentage change in its price; and
- (b) Price-elasticity of demand for its substitute, because when the price of a product increases, the demand for its substitutes increases automatically even if their prices remain unchanged.

Raising the price will be beneficial only if (i) demand for a product is less elastic; and (ii) demand for its substitute is much less elastic. Although most businessmen are intuitively aware of the elasticity of demand of the goods they make, the use of precise estimates of elasticity of demand will add precision to their business decisions.

**2. Determination of Optimum Price.** Another and a rather very important application of price elasticity is that it can be used, at least theoretically, to determine the optimum price with the objective of profit maximization for a firm facing downward sloping demand curve. How optimum price, i.e., profit maximizing price, can be determined by using the price elasticity is explained below. The necessary condition for profit maximization is given as  $MR - MC$ . That is, profit is maximized at the level of price and output at which

$$MR = MC \quad \dots(6.10)$$

As has been shown in Eq. (6.9),

$$MR = P [1 + (1/e_p)] \quad \dots(6.11)$$

By substituting Eq. (6.11) for  $MR$  in Eq. (6.10), the profit maximizing condition can be expressed as

$$P [1 + (1/e_p)] = MC \quad \dots(6.12)$$

The optimal price can be worked out from Eq. (6.12) as follows.

$$P = MC / [1 + (1/e_p)] \quad \dots(6.13)$$

Given the Eq. (6.13). if point elasticity of demand curve and marginal cost ( $MC$ ) of a firm are known, the optimal price can be easily determined. For example, suppose point elasticity of demand at point on demand curve is estimated as  $e_p = -2$  and firm's  $MC = ₹50$ . In that case, optimal price can be worked out as follows.

$$\begin{aligned} P &= 50 / [1 + (1/-2)] \\ &= ₹100 \end{aligned}$$

It may thus be concluded that the concept of price elasticity of demand can be used to manipulate the price to maximize the revenue of the firm given the demand function and to find the optimal price for profit maximization.

---

## 6.4 DEMAND AND SUPPLY EQUILIBRIUM

---

In this section, we explain how demand and supply strike a balance, how market attains **equilibrium**, and how **equilibrium price** is determined in a free market. A **free market** is one in which market forces of demand and supply are free to take their own course: there is no outside control on price, demand and supply.

### *The Concept of Equilibrium*

In a general sense, the term equilibrium means the “state of rest”. It indicates the condition where forces working in opposite direction are in balance. In the context of the market analysis, *equilibrium refers to a state of market in which the quantity demanded of a commodity equals the quantity supplied of the commodity*. The equality of demand and supply produces an **equilibrium price**.

The **equilibrium price** is the price at which quantity demanded of a commodity over a period of time equals its quantity supplied over that period. In other words, at equilibrium price, demand and supply are in equilibrium. The equilibrium price is also called *market-clearing price* because at this price the quantity that suppliers want to supply equals the quantity that buyers are willing to buy. Market is cleared in the sense that there is no unsold stock and no unsupplied demand.

### *Determination of Equilibrium Price*

The equilibrium price in a free market is determined by the market forces of demand and supply. In order to analyse how equilibrium price is determined, we have to analyse the process through which market forces bring the suppliers' plan in balance with the buyers' plan. For this purpose, let us use our example of demand and

supply schedules for shirts. Suppose that the market demand and supply schedules for shirts are given as shown in Table 6.2.

**Table 6.2** Monthly Demand and Supply Schedules for Shirts

Price per Shirt (₹)	Demand ('000 shirts)	Supply ('000 shirts)	Market Position	Effect on Price
100	80	10	Shortage	Rise
200	55	28	Shortage	Rise
300	40	40	<b>Equilibrium</b>	<b>Stable</b>
400	28	50	Surplus	Fall
500	20	55	Surplus	Fall
600	15	60	Surplus	Fall

Table 6.2 places the demand and supply schedules side by side against the same price schedule. As the table shows, there is only one price of shirts (₹ 300) at which the market is in equilibrium, i.e., the quantity demanded and the quantity supplied are equal at 40 thousand shirts. At all other prices, the shirt market is in *disequilibrium* as either demand exceeds supply or supply exceeds demand. At all prices below ₹300, demand exceeds supply showing *shortage* of shirts in the market. Likewise, at all prices above ₹300 supply exceeds demand showing *surplus supply*.

**In a free market, disequilibrium itself creates the condition for equilibrium.** When there is excess supply, it forces downward adjustments in the price and quantity supplied and demanded. When there is excess demand, it forces upward adjustments in the price and in quantity demand and supplied. The process of downward and upward adjustments in price and quantity continues till the price reaches ₹300 and quantities supplied and demanded balance at 40 thousand shirts. This process is automatic. Let us now look into the process of price and quantity adjustments.

### *Demand and Supply Adjustment*

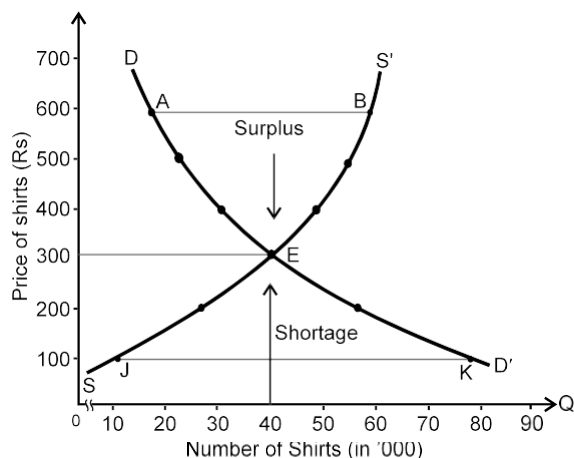
Let the price be initially set at ₹100. At this price, the quantity demanded exceeds the quantity supplied by 70 thousands shirts. The shortage gives sellers an opportunity to raise the price. Increase in price enhances the profit margin. This induces firms to produce more in order to maximize their profits. This trend will continue till price rises to ₹300. As Table 6.2 shows, at price ₹300, the buyers are willing to buy 40 thousands shirts. This is exactly the number of shirts that the sellers would like to sell at this price. At this price, there is neither shortage nor surplus of shirts in the market. This price is therefore the equilibrium price.

Similarly, at all prices above ₹300, supply exceeds demand showing surplus or excess supply of shirts in the market. The excess supply forces the competing sellers to cut down the price. Some firms find low price unprofitable and go out of market and some cut down their production. Therefore, supply of shirts goes down. On the other hand, fall in price invites more customers. This process continues

until price of shirts falls to ₹300. At this price, demand and supply are in balance and market price is in equilibrium.

*Graphical Presentation*

The determination of equilibrium price is illustrated graphically in Fig. 6.2. The demand curve  $DD$  and the supply curve  $SS$  have been obtained by plotting the demand and supply schedules, respectively, (given in Table 6.2) on the same price and quantity axes.



**Fig. 6.2** *Equilibrium of Demand and Supply: Price Determination*

As Fig. 6.2 shows, demand and supply curves for shirts intersect at point  $E$  determining the **equilibrium price** at ₹300. At this price, the quantity demanded (40 thousand shirts) equals the quantity supplied. Thus, the equilibrium price is ₹300 and equilibrium quantity is 40 thousand shirts. The equilibrium condition is not fulfilled at any other point on the demand and supply curves.

*Algebra of Demand-Supply Equilibrium*

In the previous section, we have seen graphically how the equilibrium of demand and supply is determined at the point of intersection of the demand and supply curves. If demand and supply functions are known, the equilibrium quantity and equilibrium price can also be determined algebraically. In this section, we assume linear demand and supply functions for a commodity  $X$  to show the determination of equilibrium price and quantity.

Let the **demand function** for commodity  $X$  be given as

$$Q_d = 150 - 5P_x$$

and **supply function** as

$$Q_s = 10P_x$$

We know that the equilibrium of demand and supply takes place where the quantity supplied equals the quantity demanded, i.e., where

$$Q_s = Q_d \quad \dots(6.14)$$

By substituting supply and demand functions into Eq. 6.14, we get

$$10P_x = 150 - 5P_x \quad \dots(6.15)$$

Given the Eq. (6.15) the **equilibrium price** can be worked out as follows:

$$10P_x = 150 - 5P_x \quad \dots(6.16)$$

$$\text{or } 10P_x + 5P_x = 150$$

$$15P_x = 150$$

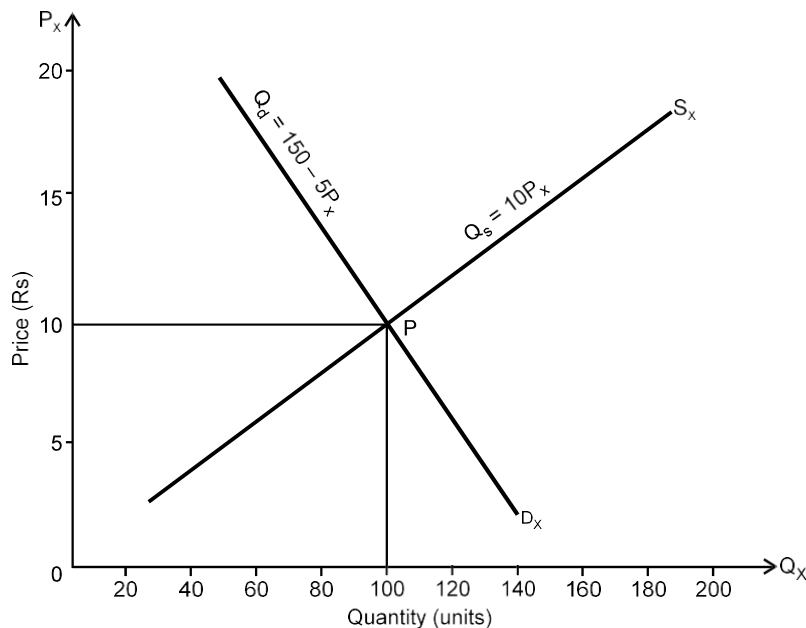
$$P_x = 10$$

Thus, at equilibrium,  $P_x = 10$ . That is, the equilibrium price is ₹10.

Given the equilibrium price  $P_x = 10$ , the equilibrium quantity supplied and the quantity demanded can be easily worked out. Equilibrium supply equals  $10(P_x) = 10(10) = 100$ . Similarly, equilibrium demand equals  $150 - 5P_x = 150 - 5(10) = 100$ .

The algebraic determination of equilibrium price and quantity can also be demonstrated graphically. It is shown in Fig. 6.3. The demand curve  $DD\phi$  has been drawn using the demand function  $Q_d = 150 - 5P_x$  and the supply curve  $SS\phi$  using the supply function  $Q_s = 10P_x$ .

The demand and the supply curves intersect at point  $P$ . A perpendicular drawn from point  $P$  to the quantity axis determines the equilibrium quantity at 100 units and a line drawn from point  $P$  to the price axis determines the equilibrium price at ₹10. At this price, the quantity demanded equals the quantity supplied at 100 units. (Fig. 6.3).



**Fig. 6.3** Determination of Equilibrium Price and Quantity

---

## 6.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

---

1. Price elasticity in the case of a multiplicative demand function remains constant all along the demand curve regardless of a change in price.
2. Price manoeuvring means changing price of the product to achieve business objective.
3. A free market is one in which market forces of demand and supply are free to take their own course: there is no outside control on price, demand and supply.
4. The equilibrium price is the price at which quantity demanded of a commodity over a period of time equals its quantity supplied over that period.

---

## 6.6 SUMMARY

---

- The price elasticity of demand for a product can be measured directly from the demand function.
- Given the demand function, point elasticity can be measured for any price.
- Price elasticity in the case of a multiplicative demand function remains constant all along the demand curve regardless of a change in price.
- If demand for a product is unit elastic ( $e_p = 1$ ) quantity demanded increases (or decreases) in the proportion of decrease (or increase) in the price. It implies that a small change in price leaves total revenue remains unchanged. Therefore, total revenue remains unaffected.
- The case of infinitely elastic demand represented by a horizontal straight line is rare. Such a demand line implies that a consumer has the opportunity to buy any quantity of a commodity and the seller can sell any quantity of a commodity, at a given price.
- It is the case of a commodity being bought and sold in a perfectly competitive market. A seller, therefore, cannot charge a higher or a lower price.

- The relationship between price-elasticity and the total revenue (TR) can be known more precisely by finding the relationship between price-elasticity and marginal revenue (MR). MR is the first derivative of TR-function and  $TR = P \cdot Q$  (where P = price, and Q = quantity sold).
- One of the most important determinants of elasticity of demand for a commodity is the availability of its close substitutes. The higher the degree of closeness of the substitutes, the greater the elasticity of demand for the commodity.
- The nature of a commodity also affects the price-elasticity of its demand. Commodities can be grouped as luxuries, comforts, and necessities.
- Another factor that influences the elasticity of demand is the proportion of income which consumers spend on a particular commodity. If proportion of income spent on a commodity is large, its demand will be more elastic. On the contrary, if the proportion of income spent on a commodity is small, its demand is less price-elastic.
- Price-elasticity of demand depends also on *the time consumers need to adjust their consumption pattern to a new price*: the longer the time available, the greater the price-elasticity. The reason is that over a period of time, consumers are able to adjust their expenditure pattern to price changes.
- Price manoeuvring means changing price of the product to achieve business objective. The concept of elasticity of demand plays a crucial role in business-decisions regarding manoeuvring of prices for the benefit of the firm.
- A Free Market is one in which market forces of demand and supply are free to take their own course: there is no outside control on price, demand and supply.
- In the context of the market analysis, *equilibrium refers to a state of market in which the quantity demanded of a commodity equals the quantity supplied of the commodity*.
- The equilibrium price in a free market is determined by the market forces of demand and supply.

---

## 6.7 KEY WORDS

---

- **Disequilibrium:** It refers to loss or lack of equilibrium or stability, especially in relation to supply, demand, and prices.
- **Determinant:** It refers to a factor which decisively affects the nature or outcome of something.
- **Inelastic:** It implies something that is insensitive to changes in price or income.



---

## 6.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

### *Short-Answer Questions*

1. What is the relationship between price elasticity, marginal revenue and total revenue?
2. Explain the application of price elasticity of demand.
3. In a free market, disequilibrium itself creates the condition for equilibrium. Briefly discuss.

### *Long-Answer Questions*

1. Discuss the process of measuring price elasticity from a linear and non-linear demand function.
2. Describe the factors that determine price elasticity of demand.
3. Briefly discuss the concept of equilibrium. How is equilibrium price determined?
4. Assess the concept of demand and supply equilibrium.

---

## 6.9 FURTHER READINGS

---

Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.

Keat Paul G. and K. Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.

Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition. Singapore: Pearson Education, Inc.

Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Publishers.

---

## BLOCK - III

### PRODUCTION AND COST

---

---

#### UNIT 7 FACTORS OF PRODUCTION

---

##### Structure

- 7.0 Introduction
- 7.1 Objectives
- 7.2 Production: Basic Concepts
  - 7.2.1 Short Run and Long Run
- 7.3 Production Possibilities of An Economy
- 7.4 Production Function
- 7.5 Answers to Check Your Progress Questions
- 7.6 Summary
- 7.7 Key Words
- 7.8 Self Assessment Questions and Exercises
- 7.9 Further Readings

---

### 7.0 INTRODUCTION

---

Factors of production are the resources people use to produce goods and services; they are the building blocks of the economy. It is an economic term that describes the inputs that are used in the production of goods or services in order to make an economic profit. The four major factors of production include land, labour, capital and entrepreneurship. In this unit, you will study the factors of production in detail.

---

### 7.1 OBJECTIVES

---

After going through this unit, you will be able to:

- Discuss basic concepts of production
  - Describe short run and long run concepts of production
  - Explain the production possibilities of an economy
  - Define production function

---

### 7.2 PRODUCTION: BASIC CONCEPTS

---

In economics, the term 'production' means a process by which inputs or factors of production (land, labour, capital, etc.) are converted or transformed into an output. In other words, production means transforming *inputs*, (labour, machines,

raw materials) into an *output*. This kind of production is called ‘manufacturing’. However, production process does not necessarily involve physical conversion of raw materials into tangible goods. In the process of production, an input may be intangible (service) and an output may be intangible too. For example, in the production of legal, medical, social and consultancy services both input and output are intangible: lawyers, doctors, social workers, consultants, hair-dressers musicians, orchestra players are all engaged in productive activity.

In economic sense, *production process* may take a variety of forms. For example, transporting a commodity from one place to another where it can be used is *production* of a commodity. A coal seller does virtually nothing more than transporting coal from coal mines to the market place. Similarly, a fisherman only catches and transports fish to the market place. Their activities too are productive activities. Transporting men and materials from one place to another is in itself a productive activity: it produces *service*. *Storing* a commodity for future sale or consumption is production. Wholesaling, retailing, packaging, assembling are all productive activities. These activities are just as good examples of production as manufacturing.

### Some Concepts

#### Input and Output

An *input* is a good or service that is used into the process of production. In the words of Baumol, “An input is simply anything which the firm buys for use in its production or other process. An ‘output’ is any commodity which the firm produces or processes for sale.” An *output* is any good or service that comes out of production process.

The term ‘inputs’ needs some more explanation. Production process requires a wide variety of inputs, depending on the nature of product. But, economists have classified inputs as (i) labour, (ii) capital, (iii) land, (iv) raw materials, and (v) time. All these variables are ‘flow’ variables in the sense that they are measured per unit of time.

#### Fixed and Variable Inputs

For the sake of analytical convenience, inputs are further classified as (i) *fixed* inputs, and (ii) *variable* inputs. A *fixed input* is one whose quantity remains constant for a certain level of output, e.g., plant, building, machinery, etc. The supply of fixed inputs remains inelastic, in the short-run. A *variable input* is defined as one whose quantity changes with change in output. The supply of such inputs (as labour and raw materials) is elastic in the short-run.

#### 7.2.1 Short Run and Long Run

Corresponding to fixed and variable inputs, economists use two other terms, viz., *short run* and *long run*. **Short run** refers to a period of time in which the supply of certain inputs (e.g., plant, building and machines, etc.) is fixed or inelastic. In the

short run, therefore, production of a commodity can be increased by increasing the use of variable inputs, like labour and raw materials.

It is worth noting that short run does not refer to any fixed time period. While in some industries it may be a matter of weeks or a few months, in some others (e.g., electric and power industry), it may mean three or more years.

**Long run** refers to a period of time in which the supply of all the inputs is elastic, but not enough to permit a change in technology. That is, in the long run, the supply of even fixed variables increases. Therefore, in the long run, production of a commodity can be increased by employing more of both variable and fixed inputs.

The economists use another term, i.e., *very long period* which refers to a period in which the technology of production is subject to change. In very long run, the production function also changes. The technological advances mean that a larger output can be created with a given quantity of inputs.

Having noted some basic concepts used in production theory, let us now look into the major questions which production theory seeks to answer.

### *The Issues in Production Theory*

Production theory seeks to analyse the input and output relations and answers the following theoretical issues.

1. If all the inputs are simultaneously increased (or decreased) at a certain rate, will the output increase (or decrease) in the same proportion? Or if, for example, the amount of each input is doubled (or halved), will the output be doubled (or halved) or will it change in a different proportion?
2. Supposing there are more than one processes of producing a commodity, how will output change (or behave) in response to change in factor proportions? Or, how will output change if one input is substituted for another?
3. How can the least-cost combination of inputs be achieved? Or, in other words, how is optimum technique of production chosen?

---

## **7.3 PRODUCTION POSSIBILITIES OF AN ECONOMY**

---

As noted earlier, societies cannot have all that they want because resources are scarce and technology is given. In reality, however, both human and non-human resources available to a country keep increasing over time and technology becoming more and more efficient and productive. Availability of human resources increases due to a natural process of increase in population, and non-human resources (especially capital goods and raw materials) increase due to creative nature of human beings. Non-human resources have been increasing due to human efforts

to create more and better of capital goods, to discover new kinds and sources of raw materials, and to create a new and more efficient technique of production. Such factors bring about a change in **production possibilities** and **production possibilities frontier** of an economy.

In this section, we will describe the *production possibilities frontier* and introduce the concept of *opportunity cost*. To begin with, we will assume a **static model** with the following **assumptions**: (i) a country’s resources consists of only labour and capital; (ii) availability of labour and capital is given; (iii) the country produces only two goods—food and clothing; and (iv) production technology for the goods is given.

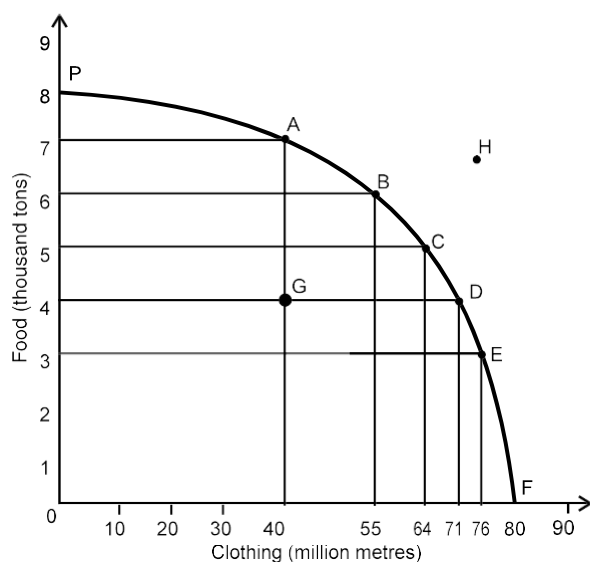
**The Production Possibilities Frontier**

*Production possibilities refer to the alternative combinations of goods and services that a society is capable of producing with its given resources and state of technology.* With reference to our model specified above, *production possibilities are the alternative combinations of maximum food and clothing that the country can produce by making full use of its labour and capital, given the technology.* For example, let us suppose that, given the availability of labour, capital and technology, the alternative production possibilities open to the country are given in Table 7.1. These production possibilities given in Table 7.1 can be presented in the form of a diagram as shown in Fig. 7.1. In this diagram, vertical axis measures food production and horizontal axis measures production of clothing. By graphing the alternative production possibilities given in Table 7.1, we locate points A, B, C, and E shown in Fig. 7.1. A number of intermediate points can be located between any two of these points. By joining these points, we get a curve *PF*. This curve is called **production possibilities frontier (PPF)**. *The production possibilities frontier (PF) shows all the alternative combinations of two goods (food and clothing) that can be produced by making full use of all the available resources (labour and capital), given the state of technology.* Each point on the *PPF* shows a different combination of two goods. For example, shows that if the country chooses point *P* on the *production possibilities frontier, PF*, it can produce 8 thousand tons of food and no clothing.

**Table 7.1** Alternative Production Possibilities

<i>Alternative</i>	<i>Food (thousand tons)</i>	<i>Clothing (million metres)</i>
P	8	0
A	7	40
B	6	55
C	5	64
D	4	71
E	3	76
F	0	80

Similarly, point *F* shows that the country can produce 80 million meters of clothing but no food. A large number of other alternative combinations of food and clothing can be located on the curve *PF* that the country can produce by making full use of its resources, given the technology. For example, point *B* shows a combination of 6 thousand tons of food and 55 million metres of clothing and point *C* shows a combination of 5 thousand tons of good and 64 million metres of clothing, and so on. The combination of food and clothing that a society chooses to produce on the *PPF* depends on the demand for food and clothing.



**Fig. 7.1** The Production Possibilities Frontier

**Implications of points away from PPF.** The *production possibilities frontier* shows the alternative combinations of two goods under the conditions that all the resources (labour and capital) are fully employed. Any point below the *PPF*, e.g., point *G*, implies underutilization or unemployment of resources. If resources are fully employed, an additional 3 thousand tons of food or 31 million metres of clothing or more of both the goods can be produced. Any point that falls beyond the *PPF*, e.g., point *H*, is unattainable for lack of resources. The scarcity of resources does not permit production of any combination of food and clothing indicated by a point outside the *PPF*.

### The Opportunity Cost

Apart from showing the possible alternative combinations of two goods, *production possibilities frontier* indicates also the *opportunity cost* of one commodity in terms of the other. Conceptually, *opportunity cost* is the benefit that is foregone to avail the benefit of another opportunity. In the present context, “The opportunity cost of an increase in the output of some product is the value of the other goods and services that must be foregone when inputs (resources) are taken away from production in order to increase the output of the product in question”. In our example, opportunity cost of food production is the quantity of clothing foregone

to produce a certain quantity of food, and *vice versa*. The concept of 'opportunity cost' can be exemplified with the help of alternative options given by the *PPF*. As can be seen in Fig. 7.1, the movement along the *production possibilities frontier*, *AF*, shows decrease in the output of one commodity and increase in the output of the other. For example, movement from point *A* to point *B* shows decrease in food production from 7 thousand tons to 6 thousand tons and increase in the production of clothing from 40 million metres to 55 million metres. It implies that one (7–6) thousand tons of food can be produced only by sacrificing 55 million metres of clothing. It means that *opportunity cost* of one thousand tons of food is 15 million metres of clothing. You can similarly find the opportunity cost of food and clothing in terms of one another between any two points on the *PPF* curve.

### Increasing Opportunity Cost and Concavity of PPF

The *production possibilities frontier* reveals another important fact that *opportunity cost* changes along the *PPF*. In Fig. 7.1, movement from point *A* downwards to points *B*, *C*, *D*, *E* and *F* shows increasing opportunity cost of clothing in terms of lost output of food. For example, movement from point *A* to point *B*, means transferring resources (labour and capital) from food production to clothing production. As a result, food production is lost by 1 thousand tons for 15 million metres of clothing. It means that the opportunity cost of 15 million metres of clothing is 1 thousand tons of food. A movement from point *B* to *C* shows that the opportunity cost of only 9 million metres of clothing, a much lower quantity, is the same one thousand tons of food. It means that opportunity cost of clothing increases as we move downwards along the *PPF*.

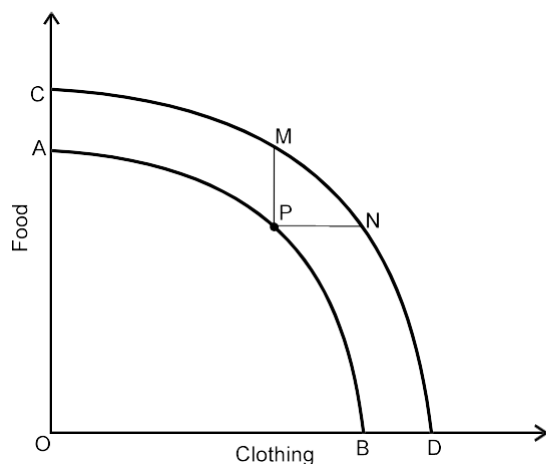
**Why is PPF Concave?** It can be seen in Fig. 7.1 that *PPF* takes the form of a *concave curve*. The *PPF* derives its concavity from the fact that opportunity cost increases along the *PPF*. Opportunity cost increases due to an economic law, i.e., the *law of diminishing returns*. The law of diminishing returns states that when more and more units of inputs are used to produce a commodity, the return on the marginal units goes on diminishing. The movement from one point on the *PPF* to another means transfer of resources from the production of one commodity to that of the other. For example, movement from point *A* towards point *F* implies transfer of resources from food production to production of clothing. As more and more resources are employed to produce clothing, marginal productivity of resources in terms of clothing goes on diminishing. The result is increase in the opportunity cost which causes concavity in the *PPF* curve.

### Shift in Production Possibilities Frontier

The *production possibilities frontier* for a country is not fixed for all times to come. In general, it keeps shifting upward for two reasons: (i) *expansion of resource*, i.e., increase in the supply of productive resources (labour and capital), and (ii) *technological improvements*. The effects of resource expansion and technological improvements on the *PPF* are explained and illustrated in Fig. 7.2.

Increase in human and non-human resources of a country, technology remaining the same, causes a parallel shift in its *PPF*. In general, resources of a country increase over time with increase in labour supply with increase in population and increase in the supply of capital. The upward shift in the *PPF* due to increase in country's resources (labour and capital) is illustrated in Fig. 7.2 assuming a given technology. Suppose that given the resources and technology of a country, its *PPF* is shown by the curve *AB* in Fig. 7.2.

Now, let the resources (labour and capital) of the country increase so that a larger quantity of labour and capital is available to produce food and clothing. With the increase in resources, the country can increase its food production by *AC* or, alternatively, production of clothing by *BD*, or a larger combination of both the goods. By joining the possible points *C* and *D*, we get a higher *PPF* as shown by the curve *CD* in Fig. 7.2. This shows an upward shift in the *PPF* from *AB* to *CD* due to increase in resources. Each point on the *production possibilities frontier CD* shows a larger combination of food and clothing. For example, suppose given its resources, the country was at producing at point *P* on the *PPF* shown by *AB*. When its resources increase, the country can increase its production of food by *PM* or of clothing by *PN* or an additional quantity of both the goods indicated by points between *M* and *N*. This kind of possibilities shows *economic growth* of the country.



**Fig. 7.2** Shift in Production Possibility Frontier

### Technological Improvement and Production Possibilities Frontier

Technological improvement refers to change in production technique so that more of goods can be produced per unit of time by using a given quantity of resources. That is, technological improvement increases the productivity of resources, both labour and capital. Technological improvement may be commodity-specific and at different points of time in different industries. In India, for example, technological



breakthrough in food production was made during the 1970s whereas technological improvement in clothing industry had started much earlier.

The shift in *PPF* indicates (i) that total food production can be increased with no change in clothing production, (ii) that more of both the goods can be produced. Similarly, part (b) shows the shift in the *PPF* when there is technological improvement in clothing industry and no such change in food industry.

**What if increase in resources and improvement in technology are simultaneous?** If *technological improvements* take place along the *resource expansion* and if technological improvements take place in both the industries is simultaneously, then the shift in *PPF* is similar to that caused by resource expansion, though the shift may not be parallel.

---

## 7.4 PRODUCTION FUNCTION

---

The tool of analysis which is used to explain the input-output relationships and gives the probable answer to the above queries is **Production Function**. Let us now briefly describe the nature and forms of a production function.

The production function describes the technological relationship between inputs and output in physical terms. It specifies the maximum quantity of a commodity that can be produced per unit of time with given quantities of inputs and technology. Beside, the production function represents the technology of a firm, of an industry or of the economy as a whole in a relevant case.

A production function may take the form of a schedule or table, a graphed line or curve, an algebraic equation or of a mathematical model. But each of these forms of a production function can be converted into the other forms.

Before we illustrate the various forms of a production function, let us note how a complex production function is simplified and the number of inputs included in the production function, as independent variables, reduced to a manageable number.

A general empirical form of production function can be expressed as

$$Q = f(L, K, LB, M, T, t, e \dots)$$

where  $Q$  = quality,  $L$  = labour,  $K$  = capital,  $LB$  = land/building,  $M$  = materials,  $T$  = technology,  $t$  = time, and  $e$  = managerial efficiency.

All these variables enter the actual production function of a firm. The economists have however reduced the number of variables used in a production

function to only two, viz., *capital* and *labour*, for the sake of convenience and simplicity in the analysis of input-output relations.

The reasons given for ignoring the other inputs are as follows: *Land/building* as an *input*, is constant for the economy as a whole, and hence it does not enter into the aggregate production function. However, land/building is not a constant variable for an individual firm or industry at least in the long run. In the case of individual firm and industries, however, land/building is lumped together with ‘capital’.<sup>2</sup> In case of ‘raw materials’, it has been observed that this input ‘bears a constant relation to output at all levels of production’. For example, cloth bears a constant relation to the number of garments, similarly for a given size of a house, the quantity of bricks, cement, steel, etc. remains constant, irrespective of number of houses constructed. This constancy of input-output relations leaves the method of production unaffected. That is why, in most production function, only two inputs—labour and capital are included. Technology, time and managerial are also assumed to be given in the short run.

We will illustrate the tabular and graphic forms of a production function when we move on to explain the laws of production. Here, let us illustrate the algebraic or mathematical form of a production function, which is most commonly used in production analysis.

To illustrate the algebraic form of production function, let us suppose that a firm employs only two inputs—capital ( $K$ ) and labour ( $L$ )—in production of a commodity. As such, the general form of its production function may be algebraically expressed as

$$Q = f(K, L) \quad \dots(7.1)$$

where  $Q$  = the quantity of output produced per time unit,  $K$  = capital, and  $L$  = labour.

The production function (7.1) implies that  $Q$  is the maximum quantity of the product that can be produced, given the total volume of capital,  $K$  and the total number of workers,  $L$ , employed to produce coal. Increasing production will require increase in  $K$  and  $L$ . Whether the firm can increase both  $K$  and  $L$  or only  $L$  depends on the time period it takes into account for increasing production, i.e., whether the firm considers the *short run* or the *long run*. As noted earlier, *short run* refers to a period during which supply of certain factors of production (viz. capital and land) is supposed to be inelastic. On the other hand, *long run* is a period of time during which supply of all the factors of production is assumed to be elastic, though not long enough to permit change in technology.

By definition, supply of capital is *inelastic* in the short-run and *elastic* in the long run. In the short run, therefore, the firm can increase its production by increasing only labour, since the supply of capital in the short run is fixed.<sup>3</sup>

In the long-run, however, the firm can employ more of both capital and labour. Accordingly, the firm would have two types of production functions: (i) *short-run production function*; and (ii) *long run production function*. The

short run production function or what may also be termed as 'single variable production function', can be expressed as

$$Q = f(L) \quad \dots(7.1)$$

In the long run-production function, both  $K$  and  $L$  are included and the function takes the form

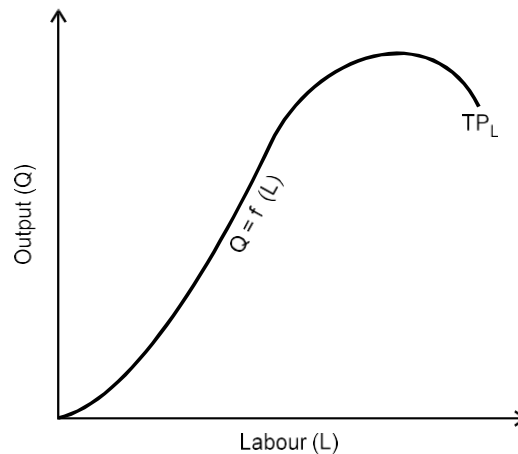
$$Q = f(K, L) \quad \dots(7.2)$$

### Assumptions

The production functions are based on certain *assumptions*:

- (i) perfect divisibility of both inputs and output;
- (ii) limited substitution of one factor for another;
- (iii) constant technology; and
- (iv) inelastic supply of fixed factors in the short-run.

If there is a change in these assumptions the production function will have to be modified accordingly.



**Fig. 7.3** Single Variable Production Function

The production function (7.1) may be graphically presented, as shown in Fig. 7.3, on a two-dimensional diagram. The vertical axis shows the quantity of output ( $Q$ ) and the horizontal axis shows the number of workers ( $L$ ) employed. When the production function is graphed, it takes a graphical form of production function. The resulting curve is called *Total Product (TP)* curve.

### Marginal Product

The laws of returns are concerned with the relation between marginal change in input and the resulting marginal change in output. Therefore, the concept of marginal product plays an important role in explaining the laws of returns. We will therefore define the marginal product of variable input, labour, and derive marginal product ( $MP$ ) curve.

From the production function (7.1), one may derive the *marginal products* ( $MP_L$ ) of labour, the variable factor. The  $MP_L$  may be defined as the change in output ( $Q$ ) resulting from a very small change ( $\Delta L$ ) in labour employed, other factors held constant. In fact, the  $MP_L$  is partial derivative of the production function with respect to labour. Thus,

$$MP_L = \frac{\Delta Q}{\Delta L},$$

or for a large change in  $L$ ,  $\frac{\Delta Q}{\Delta L} \dots(7.3)$

Geometrically,  $MP_L$  is given by the slope of the curve,  $TP_L = Q = f(L)$ .

Given the definition of  $MP_L$  the  $MP_L$  curve may be derived from the  $TP_L$ , as shown in Fig. 7.4.

By definition,  $MP_L$  is the addition to the total product resulting from a very small change in the variable input, labour (say, by one unit of labour). Or, as mentioned above, the  $MP_L$  is simply the slope of  $TP_L$ . The  $MP_L$  curve can therefore be derived by measuring the slope of  $TP_L$  at various points on it and by plotting such measures. For example, if we choose a point  $P$  on  $TP_L$  and draw a tangent  $ab$  through this point, the slope of the  $TP_L$  and that of the tangent  $ab$  at point  $P$  is the same. The ordinate  $PM$  measures the output resulting from  $OM$  labour. The contribution of the marginal labour, say  $NM$  amount of labour, can be obtained by drawing a line parallel to  $ab$  from point  $N$  through  $PM$ . Note that the parallel line intersects  $PM$  at  $P'$ . Thus,  $P'M$  is the  $MP$  of  $NM$  labour. This process may be repeated for different points chosen on the  $TP_L$  and  $MP$  of labour obtained. By joining the resultant points (say,  $P'$ ,  $Q'$  and  $T$ ), we get the  $MP_L$  curve.

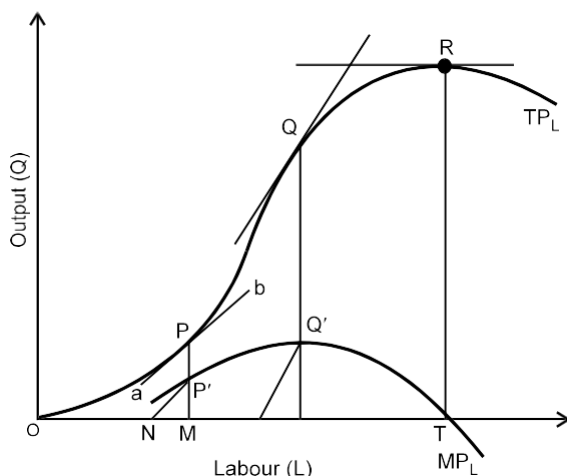


Fig. 7.4 Derivation of  $MP_L$  Curve

### Average Product

Another important concept used in discussions on production theory, though not much of theoretical importance, is *average (physical) product*. In our example the average product of labour ( $AP_L$ ) may be defined as

$$AP_L = \frac{Q}{L} = \frac{f(L)}{L} \dots (7.4)$$

The  $AP_L$  can also be derived from the  $TP_L$  curve or the function  $Q = f(L)$ .

Suppose that we want to measure  $AP_L$  at point  $P$  on the  $TP_L$  in Fig. 7.5.

At point  $P$  on the  $TP_L$  curve, output is  $PN = OM$  from the total labour employed  $ON = MP$ . Thus the  $AP_L$  at point  $P$  is

$$AP_L = \frac{PN}{ON} = \frac{MO}{MP}$$

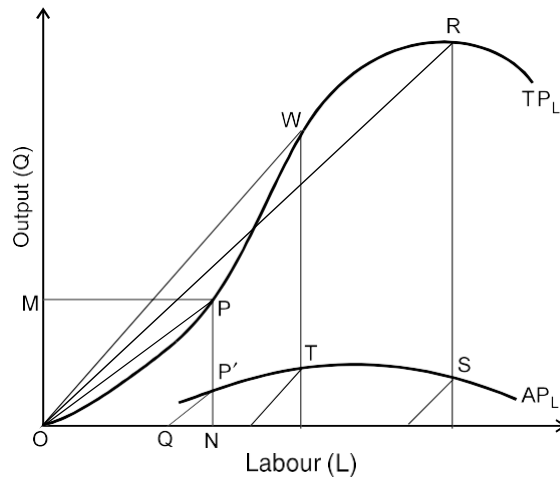


Fig. 7.5 Derivation of  $AP_L$  Curve.

This measure is geometrically obtained by drawing a line from points  $P$  to the origin  $O$ , as shown by line  $OP$ . The slope of line  $OP$  is  $PN/ON$ . This is nothing but the  $AP_L$ . Thus the slope of line  $OP$  measures the  $AP_L$ , i.e., product per unit of labour. To measure the  $AP_L$  at the total output  $PN$  and total labour  $ON$ , let us suppose that  $QN$  measures one unit of labour. Now if we draw a line parallel to  $OP$  from  $Q$  through the line  $PN$ , the point of intersection gives the measure of  $AP_L$ . Note that  $QP'$  which is parallel to  $OP$  intersects  $PN$  at  $P'$ . Thus,  $NP'$  is the measure of  $AP_L$  for point  $P$ . The same procedure may be repeated for all the points chosen on the  $TP_L$  (say,  $W$  and  $R$ ) and  $AP_L$  measured for the respective points as shown by points  $P'$ ,  $T$  and  $S$ . By joining these points, we get the  $AP_L$  curve.

## **7.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS**

1. Long run refers to a period of time in which the supply of all the inputs is elastic, but not enough to permit a change in technology.
2. The opportunity cost of an increase in the output of some product is the value of the other goods and services that must be foregone when inputs (resources) are taken away from production in order to increase the output of the product in question.
3. The production function describes the technological relationship between inputs and output in physical terms.
4. One assumption of the production function is the perfect divisibility of both inputs and output.

## **7.6 SUMMARY**

- In economics, the term ‘production’ means a process by which inputs or factors of production (land, labour, capital, etc.) are converted or transformed into an output.
- In other words, production means transforming inputs, (labour, machines, raw materials) into an output. This kind of production is called ‘manufacturing’.
- However, production process does not necessarily involve physical conversion of raw materials into tangible goods. In the process of production, an input may be intangible (service) and an output may be intangible too.
- In economic sense, production process may take a variety of forms. For example, transporting a commodity from one place to another where it can be used is production of a commodity.
- In the words of Baumol, “An input is simply anything which the firm buys for use in its production or other process. An ‘output’ is any commodity which the firm produces or processes for sale.”

- But, economists have classified inputs as (i) labour, (ii) capital, (iii) land, (iv) raw materials, and (v) time. All these variables are ‘flow’ variables in the sense that they are measured per unit of time.
- For the sake of analytical convenience, inputs are further classified as (i) fixed inputs, and (ii) variable inputs. A fixed input is one whose quantity remains constant for a certain level of output, e.g., plant, building, machinery, etc. The supply of fixed inputs remains inelastic, in the short-run.
- A variable input is defined as one whose quantity changes with change in output. The supply of such inputs (as labour and raw materials) is elastic in the short-run.
- Short run refers to a period of time in which the supply of certain inputs (e.g., plant, building and machines, etc.) is fixed or inelastic. In the short run, therefore, production of a commodity can be increased by increasing the use of variable inputs, like labour and raw materials.
- Long run refers to a period of time in which the supply of all the inputs is elastic, but not enough to permit a change in technology. That is, in the long run, the supply of even fixed variables increases.
- Production possibilities refer to the alternative combinations of goods and services that a society is capable of producing with its given resources and state of technology.
- Production possibilities are the alternative combinations of maximum food and clothing that the country can produce by making full use of its labour and capital, given the technology.
- The production possibilities frontier shows the alternative combinations of two goods under the conditions that all the resources (labour and capital) are fully employed.
- “The opportunity cost of an increase in the output of some product is the value of the other goods and services that must be foregone when inputs (resources) are taken away from production in order to increase the output of the product in question”.
- The production function describes the technological relationship between inputs and output in physical terms. It specifies the maximum quantity of a commodity that can be produced per unit of time with given quantities of inputs and technology.

---

## 7.7 KEY WORDS

---

- **Production:** It refers to the action of making or manufacturing from components or raw materials, or the process of being so manufactured.

- **Intangible:** Something that cannot be assessed, felt, measured, or moved because it has no physical substance.
- **Input:** An input is a good or service that is used into the process of production.

---

## 7.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

### Short-Answer Questions

1. Write a note on short run and long run as used in economics.
2. Briefly discuss the issues in production theory.
3. How does technological improvement impact production possibilities frontier?

### Long-Answer Questions

1. Give a detailed description of the Production Possibilities Frontier (PPF). Discuss the relationship between increasing opportunity cost and concavity of PPF.
2. Explain production function in detail. What do you understand by marginal and average product?

---

## 7.9 FURTHER READINGS

---

Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.

Keat Paul G. and K. Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.

Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition. Singapore: Pearson Education, Inc.

Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Publishers.



---

## UNIT 8 PRODUCTION FUNCTION

---

### Structure

- 8.0 Introduction
- 8.1 Objectives
- 8.2 Laws of Production
  - 8.2.1 The Law of Returns to Variable Proportions—Production with One Variable Input
- 8.3 Isoquants
  - 8.3.1 Properties of Isoquant Curves
- 8.4 Laws of Returns to Scale: Long Run Analysis of Production
- 8.5 Producer's Equilibrium or Optimum Combination of Inputs
  - 8.5.1 Optimising Combination of Inputs
  - 8.5.2 Choice of Optimal Expansion Path
  - 8.5.3 Change in Input Prices and Input Combination
- 8.6 Answers to Check Your Progress Questions
- 8.7 Summary
- 8.8 Key Words
- 8.9 Self Assessment Questions and Exercises
- 8.10 Further Readings

---

### 8.0 INTRODUCTION

---

In the previous unit, you studied about the factors of production as well as the production function.

This unit further discusses the production function in detail. Laws of Production in economics deals with the concepts of cost and producer's equilibrium. It is an important aspect of economics as it helps a business determine the level of output that leads to maximum profits. It also defines the various variable and fixed costs of the firm. These concepts are explained in detail in this unit.

---

### 8.1 OBJECTIVES

---

After going through this unit, you will be able to:

- Explain the law of variable proportions and the law of returns
- Describe the properties isoquants
- Assess the meaning of producer's equilibrium

---

## 8.2 LAWS OF PRODUCTION

---

The traditional theory of production studies the marginal input-output relationships under: (i) short run; and (ii) long run. In the short run, input-output relations are studied with *one* variable input, while other inputs are held constant. The laws of production under these assumptions are called 'The Law of Variable Proportions'. In the long run, input-output relations are studied assuming all the input to be variable. The long run input-output relations are studied under 'Law of Returns to Scale'.

### 8.2.1 *The Law of Returns to Variable Proportions—Production with One Variable Input*

The laws of returns to variable proportions explain the relationship between the variable input and the output in the short term. As noted above, certain factors of production (viz., land and capital equipments such as plant and machinery) are available in short supply in the short run. Such factors are known as **fixed factors**. On the other hand, the factors which are available in unlimited supply even in the short run are known as **variable factors**. In short run, therefore, the firms can employ only a limited or fixed quantity of fixed factors and an unlimited quantity of the variable factor. In other words, firms can employ in the short run, varying quantities of variable inputs against a given quantity of fixed factors. This kind of change in input combinations leads to variation in factor proportions. The laws which bring out the relationship between varying factor-proportions and output are therefore known as the Law of Variable Proportions. or what is more popularly known as the Law of Diminishing Returns.

The Law of Diminishing Returns states that if more and more units of a variable input are applied to a given quantity of fixed inputs, the total output may initially increase at an increasing rate, but beyond a certain level, output increases at a diminishing rate. Precisely, marginal increase in total output eventually decreases when additional units of variable factors are applied to a given quantity of the fixed factors. The main reason behind the operation of this law is the decreasing labour-capital ratio. Given the quantity of fixed factor (capital), with increasing variable input (labour) capital-labour ratio goes on decreasing. That is, each additional worker has less and less tools and equipments to work with. Consequently, the productivity of the marginal worker eventually decreases. As a result, the total output increases but at a diminishing rate beyond a point.

To illustrate the Law of Diminishing Returns let us assume that the firm (in our earlier example) possesses a set of machinery as its capital ( $K$ ) which is fixed

in the short run and that it can increase only the number of workers to increase its production. Thus, the short-run production function for the firm will be of the form given in Eq. i.e.,

$$Q = f(L)$$

Let us assume that the labour-output relationship is given by the following production function

$$Q = -L^3 + 10L^2 + 20L^1 \quad \dots(8.1)$$

Given the production function (8.1), we may substitute numerical values for  $L$  in the function and workout a series of  $Q$ , i.e., the quantity of output that can be produced with different number of workers. For example, if we substitute 5 for  $L$ , the production function (8.1) will read as

$$\begin{aligned} Q &= -5^3 + 10 \times 5^2 + 20 \times 5 \\ &= -125 + 250 + 100 \\ &= 225 \end{aligned}$$

A tabular array of output levels associated with different number of workers, from 1 to 10, in our hypothetical example is given in Table 8.1 (Cols. 1 and 2). From the table, we derive the marginal product ( $MP_L$ ) and the average product ( $AP_L$ ) schedules, as given in the table (Cols. 3 and 4).

**Table 8.1** Schedule Illustrating Law of Diminishing Returns

No. of workers (tons)	Total product (TP)(tons)	Marginal product ( $MP_L$ ) physical units (tons)	Average product (N) ( $AP_L$ ) (tons)	Stage of production
1	2	3	4	5
1	29	29	29	
2	72	43	36	
3	133	51	41	I
4	176	53	44	
5	225	49	45	
6	264	39	44	
7	287	23	41	II
8	288	1	36	
9	261	(-)27	29	
10	200	(-)61	20	III

Note:  $MP_L = TP_n - TP_{n-1}$  where  $n$  is the number of workers. For example, productivity of 7th worker may be known as  $TP_n - TP_{n-1} = 287 - 264 = 39$  and  $AP_L = TP/N$ .

The labour-output data contained in Table 8.1 is presented graphically in Fig. 8.1. The  $TP_L$  and  $MP_L$  schedules demonstrate the law of diminishing returns. As the curve  $TP_L$  indicates, the total output continues to increase at an increasing rate till the employment of the 4th worker, the rate of increase in the total output, i.e., marginal product ( $MP_L$ ), is increasing. Beyond the employment of 4th worker, although  $TP_L$  continues to increase (until 8th worker) the rate of increase in  $TP_L$  (i.e., marginal addition to  $TP_L$ ) begins to fall and continues to fall, ultimately becoming negative.

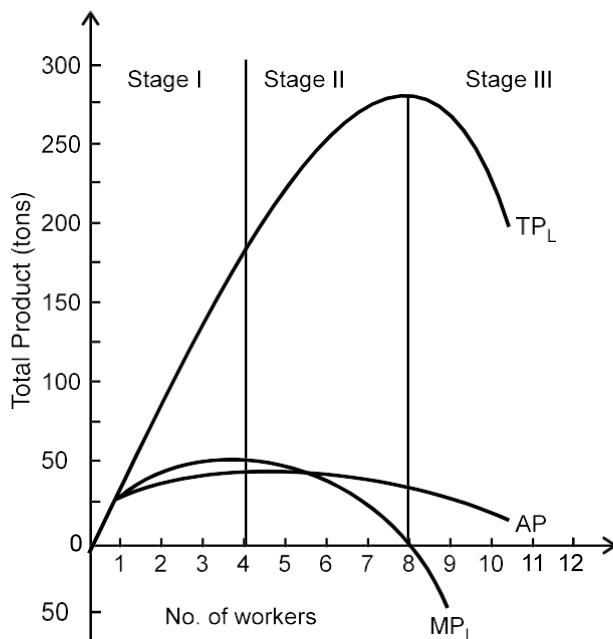


Fig. 8.1 Total Average and Marginal Products

### Three Stages in Production

Table 8.1 and Fig. 8.1 present the three usual stages in the application of the laws of diminishing returns. In Stage I,  $MP_L$  continues to increase making  $TP_L$  increase at an increasing rate. In Stage II,  $MP_L$  starts falling so that  $TP_L$  increases at a decreasing rate until  $MP$  becomes negative when  $TP$  starts falling.

The reasons which underlie the application of the laws of returns to variable proportions in Stages I and II may be described as follows.

The law of increasing returns operates because of *underutilisation of the fixed factor*, i.e., capital. Let us suppose that optimum capital-labour combination is 1:4. It implies that if less than 4 workers are employed, the plant or machine would remain underutilised. When more and more workers are added, utilisation of machine increases and also the productivity of additional workers. Another

reason for increase in labour productivity is that employment of additional workers lead to advantages of *division of labour*, until optimum capital-labour combination is reached.

Once the optimum capital-labour ratio is reached, employment of additional workers means substitution of capital with labour. But technically, one factor can substitute another only to a limited extent. In other words, there is a limit to which one input can be substituted for another. That is, the elasticity of substitution between inputs is not infinite. Hence, to replace the same amount of capital and to achieve the labour productivity at the optimum level of capital-labour combination, more and more workers will have to be employed. Naturally, per worker marginal productivity decreases.

### Assumptions

The application of the law of diminishing returns is subject to the following assumptions:

- (a) The state of technology is given—it does not change in the course of application of the law.
- (b) Input prices remain unchanged.
- (c) The variable factors are homogeneous.

### Application of the Law of Diminishing Returns

The Law of Diminishing Returns is an empirical law, frequently observed in the various production activities. This law however may not apply universally to all kinds of productive activities since the law is not as perfect as the laws of physics. In some productive activities, it may operate quickly; in some, its operation may be delayed; and in some others, it may not appear at all. The law of diminishing returns has been found to appear in agricultural production more quickly than in industrial production, because in the former, a natural factor (land), plays a predominant role, while in the latter, man-made factors play the major role. Despite the limitations of the law, if increasing units of an input are applied to the fixed factors, the marginal returns to the variable input eventually decrease.

---

## 8.3 ISOQUANTS

---

In this section, we will discuss the input-output relationship under the condition that *both labour and capital are variable factors*. This is a long-run phenomenon. Since in the long run, supply of both the inputs is supposed to be elastic, firms can hire a larger quantity of both labour and capital. Employment of labour and capital in larger quantity increases the scale of production. The long-run production theory is therefore also called as the **laws of returns to scale**.

The long run technological relationship between input and output is generally expressed through a long-run production function. A simple graphical method of presenting production function is Isoquant curve. We will, therefore, first explain the 'Isoquant' curve and certain important concepts used in production theory. The laws of returns to scale will then be explained through various kinds of production functions and graphically through Isoquants.

### *Isoquant Curves*

The term 'isoquant' has been derived from *iso* meaning 'equal' and *quant* meaning 'quantity'. The 'isoquant curve' is therefore also known as 'Equal Product Curve' or 'Production Indifference Curve'. An isoquant curve is a locus of points representing the various combinations of two inputs—capital and labour—yielding the same output. The fact that different input combinations can produce the same output is based on the assumption that capital and labour can be substituted for one another but at a diminishing rate.

To draw an isoquant curve, the following assumptions are made:

- (i) that a producer uses only two inputs, labour ( $L$ ) and capital ( $K$ ), to produce a commodity  $X$ ;
- (ii) that  $L$  and  $K$  can be substituted for one another at a diminishing rate;
- (iii) that the technology of production is given for the period under reference; and
- (iv) that the production function of the firm is continuous, i.e., labour and capital are perfectly divisible and substitutable.

Given the assumptions, the production function takes the following form:

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

The production function being continuous, it cannot be conveniently presented in a tabular form. It can however be conveniently presented graphically. In Fig. 8.2, the vertical and horizontal axes measure  $K$  and  $L$ , respectively. Given the production function and the technology, it is technically possible to form a number of capital-labour combinations that can produce a given quantity of a commodity,  $X$ . When such combinations are plotted, it gives the **isoquant**. Consider, for example, the isoquant labelled  $Q = 100$ . Each point on this curve shows a different capital-labour combination that can produce 100 units of  $X$  per unit of time. For example, points  $A$  and  $B$  represent two alternative combinations of capital and labour capable of producing 100 units of  $X$ . That is,  $OK_2$  units of capital and  $OL_1$  units of labour (as indicated by point  $A$ ) or  $OK_1$  units of capital and  $OL_2$  units of labour (as indicated by point  $B$ ) can be alternatively used to produce 100 units of  $X$  per unit of time. All other capital-labour combinations represented by this curve can be used to produce the same quantity of  $X$ .

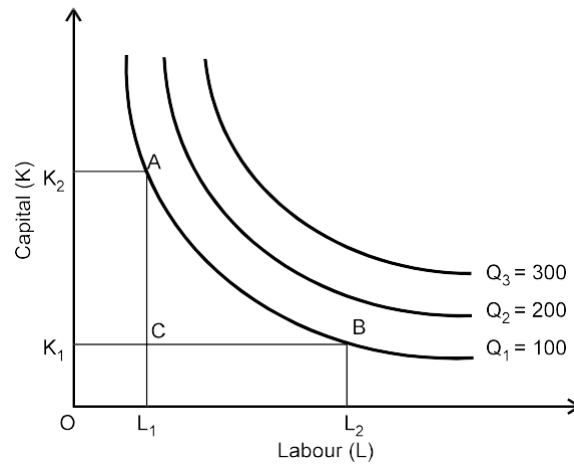


Fig. 8.2 Isoquants

It is important to note here that movement along an isoquant means substitution of one factor for another. For example, movement from point A to B means that  $L_1L_2 (= CB)$  units of labour is substituted for  $K_1K_2 (= CA)$  units of capital. It also implies that  $K_1K_2$  units of capital can produce as much as  $L_1L_2$  units of labour. The rate at which one factor can substitute another is called *marginal rate of technical substitution*. This concept is of great significance in production analysis. Therefore, we will explain this concept in detail.

### Marginal Rate of Technical Substitution

The marginal rate of technical substitution (*MRTS*) is the rate at which one input can be substituted for another without changing the level of output. The rate at which one input can be substituted for another, holding the output constant, is given by the slope of the isoquant. Since the slope of an isoquant moving down the isoquant is given by  $-\frac{\partial K}{\partial L}$ ,

$$MRTS = -\frac{\partial K}{\partial L}$$

= Slope of the Isoquant

The condition that the total output should remain constant implies that marginal product of  $K$  (i.e.,  $MP_K$ ) must equal marginal product of  $L$  (i.e.,  $MP_L$ ). That is,

$$(-\partial K \times MP_K) = (\partial L \times MP_L) \quad \dots(8.2)$$

By rearranging Eq. 8.2, we get

$$-\frac{\partial K}{\partial L} = \frac{MP_L}{MP_K}$$

Since  $\frac{\Delta K}{\Delta L} = MRTS$ ,

$$\frac{MP_L}{MP_K} = MRTS \quad \dots(8.3)$$

Thus, *MRTS* of *L* for *K* is the ratio of the marginal product of labour ( $MP_L$ ) to the marginal product of capital ( $MP_K$ ).

To illustrate the *MRTS* numerically, let us suppose that a given production function may be presented in a tabular form as given in Table 8.2. The table presents 5 alternative combinations of *K* and *L* that can be used to produce a *given quantity*, say 10 units, of a commodity.

**Table 8.2** Alternative Methods of Producing 10 Units of a Commodity

<i>K</i>	+	<i>L</i>	$\Delta K$	$\Delta L$	$MRTS = \Delta K / \Delta L$
10	+	2	-2	2	-1.0
8	+	4	-3	6	-0.5
5	+	10	-4	10	-0.4
1	+	20			

Note that as we move down the table, the *MRTS* declines. This is an important factor in determining the shape of the isoquant. The downward movement on an isoquant indicates substitution of labour for capital. The amount of capital *decreases* while the number of workers *increases*, so that output remains constant. The units of labour which can substitute one unit of capital go on increasing. As a result, the *MRTS* ( $= - \Delta K / \Delta L$ ) decreases. The reason is that both the factors are subject to the law of diminishing marginal return. As the number of labour increases, its marginal product decreases. On the other hand, with the decrease in the quantity of capital, its marginal productivity increases. Thus, labour whose marginal productivity is decreasing substitutes capital whose marginal productivity is increasing. Therefore, to substitute each subsequent unit of capital, more and more units of labour are required to maintain the production at the same level. That is why the *MRTS* decreases.

### 8.3.1 Properties of Isoquant Curves

Like indifference curves, Isoquants have the following properties.

(a) **Isoquants Have a Negative Slope.** An isoquant has a negative slope in the economic region or in the relevant range. Economic region is the region on the isoquant plane in which substitution between inputs is technically possible. It is also known as the profit maximising region. The negative slope of the isoquant



implies that if one of the inputs is reduced, the other input has to be so increased that the total output remains unaffected, i.e., the reduction in production due to reduction in the quantity of an input is exactly offset by the output resulting from increase in the other input. For example, movement from  $A$  to  $B$  on the isoquant  $Q_1 = 100$  (Fig. 8.2) means that if  $K_1K_2$  of units of capital are removed from the production process,  $L_1L_2$  units of labour have to be employed to maintain the same level of output.

(b) **Isoquants Are Convex to the Origin.** Convexity of isoquants implies not only the substitution of one factor for another but also diminishing marginal rate of technical substitution ( $MRTS$ ). As mentioned above, the  $MRTS$  is the rate at which marginal unit of an input can be substituted for the other input so that the level of output remains the same. The  $MRTS$  decreases because, by assumption, no factor is a perfect substitute for another. For this reason, more and more units of an input are needed to replace each successive unit of other input. (See Table 8.2).

(c) **Isoquants Cannot Intersect or be Tangent to Each Other.** The intersection or tangency of two isoquants implies that a certain quantity of a commodity can be produced with a smaller input combination as well as with a larger input combination. This is not consistent with the theory of production so long as marginal productivity of an input is greater than zero. In Fig. 8.3, two isoquants intersect each other at point  $M$ . Consider two other points—point  $J$  on isoquant marked 100 and point  $N$  on isoquant marked 200. One can easily infer that the quantity that can be produced with the combination of  $K$  and  $L$  at point  $M$ , can be produced also with factor combination at  $J$  and  $N$ . On the isoquant 100, factor combinations at points  $M$  and  $J$  are equal in terms of productivity. And, on the isoquant 200, factor combination at points  $M$  and  $N$  are equal in terms of productivity. Since point  $M$  is common to both the isoquants, it follows that input combinations at  $M$  and  $N$  are equal in terms of productivity. It means that in terms of output,

$$OL_2 + JL_2 = OL_2 + NL_2$$

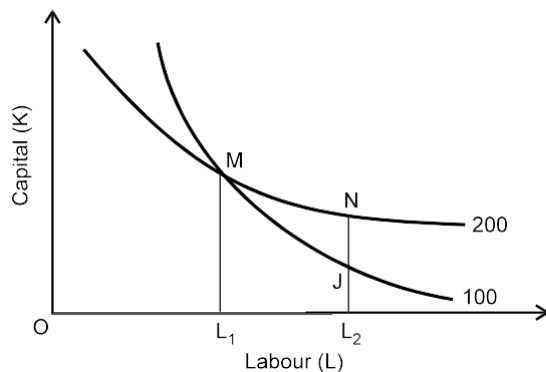
Since  $OL_2$  is common to both the sides, it means, that

$$JL_2 = NL_2$$

But it is **not true** because, as shown in Fig. 8.3,

$$JL_2 < NL_2.$$

But intersection between isoquants would mean that  $JL_2 = NL_2$  which is not consistent with theory of production. That is why isoquants will not intersect or be tangent to each other, unless marginal productivity of an input is zero.

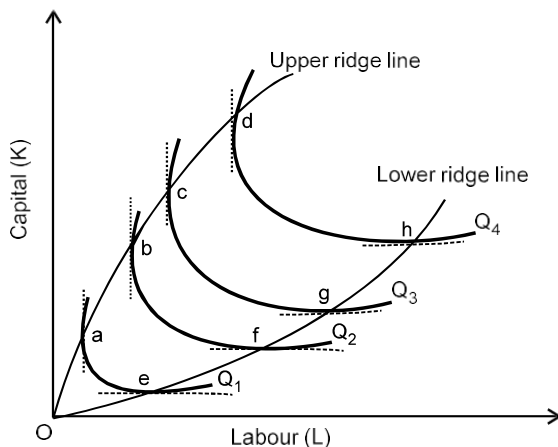


**Fig. 8.3** Intersecting Isoquants

**Isoquant Map and Economic Region**

One way to present a production function on a two-dimensional plane is to use its *isoquant map*. An *isoquant map* is a set of isoquants presented on a two-dimensional plane as shown by isoquants  $Q_1$ ,  $Q_2$ ,  $Q_3$  and  $Q_4$  in Fig. 8.4. Each isoquant shows various combinations of two inputs that can be used to produce a given level of output. An upper isoquant is formed by a greater quantity of one or both of the inputs than that indicated by the lower isoquants. For example, isoquant  $Q_2$  indicates a greater input-combination than that shown by isoquant  $Q_1$ , and so on.

Also, since upper isoquant indicates a larger input-combination than the lower ones, each successive upper isoquant indicates a higher level of output than the lower ones. For example, if isoquant  $Q_1$  represents an output equal to 100 units, isoquant  $Q_2$  represent an output greater than 100 units. As one of the properties of isoquants, no two isoquants can intersect or be tangent to one another.



**Fig. 8.4** Isoquant Map

Conventional economic theory concentrates on the economically efficient range of output, i.e., till marginal productivity ( $MP$ ) of an input is positive. Zero  $MP$  of an input is the limit of efficient range of output. This phenomenon can be presented through isoquants.

It is noteworthy that the whole isoquant map or production plane is not technically efficient, nor is every point on isoquant technically efficient. The reason is that, on a convex isoquant, the  $MRTS$  decreases along the isoquant. The limit to which the  $MRTS$  can decrease is zero. The zero  $MRTS$  implies that there is a limit to which one input can substitute another. It determines also the minimum quantity of an input which must be used to produce a given output. Beyond this point, an additional employment of one input will necessitate employing additional units of the other input. Such a point on an isoquant may be obtained by drawing a tangent to the isoquant and parallel to the vertical and horizontal axes, as shown by dashed lines in Fig. 8.4. By joining the resulting points  $a, b, c$  and  $d$ , we get a line called the *upper ridge line*,  $Od$ . Similarly, by joining the points  $e, f, g$  and  $h$ , we get the *lower ridge line*,  $Oh$ . The ridge lines are locus of points on the isoquants where the marginal products of the inputs are equal to zero. The upper ridge line implies that  $MP$  of capital is zero along the line,  $Od$ . The lower ridge line implies that  $MP$  of labour is zero along the line,  $Oh$ .

The area between the two ridge lines,  $Od$  and  $Oh$ , is called **Economic Region** or ‘technically efficient region’ of production. Any production technique, i.e., capital-labour combination, within the economic region is technically efficient to produce a given output. And, any production technique outside this region is technically inefficient since it requires more of both inputs to produce the same quantity.

For example, suppose that the quantity represented by isoquant  $Q_2$  is to be produced. We have two points  $b$  and  $f$ , on the isoquant  $Q_2$ , which fall on the ridge lines. Consider first the point  $b$ , i.e., the point of intersection between the isoquant  $Q_2$  and the upper ridge line. Point  $b$  indicates that a minimum of capital is required to produce  $Q_2$ . Any smaller amount of capital, given the labour input at point  $b$ , would be insufficient to produce  $Q_2$ . Beyond point  $b$ , producing  $Q_2$  would require more of both inputs, capital and labour, which is technically inefficient. It would mean uneconomic use of resources. It may be inferred from the above (i) that at point  $b$ ,  $MP$  of capital is zero, and (ii) that further substitution of capital for labour is technically inefficient.

We have discussed above the concept and measure of marginal rate of technical substitution (*MRTS*). The *MRTS* refers only to the slope of the isoquant, i.e., to the ratio of only marginal changes in the inputs. It does not reveal how 'difficult' or 'easy' it is to substitute one input for another. Besides, the measurement of the *MRTS* depends on the units of the measurement of the factors, which does not tell much about the substitutability of factors.

Economists have devised a method of measuring the degree of substitutability of factors, called the *Elasticity of Substitution*. The elasticity of substitution ( $\sigma$ ) is formally defined as *the percentage change in the capital-labour ratio (K/L) divided by the percentage change in the marginal rate of technical substitution (MRTS)*, i.e.,

$$\sigma = \frac{\text{Percentage change in } K/L}{\text{Percentage change in } MRTS}$$

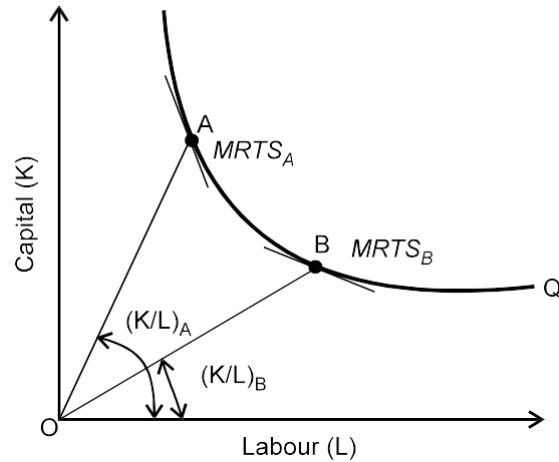
or

$$\sigma = \frac{\% (K/L) / (K/L)}{\% (MRTS) / MRTS}$$

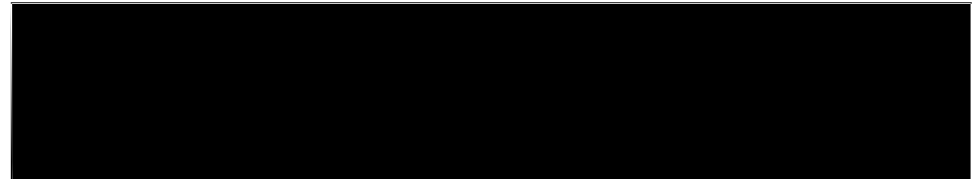
Since along an isoquant, *K/L* and *MRTS* move in the same direction, the value of  $\sigma$  is always positive. Besides, the elasticity of substitution ( $\sigma$ ) is "a pure number independent of the units of the measurement of *K* and *L*, since both the numerator and the denominator are measurement of *K* and *L*, and since both the numerator and the denominator are measured in the same units."

The concept of elasticity of substitution is graphically presented in Fig. 8.5. The movement from point *A* to point *B* on the isoquant *Q*, gives the ratio of change in the *MRTS*. The rays *OA* and *OB* represent two techniques of production with different factor intensities, i.e., *K/L* while process *OA* is capital intensive, the process *OB* is labour intensive. The shift of *OA* to *OB* gives the change in factor intensity. The ratio between the two changes measures the **elasticity of substitution**.

The value of elasticity depends on the curvature of the isoquants. It varies between 0 and  $\infty$ , depending on the nature of the production function, which determines the curvature of the various kinds of isoquants. For example, in case of a fixed-proportion production function yielding an L-shaped isoquant, (see Fig. 8.6)  $\sigma = 0$ . If production function is such that resulting isoquant is linear (Fig. 8.5),  $\sigma = \infty$ . And, in case of a homogeneous production function of degree 1 of the Cobb-Douglas type,  $\sigma = 1$ .



*Fig. 8.5 Graphic Derivation of Elasticity of Substitution*




---

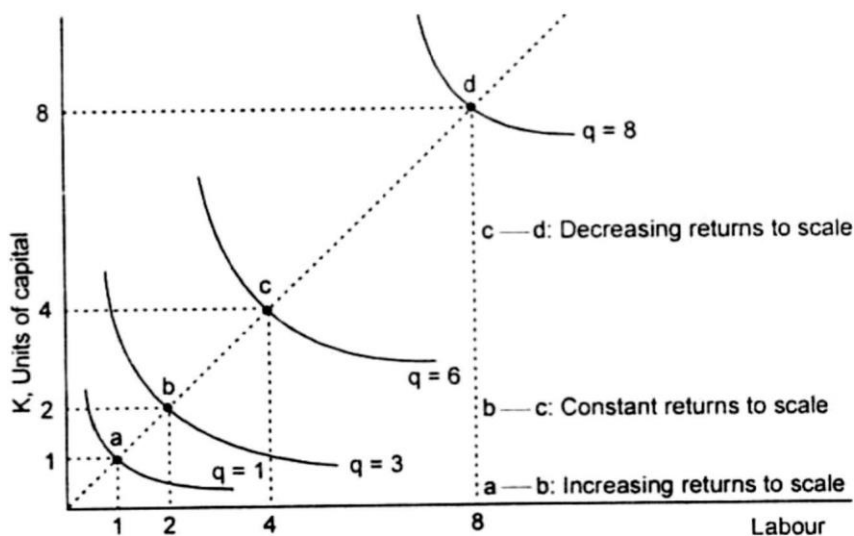
## 8.4 LAWS OF RETURNS TO SCALE: LONG RUN ANALYSIS OF PRODUCTION

---

The laws of returns to scale explain the behaviour of the total output in response to changes in the scale of the firm, i.e., in response to a simultaneous and proportional increase in all the inputs it uses. More precisely, the laws of returns to scale explain how a simultaneous and proportionate increase in both labour ( $L$ ) and capital ( $K$ ) affects the total output at various levels of input combination.

When a firm increases all its inputs proportionately, technically, there are three possibilities, i.e., the total output may increase proportionately, more than proportionately, or less than proportionately. If increase in the total output is proportional to the increase in inputs, it means **constant returns to scale**. For example, if inputs are doubled, then output is also doubled. If increase in the output is greater than the proportional increase in the inputs, it means **increasing returns to scale**. If increase in output is less than proportional to the increase in inputs, it means **diminishing returns to scale**. We will now illustrate the laws of returns to scale first through production function and then through isoquant curves.

The three laws of returns to scale are now explained with the help of a diagram below:



*Fig. 8.6 Laws of Returns to Scale*

Fig 8.6 shows that when an organization employs one unit of labour and one unit of capital, point a, it produces 1 unit of quantity as is shown on the  $q = 1$  isoquant. When the firm doubles its outputs by employing 2 units of labour and 2 units of capital, it produces more than double from  $q = 1$  to  $q = 3$ . Therefore, the production function has increasing returns to scale in this range. Another output from quantity 3 to quantity 6. At the last doubling point c to point d, the production function has decreasing returns to scale. The doubling of output from 4 units of input, causes output to increase from 6 to 8 units increases of two units only.

### *The Causes of Increasing Returns to Scale*

There are at least three plausible reasons for increasing returns to scale.

- (i) **Technical and Managerial Indivisibilities.** Certain inputs, particularly capital equipments and managerial skills, used in the process of production are available in a given minimum size. Such inputs cannot be divided into a smaller size to suit a smaller scale of production. For example, half a turbine cannot be used; a quarter or a part of a locomotive engine cannot be used; one-third or a part of composite harvester and earth-mowers cannot be used. Similarly, half of a production manager cannot be employed, if parttime employment is not acceptable to the manager. Because of indivisibility of

such factors, they have to be employed in a minimum quantity even if scale of production is relatively small. Therefore, when scale of production is increased by increasing all inputs, the productivity of indivisible factors increases exponentially. This results in increasing returns to scale.

- (ii) **Higher Degree of Specialisation.** Another factor causing increasing returns to scale is higher degree of specialisation of both labour and machinery, which becomes possible with increase in scale of production. The use of specialised labour and machinery increases productivity per unit of inputs. Their cumulative effects contribute to the increasing returns to scale. Besides, managerial specialisation contributes a great deal to increasing returns to scale.
- (iii) **Dimensional Relations.** Increasing returns to scale is also a matter of dimensional relations. For example, when the size of a room ( $15 \times 10 = 150$  sq ft) is doubled to  $30 \times 20$ , the area of the room is more than doubled, i.e.,  $30 \times 20 = 600$  sq ft. When diameter of a pipe is doubled, the flow of water is more than doubled. Following this dimensional relationship, when the labour and capital are doubled, the output is more than doubled.

---

## 8.5 PRODUCER'S EQUILIBRIUM OR OPTIMUM COMBINATION OF INPUTS

---

A profit maximising firm seeks to minimise its cost for a given output or to maximise the output from a given total costs. This objective can be achieved only by finding *optimum combination of inputs* or *what is also called least-cost combination of inputs*, given the input prices. Given the technology, a given output can be produced with different input combinations. But all input-combinations do not conform to the optimality rule of input combination. In this unit, we will show how a firm can find the optimum or the least-cost combination of inputs for a given output.

To look at this issue more closely, let us consider the information contained in Fig. 8.7. As isoquant  $I_1 = 100$  shows, 100 units of a commodity,  $X$ , can be produced with all the combinations of  $K$  and  $L$  that can be formed on the isoquant  $I_1$ . For example, points  $A$ ,  $B$  and  $C$ , represent three different combinations of  $K$  and  $L$ : (i)  $OK_3 + OL_1$ ; (ii)  $OK_2 + OL_2$ ; and (iii)  $OK_1 + OL_3$ . All these combinations of  $K$  and  $L$  can produce 100 units of  $X$ . Similarly, many other combinations of

capital and labour can be formed on the isoquant  $I_1$  that can produce 100 units of commodity  $X$ . Therefore, any of these combinations may be chosen for producing of 100 units of  $X$ . But, given the input prices—interest and wages—the total cost of production varies from point to point on an isoquant and only one of the combinations gives the minimum cost, not necessarily any of  $A$ ,  $B$  and  $C$ . The problem now is how to find an input-combination that results in the minimum cost of production. The least-cost-input combination can be determined by combining firm's production and cost functions. We know that firm's production function is represented by **isoquants**. What we need here is to devise firm's cost functions and draw the **Isocosts**.

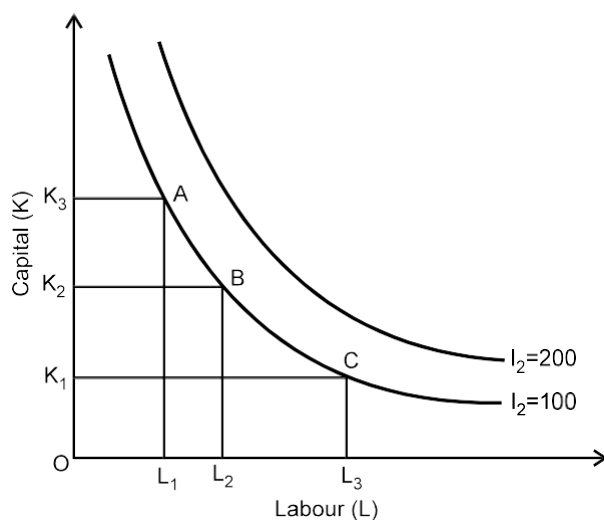


Fig. 8.7 Input Combination

### Budgetary Constraints and Isocosts

To construct the cost function, let us assume that a firm plans to incur a total cost,  $C$ , on both  $K$  and  $L$  and that  $P_k$  and  $P_l$  are the unit costs of  $K$  and  $L$ , respectively. Given these conditions, firm's cost function may be expressed as

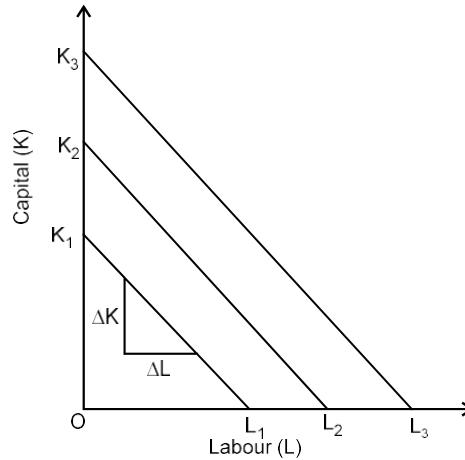
$$C = K \cdot P_k + L \cdot P_l \quad \dots(8.4)$$

From Eq. (8.4), the quantity of capital,  $K$  (and also of the number of workers  $L$ ) that can be purchased out of the total cost,  $C$ , can be easily obtained as shown below:

$$K = \frac{C}{P_k} - \frac{P_l}{P_k} L \quad \dots(8.5)$$



Equation 8.5 yields a curve which represents the alternative combination of  $K$  and  $L$  that can be purchased given the total cost  $C$ . This curve is known as **isocost curve**. The isocost is also known as the **budget line**, or the **budget constraint line**.



**Fig. 8.8** Isocosts

Consider a family of three isocosts in Fig. 8.8. They are drawn on the assumption that a firm has the option of spending its total cost,  $C$ , either on  $K$  or on  $L$ , or on both. If the resources are spent on  $K$ , or alternatively on  $L$ , the firms can buy either  $OK_1$  units of  $K$  or  $OL_1$  units of  $L$  as shown below:

$$OK_1 = \frac{C}{P_k} \frac{P_l}{P_k} L$$

where  $L = 0$

and  $OL_1 = \frac{C}{P_l} \frac{P_k}{P_l} K$

where  $K = 0$ .

The line connecting points  $K_1$  and  $L_1$  is termed as **isocost line**. It shows the whole range of combinations of  $K$  and  $L$  that can be bought, given the total cost and factor prices. Similarly, if the firm decides to spend more than  $C$ , its isocost line will shift upwards to  $K_2L_2$  or to  $K_3L_3$ , and so on. The isocosts,  $K_2L_2$  and  $K_3L_3$  show the upward movement of isocosts when firm spends more than  $C$ , given the labour and capital prices.

It is important to note here that the slope of the isocosts (i.e.,  $\Delta K/\Delta L$ ) gives the marginal rate of exchange (*MRE*) between  $K$  and  $L$ . Since factor prices are constant, marginal rate of exchange remains constant all along the line.

### 8.5.1 Optimising Combination of Inputs

Having introduced the isocosts, we may now combine isoquants and isocosts to show the optimal input combination or least-cost-combination of inputs. The optimal input combination is determined by the 'least-cost criteria'. In this section, we describe first the general conditions of the least-cost criteria. The least-cost criteria will then be applied to the given cost condition.

#### The Least Cost Criteria

The general criteria for least-cost input combination can be expressed in both physical and value terms. Given the two inputs  $K$  (capital) and  $L$  (labour) the criterion in physical terms is given by

$$\frac{\Delta K}{\Delta L} = \frac{MP_L}{MP_K} \quad \dots(8.6)$$

where  $\Delta K/\Delta L$  is the exchange ratio between  $K$  and  $L$ , and  $MP_L/MP_K$  is the ratio of

$l \quad k$

marginal productivities of  $L$  and  $K$ . This is an input combination at which factor exchange ratio (given factor prices) equals the marginal productivity ratios. This rule gives the *least-cost input combination*.

In terms of money value, the criterion for the least-cost or optimal input combination may be expressed as

$$\frac{MP_L}{MP_K} = \frac{P_L}{P_K}$$

or

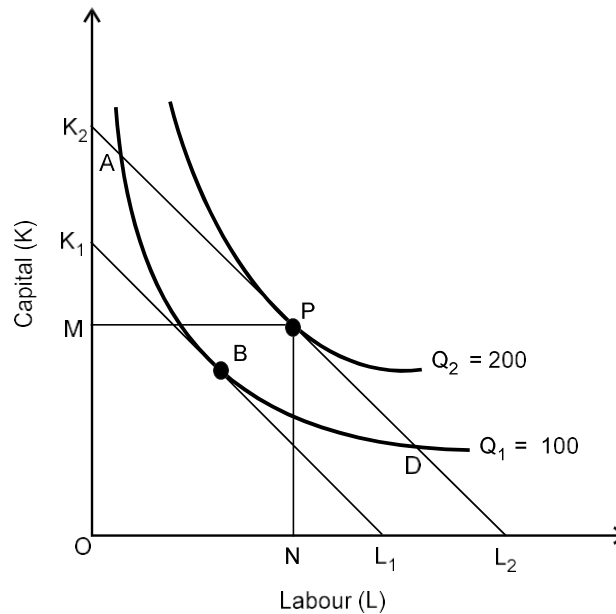
$$\frac{MP_L}{P_L} = \frac{MP_K}{P_K} \quad \dots(8.7)$$

where  $P_L$  and  $P_K$  stand for prices of labour and capital, respectively.

In Eq. (8.6),  $-\Delta K/\Delta L$  = slope of the isocost, and  $MP_L/MP_K$  = slope of the isoquants. It means that the least-cost combination exists at a point where isoquant is tangent to the isocost. The least-cost-combination of  $K$  and  $L$  is shown in Fig. 8.9. The isoquant  $Q_2 = 200$  is tangent to isocost  $K_2L_2$  at point  $P$ . At this point the combination of  $K$  and  $L$  is  $OM$  of  $K$  plus  $ON$  of  $L$ . This combination of  $K$  and  $L$  is optimal since it satisfies the least-cost-criterion, i.e.,

$$\frac{MP_K}{MP_L} = -\frac{\Delta L}{\Delta K}$$

Thus, the *necessary condition* of minimum cost is satisfied at the point of tangency between the isoquant and the isocost. At the points of tangency, the marginal exchange ratio of inputs is equal to the ratio of their marginal productivity, i.e., at this point  $P_L/P_K = MP_L/MP_K$ .



**Fig. 8.9** The Least Cost Combination of Inputs

There is however another condition, called *second order* condition, that must be satisfied simultaneously. Note that the least cost condition is satisfied also on points *A* and *D*, the points of interaction between isoquant  $Q_1 = 100$  and isocost  $K_2L_2$ . Note also that points *A*, *D* and *P* are on the same isocost, but on different isoquants. While point *P* is associated with an output  $Q_2 = 200$ , points *A* and *D* being on a lower isoquant are associated with output of 100 units. It means that, given the total cost, a firm can produce 100 units as well as 200 units. Therefore, if the firm chooses input combinations at point *A* or *D* it will produce only 100 units whereas it could produce 200 units at the same cost at point *P*. That is, only point *P* satisfies the second order condition of cost minimization.

Physical criterion (discussed above) can be translated into value terms by multiplying the factor exchange ratios with factor prices and *MRTS* with product price (*P*). In fact, factor price ratios are the same as the reciprocal of factor ratios, i.e.,  $P_l/P_k = \frac{\partial K}{\partial L}$ , and

$$MRTS = \frac{MP_l}{MP_k} = \frac{(MP_l)P}{(MP_k)P} = \frac{MRP_l}{MRP_k} \quad \dots(8.8)$$

where  $MRP$  = marginal revenue productivity of the factor and  $P$  = product price. Thus, the least-cost criterion given in Eq. (8.8) can be put in terms of input and output price as

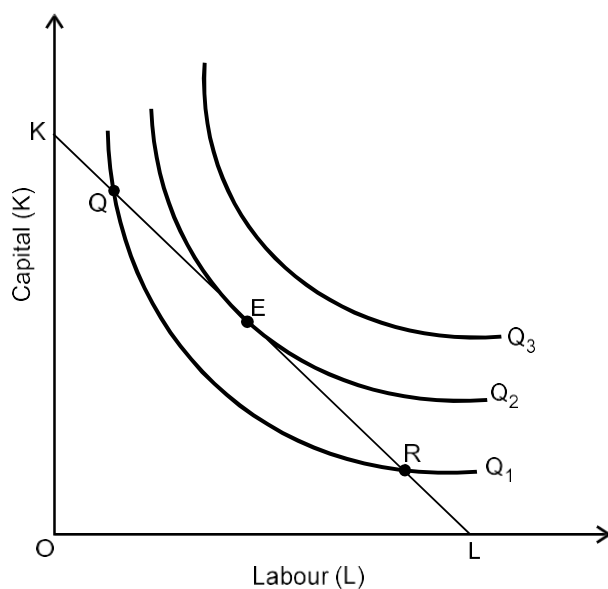
$$\frac{P_l}{P_k} = \frac{MRP_l}{MRP_k}$$

or 
$$\frac{MRP_l}{P_l} = \frac{MRP_k}{P_k} \quad \dots(8.9)$$

It may be inferred from Eq. (8.9) that the *least-cost or optimal input combination requires that the MRP ratios of inputs should be equal to their price ratios.*

### Maximisation of Output for a Given Cost

In an alternative situation, a firm faced with a *resource constraint* may seek to maximise the output. This is corollary of cost minimization hypothesis. Maximization of output at a given cost is illustrated in Fig. 8.10. In the figure, the isocost of the firm is given as shown by line  $KL$  in Fig. 8.10. The firm has to maximise the output subject to the cost constraint.



**Fig. 8.10** Output Maximisation for a Given Cost

The **first order condition** for maximising output is that the slope of the isocost must be equal to the slope of the isoquant. That is, output is maximized where

$$\frac{P_l}{P_k} = \frac{MP_l}{MP_k}$$

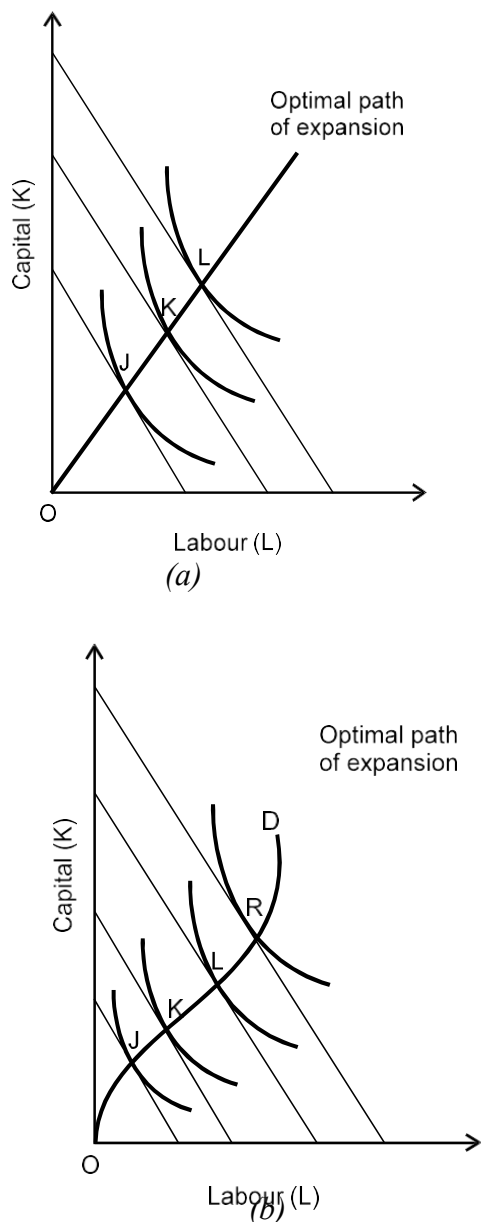
The **second order condition** requires that the first order condition must be satisfied at the highest possible isocost. Both these conditions are fulfilled at point  $P$  on the isoquant  $Q_2$ . Thus,  $Q_2$  is the maximum output attainable under the given cost condition. Although the *first order* condition is fulfilled also at points  $Q$  and  $R$  on isoquant  $Q_1$ , the output is not maximum at these points because the second order condition is not fulfilled. The points  $Q$  and  $R$ , being on a lower isoquant ( $Q_1$ ) denote an output smaller than  $Q_2$ .

### 8.5.2 Choice of Optimal Expansion Path

In the long-run, all inputs are variable. There is no constraint to the expansion of the output. The firms can employ capital and labour as much as they want in order to maximise their profit in the long-run. But, a profit maximising firm would employ capital and labour in the optimal proportion that is economically most efficient. Given the production function and input prices, the optimal factor proportion is determined by the point of tangency between the isocosts and isoquants. In other words, the optimality of factor-proportion requires that at each successive employment of labour and capital, the factor-price ratio ( $P_k/P_l$ ) equals the *MRTS*. The expansion of input and output through the points of optimal factor proportions gives the **Optimal Expansion Path**.

The **optimal paths of expansion** is illustrated in Fig. 8.11 under homogeneous and non-homogeneous production functions. If the production function is homogeneous of degree 1, the expansion path is a straight line ( $OB$ ) from the origin, as shown in Fig. 8.11(a). The line  $OB$  is obtained by joining the tangential points  $J$ ,  $K$  and  $L$ , each of which represents the optimal factor combination for a given level of output. That is, each point  $J$ ,  $K$  and  $L$ , represents the equilibrium point at different levels of output. Note that all along the isocline, *MRTS* is constant.

If production function is of non-homogeneous type, the expansion path is represented by the curve  $OD$  in Fig. 8.11(b). The expansion path represents the equilibrium path of output expansion.



**Fig. 8.11** *Optimal Paths of Expansion: (a) Homogeneous Production Function; (b) Non-homogeneous Production Function*

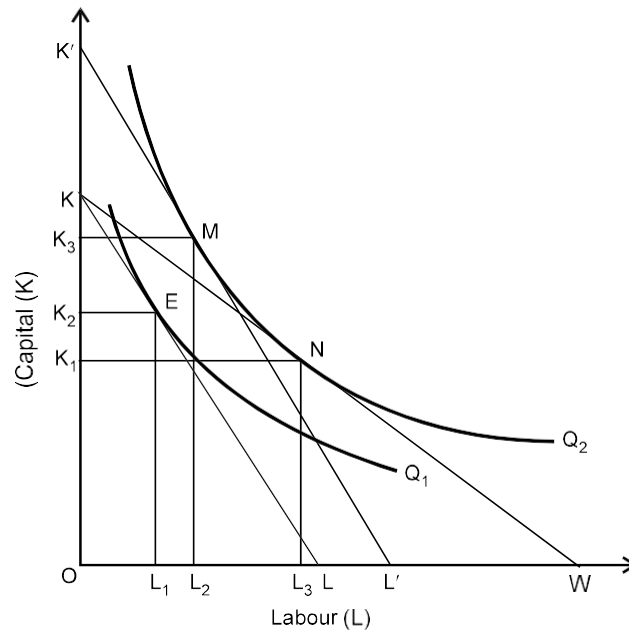
### 8.5.3 Change in Input Prices and Input Combination

We have discussed the least-cost combination of inputs assuming constant input prices. But, if input prices change, it will change the optimum input combination and also the level of output, given the total cost. It may be noted at the outset that if all input prices change in the same proportion, the relative prices of inputs

remains unaffected. But, relative prices of the inputs will change when input prices change in different proportions and in the same direction, or change unproportionately in the opposite direction or price of only one input changes while prices of other inputs remain constant. A change in relative input-prices changes both input combination and the level of output. The change in input combinations results from the *substitution effect* of change in relative prices of inputs. The change in relative prices of inputs implies that one input has become cheaper in relation to the other. The cost-minimising firms therefore substitute the relatively cheaper input for the costlier one. This is known as the *substitution effect* of change in relative input prices.

To explain the effect of change in factors prices on the input combination, let us make the following assumptions to begin with

- (i)  $P_k$  and  $P_l$  are given;
- (ii) Total resources of the firm are given; and
- (iii) Firm's initial input combination is given.



**Fig. 8.12** Substitution Effect and Input Combination

Given the assumptions, the initial equilibrium conditions of the firm are depicted in Fig. 8.12. The firm minimises its cost of point  $E$  where the firm combines  $OK_2$  of  $K$  and  $OL_1$  of labour to produce output represented by  $Q_1$ .

Given the initial conditions, let us suppose that  $P_l$  decreases while  $P_k$  remains constant so that the isocost  $KL$  shifts to  $KW$ . The isocost  $KW$  is tangent to isoquant  $Q_2$  at point  $N$ . At this point firm's new combination of inputs is  $OK_1 + OL_3$ . Thus, as a result of decrease in  $P_l$ , the firm reduces its  $K$  by  $K_1K_2$  and increases  $L$  by  $L_1L_3$ . In other words, the firm substitutes  $L_1L_3$  of labour for  $K_1K_2$  of capital with the objective of maximizing its output. This change in input combination is the result of *Price Effect*.

The price effect is indicated by movement from point  $E$  to  $N$ . Note that after decrease in  $P_l$ , the firm reduces its  $K$  by  $K_1K_2$  and adds  $L_1L_3$  to its labour input. Given the slope of the isoquants, it is obvious that  $L_1L_3$  of  $L$  is much greater than  $K_1K_2$  of  $K$  can substitute. It means that  $L_1L_3$  is not the *substitution effect*.

To find the **substitution effect**, let us find how much additional labour the firm will employ if its resources increase so that the firm reaches the isoquant,  $Q_2$ , input prices remaining the same. This can be found by drawing an isocost parallel to  $KL$  and tangent to  $Q_2$ , as shown by isocost  $K'L'$ . The isocost  $K'L'$  is tangent to isoquant  $Q_2$  at point  $M$ . It means that if  $P_k$  and  $P_l$  remain constant and firm's resources increase, it will settle itself at point  $M$  where its input combination will be  $OK_3$  of  $K$  and  $OL_2$  of  $L$ . This combination may be said to have resulted from the **budget effect** or **resource effect**, or the **output effect**. If we deduct the budget effect on labour from the price effect, we get the substitution effect, as given below.

Substitution effect

$$= \text{Price effect} - \text{Budget effect}$$

Since price effect =  $L_1L_3$ ,

and, budget effect =  $L_1L_2$ .

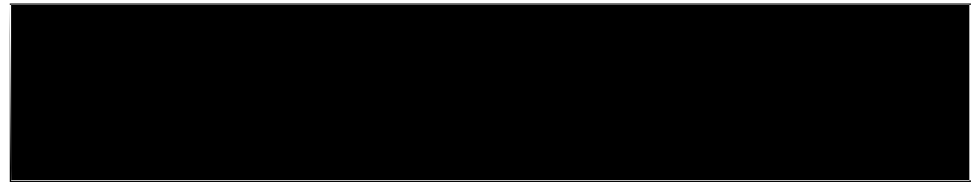
$$\text{Substitution effect} = L_1L_3 - L_1L_2 = L_2L_3.$$

Thus, we find that as a result of change in price of one input, input-combination of the firm changes: the firm employs more of cheaper input and less of the costlier one. Besides, the level of output also changes. If price of an input decreases, the level of output increases, and vice versa. It is also noteworthy



that the total effect of change in input price has two components: (i) substitution effect; and (ii) output effect. Thus, in our example, total price effect =  $L_2L_3 + L_1L_2$ .

This concludes our brief discussion on the traditional production theory, production function, law of variable proportions, law of returns to scale, and the choice of least-cost input combination. These aspects have been explained in physical terms, i.e., in terms of physical quantities of input and output. In the next chapter, we shall discuss the theory of cost—the monetary aspects of production theory.



---

## 8.6 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

---

1. The marginal rate of technical substitution (MRTS) is the rate at which one input can be substituted for another without changing the level of output.
  2. An isoquant curve is a locus of points representing the various combinations of two inputs—capital and labour—yielding the same output.
  3. A production function is said to be homogeneous when all the inputs are increased in the same proportion and this proportion can be factored out.
  4. One factor that causes increasing returns to scale is higher degree of specialisation of both labour and machinery, which becomes possible with increase in scale of production.
- 

## 8.7 SUMMARY

---

- The traditional theory of production studies the marginal input-output relationships under: (i) short run; and (ii) long run.
- In the short run, input-output relations are studied with one variable input, while other inputs are held constant. The laws of production under these assumptions are called ‘The Laws of Variable Proportions’.

- The laws of returns to variable proportions explain the relationship between the variable input and the output in the short term.
- Factors which are available in unlimited supply even in the short run are known as variable factors.
- The Law of Diminishing Returns states that if more and more units of a variable input are applied to a given quantity of fixed inputs, the total output may initially increase at an increasing rate, but beyond a certain level, output increases at a diminishing rate.
- The main reason behind the operation of this law is the decreasing labour-capital ratio. Given the quantity of fixed factor (capital), with increasing variable input (labour) capital-labour ratio goes on decreasing.
- The law of increasing returns operates because of underutilisation of the fixed factor, i.e., capital.
- The term 'isoquant' has been derived from iso meaning 'equal' and quant meaning 'quantity'. The 'isoquant curve' is therefore also known as 'Equal Product Curve' or 'Production Indifference Curve'.
- The marginal rate of technical substitution (MRTS) is the rate at which one input can be substituted for another without changing the level of output.
- The laws of returns to scale explain the behaviour of the total output in response to changes in the scale of the firm, i.e., in response to a simultaneous and proportional increase in all the inputs it uses.
- When a proportional change in output equals the proportional change in inputs, it exhibits constant returns to scale.
- A profit maximising firm seeks to minimise its cost for a given output or to maximise the output from a given total costs. This objective can be achieved only by finding optimum combination of inputs or what is also called least-cost combination of inputs, given the input prices.

---

## 8.8 KEY WORDS

---

- **Capital:** It refers to wealth in the form of money or other assets owned by a person or organization or available for a purpose such as starting a company or investing.
- **Isocost line:** This line shows all combinations of inputs which cost the same total amount.
- **Linear production function:** It is the simplest form of a production function and describes a linear relation between the input and the output.

---

## 8.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

### Short-Answer Questions

1. What are the three stages in production?
2. Write a short note on elasticity of substitution.
3. Briefly explain law of returns to scale through Cobb-Douglas production function.
4. What are the causes of increasing returns to scale?

### Long-Answer Questions

1. Give a detailed explanation of the Law of Diminishing Returns. State its application.
2. What are isoquants? Discuss the properties of isoquants.
3. Explain the Law of Returns to Scale in detail.
4. Discuss the concept of producer's equilibrium.

---

## 8.10 FURTHER READINGS

---

Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.

Keat Paul G. and K.Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.

Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition. Singapore: Pearson Education, Inc.

Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Publishers.

---

*Structure*

- 9.0 Introduction
- 9.1 Objectives
- 9.2 Cost Concepts
  - 9.2.1 Cost in Short and Long Run and their Importance
- 9.3 Cost Functions and Cost Curves: Meaning
  - 9.3.1 Types of Cost Functions
- 9.4 Answers to Check Your Progress Questions
- 9.5 Summary
- 9.6 Key Words
- 9.7 Self Assessment Questions and Exercises
- 9.8 Further Readings

---

**9.0 INTRODUCTION**

---

A cost function is a function of input prices and output quantity whose value is the cost of making that output given those input prices, often applied through the use of the cost curve by companies to minimize cost and maximize production efficiency. There are a variety of different applications to this cost curve which include the evaluation of marginal costs and sunk costs. In economics, the cost function is primarily used by businesses to determine which investments to make with capital used in the short and long term.

In this unit, you will study about cost function and the importance of cost in short run and long run.

---

**9.1 OBJECTIVES**

---

After going through this unit, you will be able to:

- Describe cost and its types
- Explain the importance of cost in short run and long run
- Discuss the types of cost functions and cost curves

---

**9.2 COST CONCEPTS**

---

Before learning about cost function, we will study different types of costs.

**Actual costs** are those which are actually incurred by the firm in payment for labour, material, plant, building, machinery, equipments, travelling and transport, etc. The total money expenses recorded in the books of accounts are, for all

practical purposes, the actual costs. Actual cost concept comes under the accounting cost concept.

**Opportunity cost** is another *fundamental* cost concept used in business decisions. The concept of ‘opportunity cost’ is related to scarcity concept. The opportunity cost is the return expected from the second best use of the resources, which is foregone for availing the gains from the best use of the resources. For example, suppose, a businessman with his limited resources can buy either a printing machine or a lathe. From the printing machine he expects an annual income of ₹ 20,000, and from the lathe he expects ₹ 15,000. The rational businessman will obviously invest his money in the printing machine and forego the expected income from the lathe. The opportunity cost of his income from printing machine is the expected income from the lathe, i.e., ₹ 15,000. The opportunity cost arises because of the foregone opportunity. In assessing the opportunity costs, both *explicit* and *implicit* costs are taken into account.

Associated with the concept of opportunity cost is the concept of *economic rent* or *economic profit*. In our example, economic rent of the printing machine is the excess of its earning over the income from the lathe. Given the returns from printing machine and lathe, economic rent = ₹ 20,000 – ₹ 15,000 = ₹ 5,000. The business implication of this concept is that investing in printing machine is preferable so long as its economic rent is greater than zero. Also, if firms know the economic rent of the various alternative uses of their resources, the choice of the best investment avenue will be a problem.

### *Business Cost and Full Cost*

**Business costs** include all the expenses which are incurred in carrying out a business. The concept of business cost is similar to the actual or real cost. Business costs “include all the payments and contractual obligations made by the firm together with the book cost of depreciation on plant and equipment.” Both these concepts are used in calculating the profits and losses in the business in filing returns for income-tax, and for other legal purposes.

The concept of **full cost** includes two other costs: *opportunity cost* and *normal profit*. Opportunity cost, as noted above, includes the expected earning from the next best use of the resources or the market rate of interest on the total money capital, and also the value of entrepreneurs own services which are not charged in the current business. Normal profit is a necessary minimum earning, in addition to opportunity cost, which a firm must get to remain in its present occupation.

### *Explicit and Implicit or Imputed Costs*

**Explicit costs** are those which fall under actual or business costs entered in the books of accounts. The payments on account of wages, salaries, utilities, interest, rent, purchase of materials, licence fee, insurance premium and depreciation charges

are the examples of explicit costs. These costs involve cash payment and are clearly reflected by the normal accounting practices. In contrast with these costs, there are certain other costs which do not take the form of cash outlays, nor do they appear in the accounting system. Such costs are known as **implicit** or **imputed** costs. Implicit costs may be defined as the earning of owner's resources employed in their best alternative uses. For example, suppose an entrepreneur does not utilise his services in his own business and works as a manager in some other firm on a salary basis. If he joins his own business, he foregoes his salary as manager. This loss of salary which is opportunity cost of his services utilised in his own firm becomes an implicit cost of his own business. It is implicit because the income foregone by the entrepreneur is not charged as the explicit cost of his own business. The implicit cost includes implicit wages, implicit rent, implicit interest etc. Although implicit costs are not taken into account while calculating the loss or gains of the business, these costs do figure in business decisions.

### *Total, Average and Marginal Costs*

**Total cost** represents the value of the total resources used in the production of goods and services. It refers to the total outlays of money expenditure, both explicit and implicit on the resources used to produce a given output. For theoretical purpose, total cost includes payments for labour, capital, land and opportunity cost. The total cost for a given output is obtained from the cost function.

**Average cost** is of statistical nature. It is obtained simply by dividing the total cost ( $TC$ ) by the total output ( $Q$ ), i.e.,  $TC/Q =$  average cost.

**Marginal cost** is the addition to the total cost on account of producing one additional unit of the product. Or, marginal cost is the cost of marginal unit produced. Marginal cost ( $MC$ ) is also defined as  $DTC/DQ$ .

Total, average and marginal cost concepts used in the economic analysis of the firm's productive activities shall be discussed in detail in the following section.

### *Fixed and Variable Costs*

**Fixed costs** are those which are fixed in volume for a certain given output. Fixed costs do not vary with the variation in the output between zero and a certain level of output. The costs that do not vary over a certain level of output are known as fixed cost. Fixed costs include cost of (i) managerial and administrative staff; (ii) depreciation of machinery, building and other fixed assets; and (iii) maintenance of land, etc. The *concept of fixed is associated with short-run*.

**Variable costs** are those which vary with the variation in the total output. Variable costs are the function of the output. Variable costs include cost of raw materials, running cost of fixed capital, such as fuel, ordinary repairs, routine maintenance expenditure, direct labour charges associated with the level of output, and the costs of all other inputs that vary with output.

These cost concepts are economic in nature and are used in economic analysis of costs behaviour in relation to output.

### *Short-run and Long-run Costs*

Two other important cost concepts which are analogous with variable and fixed costs and often figure in economic analysis are **short-run** and **long-run costs**. Short-run costs can be defined as the costs which vary with the variation of output, the size of the firm remaining the same. In other words, short-run costs are the same as variable costs. Long-run costs, on the other hand, can be defined as the costs which are incurred on the fixed assets, like plant, building machinery, land, etc. Such costs have long-run implication in the sense that these costs are not used up in the single batch of production, and are used over time in the process of production. Long-run costs are, by implication, the same as fixed costs. In the long-run, however, even the fixed costs become variable costs as the size of the firm or scale of production increases. Broadly speaking, the short-run costs are those associated with variable costs in the utilisation of fixed plant or other facilities, whereas long-run cost-behaviour encompasses changes in the size and kind of plant.

### *Private and Social Costs*

We have so far discussed the cost concepts that are related to the functioning of the firm as a production unit, and are used in the cost-benefit analysis of the business decisions. There are however certain other costs which arise due to functioning of the firm but do not normally figure in the business decisions nor are such costs explicitly paid by the firms. Some such costs are paid by the society. Thus, the total cost generated by the firm's decision may be divided into two categories: (a) those paid out or provided for by the firms; (b) those not paid by the firms, including use of resources freely available *plus* the disutility created in the process of production. The costs of the category (a) are known as **private costs**, and costs of category (b) are known as **external** or **social costs**. For instance, Private firms, situated closer Yamuna river discharge their wastes into the Yamuna river causing water-pollution; mills and factories located in a city cause air-pollution by emitting smoke; plying cars, buses, trucks, etc. causes both air and noise pollution. Such pollutions cause tremendous health hazards which involve cost to the society as a whole. Such costs do not figure in the cost structure of the firms and hence are termed **external costs** from the firm's point of view, and *social cost* from society's point of view.

**Private costs** are those which are actually incurred or provided for by an individual or a firm on the purchase of goods and services from the market. For a firm, all the actual costs, both explicit and implicit, are private costs. Private costs are internalised in the sense that "the firm must compensate the resource owner in order to acquire the right to use the resource." It is only the internalised cost that is incorporated in the firm's total cost of production.

**Social cost**, on the other hand, implies the cost which a society bears on account of production of a commodity. Social cost includes both *private cost* and the *external cost*. External cost includes (a) the cost of ‘resources for which the firm is not compelled to pay a price,’ e.g., atmosphere, rivers, lakes and also for the use of public utility services like roadways, drainage system, etc.; and (b) the cost in the form of ‘disutility’ caused by air, water, and noise pollution, etc. The cost of category (b) is generally assumed to be equal to the total private and public expenditure incurred to safeguard the individual and public interest against the various kinds of health hazards created by the production system. But private and public expenditure serve only as an indicator of trends in ‘public disutility’, it does not give the exact measure of the public disutility.

Having described cost concepts used in business-decision-making, we turn to discuss the theory of production cost, i.e., the theory of cost-output relations. Cost theory is generally discussed in the short-run and long-run framework. Let us begin with short-run cost-output relationship.

### 9.2.1 Cost in Short and Long Run and their Importance

In this section, we will learn about the short-run and long-run cost output relations.

#### Short-Run Cost-Output Relations

In this section, we discuss the relationship between output and costs and the behaviour of cost in relation to the change in output. Cost-output relations are expressed through a *cost function*. Therefore, before we proceed, let us have a look at the cost function and ‘how cost function is constructed’.

Cost functions depend on (i) production function; and (ii) market-supply function of inputs. Production function specifies the technical input-combination and its relation to the output. Production function of a firm combined with the supply function of inputs or prices of inputs determines the cost function of the firm. Thus, cost function is a function derived from the production function and the market supply function.

Cost-output relations depend on the nature of cost function. Change in cost function causes a change in the cost-output relations. Since cost function is dependent of the production function, it may change due to change in the latter. Since a production function can take different forms depending on what variables are held constant, cost functions can also take different forms. Whether certain variables in the function can be held constant or not depends on whether short-run or long-run is considered for constructing the production function. Accordingly, there are two kinds of cost functions: (a) *short-run cost function*, and (b) *long-run cost-function*.

The cost-function may be symbolically written as

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



where  $C$  = total cost  
 $Q$  = quantity produced  
 $T$  = technology

$P_f$  = factor price  
 $K$  = capital, the fixed factor

Since, in the *short-run*, all determinants of cost other than  $Q$  are constant, the *short-run cost-function* may be specified as

$$C = f(Q)$$

Let us now explain the cost-output relations in the short-run, known as the *traditional theory of cost of production*.

### Relationship between Cost and Output in Short Run

The basic cost concepts used in the analysis of cost behaviour are total, average, and marginal costs. The total cost ( $TC$ ) is defined as the total actual cost that must be incurred to produce a given quantity of output. The short-run  $TC$  is composed of two major elements: *total fixed cost (TFC)*; and *total variable cost (TVC)*. Thus,

$$TC = TFC + TVC$$

As mentioned earlier,  $TFC$  (i.e., the cost of plant, building, equipment, etc.) remains fixed in the short-run for certain level of output, whereas  $TVC$  varies with the variation in the output.

For a given quantity of output,  $Q$ , the average total cost ( $ATC$  or  $AC$ ), average fixed cost ( $AFC$ ) and average variable cost ( $AVC$ ) can be defined as follows:

$$(i) \quad AC = \frac{TC}{Q}$$

$$(ii) \quad AFC = \frac{TFC}{Q}$$

$$(iii) \quad AVC = \frac{TVC}{Q}$$

Since  $TC = TFC + TVC$ ,

$$\begin{aligned} AC &= \frac{TC}{Q} = \frac{TFC}{Q} + \frac{TVC}{Q} \\ &= AFC + AVC \end{aligned}$$

Marginal cost ( $MC$ ) is defined as the change in the total cost divided by the change in the total output, i.e.,

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

Since  $\Delta TC = \Delta TFC + \Delta TVC$  and in short run  $\Delta TFC = 0$ ,

$\Delta TC = 0 + \Delta TVC$ . Thus, in short run,

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

The short-run cost-output relationship is presented in Table 9.1, through a hypothetical cost-function which assumes production of a single commodity with one variable input. The relationship between output and costs, as presented in the table, may be summarised as follows:

1. As output increases, the *TFC* remains constant (by assumption).
2. As output (*Q*) increases, *TVC* increases but at varying rates. It increases first at a decreasing rate (till *Q* = 20 units) and then at an increasing rate.
3. As output increases, the *TC* (= *TFC* + *TVC*) increases first at a diminishing rate and then at an increasing rate following the rates of increase in the *TVC*.
4. With increase in *Q*, the *AFC* = *TFC*/*Q* decreases continuously because *TFC* remains constant for the whole range of output.
5. As *Q* increases, the *AVC* decreases till the rate of increase in *TVC* decreases. Beyond that it increases.
6. The *MC* which is defined as  $\Delta TC/\Delta Q$ , decreases till the output of 30 units and then increases. The *MC* also follows the *TVC* in the short-run.

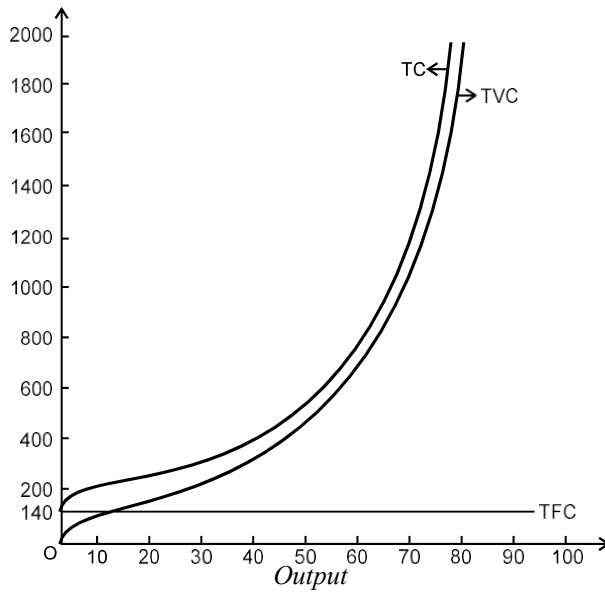
**Table 9.1** Short-Run Cost-Output Relations

<i>Q</i> (Units)	<i>TFC</i>	<i>TVC</i>	<i>TC</i>	<i>AFC</i>	<i>AVC</i>	<i>AC</i>	<i>MC</i>
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0	140	—	140	—	—	—	—
10	140	70	210	14.0	7.0	21.0	7
20	140	110	250	7.0	5.5	12.5	4
30	140	180	320	4.7	6.0	10.7	7
40	140	280	420	3.5	7.0	10.5	10
50	140	450	590	2.8	9.0	11.8	17
60	140	720	860	2.3	12.0	14.3	27
70	140	1120	1260	2.0	16.0	18.0	40
80	140	1680	1820	1.8	21.0	22.8	56

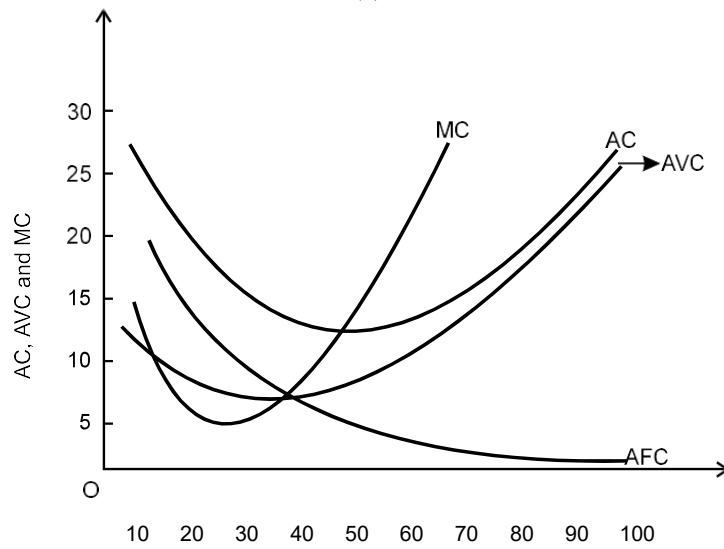
**Graphic Presentation**

The short-run output-cost relations can also be presented graphically. Figure 9.1(a) shows the relationship between output and *TFC*, *TVC* and *TC*. The *TFC* remains fixed for the whole range of output and hence, takes the form of a horizontal line—*TFC*. The *TVC* curve shows that the total variable cost first increases at a decreasing

Figure 9.1(b) brings out the laws of production costs more clearly. The law of diminishing returns states that if more and more units of a variable input are applied to the inputs held constant, the returns from the marginal units of the variable inputs initially increase but eventually decrease. The same law can also be interpreted in terms of decreasing and increasing costs. The law can be stated as when more and more units of a variable input are applied to the given amount of the fixed input, the marginal cost initially decreases, but eventually increases. Both interpretations of the law yield the same conclusion—the former in terms of marginal productivity of the variable input, and the latter, in terms of cost. Furthermore, the former is expressed through production function while the latter through a cost function.



(a)



(b)

Fig. 9.1 Short-run Cost Curves

Now Fig. 9.1(b) can be interpreted in the light of the interpretation of the laws of returns in terms of cost of production. As the figure shows, in the initial stage of production, both  $AFC$  and  $AVC$  are declining because of internal economies. Since  $AFC$  and  $AVC$  are both declining, the  $AC$  is also declining. But beyond a certain level of output (i.e., 20 units), while  $AFC$  continues to fall,  $AVC$  starts increasing because of a faster marginal increase in the  $TVC$ . Consequently, the rate of fall in  $AC$  decreases. The  $AC$  reaches its minimum when output increases to 40 units. Beyond this level of output,  $AC$  starts increasing rapidly due to the Law of Diminishing Returns coming into operation. The  $MC$  curve represents the patterns of change in both the  $TVC$  and  $TC$  curves as output changes. A downward trend in the  $MC$  shows increasing marginal productivity of the variable input which is mainly because of *internal economies* resulting from increase in production. On the other hand, rise in  $MC$  marks the disappearance of internal economies and appearance of internal diseconomies.

### *The Relationship between AC and AVC*

- (a) Since  $AC = AFC + AVC$ ,  $AC$  falls when  $AFC$  and  $AVC$  fall.
- (b) When  $AFC$  falls but  $AVC$  increases, change in  $AC$  depends on the rate of change in  $AFC$  and  $AVC$ , on the following pattern:
  - (i) if decrease in  $AFC >$  increase in  $AVC$ ,  $AC$  falls;
  - (ii) if decrease in  $AFC =$  increase in  $AVC$ ,  $AC$  remains constant; and
  - (iii) if decrease in  $AFC <$  increase in  $AVC$ ,  $AC$  increases.

### *The Relationship between MC and AC*

Much more important is the relationship between  $AC$  and  $MC$ . It may be described as follows:

- (i) When  $MC$  falls,  $AC$  follows. But the rate of fall in  $MC$  is greater than that of  $AC$  because, in case of  $MC$ , the decreasing marginal cost is attributed to a single marginal unit while, in case of  $AC$ , the decreasing marginal cost is distributed over the whole output. So long as  $MC$  curve lies below the  $AC$  curve,  $MC$  pulls  $AC$  downwards and when  $MC$  is above  $AC$ , it pulls the latter upwards.
- (ii) Similarly, when  $MC$  increases,  $AC$  also increases but at a lower rate for the reason given above. There is however an exceptional range of output where this relationship does not exist— $MC$  increases while  $AC$  continues to fall. This range of output lies between 20 and 40 units. Over this range of output,  $MC$  starts increasing while  $AC$  continues to decrease. The reason is when  $MC$  falls, it falls at a rate higher than the rate of fall in the  $AC$ . When  $MC$  starts increasing, it does so at a relatively lower rate which is not sufficient to push the  $AC$  up. That is why  $AC$  continues to fall over some range of output even if  $MC$  falls.

- (iii)  $MC$  intersects  $AC$  at its minimum. The reason is, when  $MC$  decreases it pulls  $AC$  down, and when  $MC$  increases, it pushes  $AC$  up. And when  $AC$  is at its minimum, it is neither being pulled down nor being pushed up, by the  $MC$ . It follows that  $MC = AC$  at its minimum.

### Optimum Output and Cost Curves

In the short run, optimum level of output is one which can be produced at a minimum average cost, given the technology. The minimum level of  $AC$  is determined by the point of intersection between  $AC$  and  $MC$  curves. At this level of output  $AC = MC$ ,  $AC$  being the minimum. Any other level of production, below or beyond this level, will be inoptimal. For, if production is less than 40 units [Fig. 9.1(b)], it will leave some scope for reducing  $AC$  by producing more, because  $MC < AC$ . Similarly, if production is greater than 40 units,  $AC$  can be reduced by reducing output. Thus, the cost curves can be useful in finding the optimum level of output. It may be noted here that optimum level of output is not necessarily the maximum-profit output. Profits cannot be known unless firm's revenue curves are known.

### Long-Run Cost-Output Relations

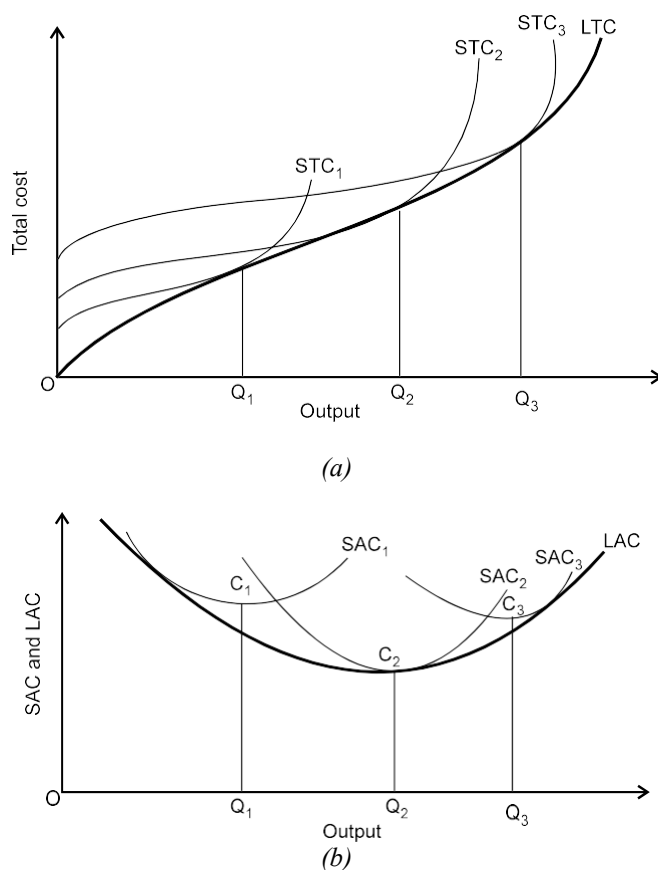
From cost theory point of view, long run is the period during which all the inputs become variable. In the short run, at least one input (especially capital) remains fixed. The variability of inputs is based on the assumption that, in the long-run, supply of all the inputs, including those held constant in the short-run, becomes elastic. The firms are, therefore, in a position to expand the scale of their production in the long-run by hiring or purchasing larger quantities of all the inputs. The long-run cost-output relations therefore imply the relationship between the total costs and the total output, whereas in the short-run this relationship is essentially one between the total output and the variable costs.

To understand the long-run cost-output relations and to derive long-run cost curves it is helpful to imagine that, operationally, a long-run is composed of a series of short-run production decisions. As a corollary of this, long-run cost curves would be composed of a series of short-run cost curves. We may now draw the long-run cost curves and study their relationship with output.

### Long-Run Total and Average Cost Curves

To derive the *long-run total cost curve*, suppose that a firm having only one plant size has its short-run total cost curve as given by  $STC_1$  in Fig. 9.2(a) and the corresponding  $SAC_1$ , in Fig. 9.2(b). Let us also suppose that the firm decides to add two more plants to its size, one after another, so that the two respective short-run total cost curves are added to  $STC_1$  in the manner as given by  $STC_2$  and  $STC_3$  in Fig. 9.2(a), and the two respective short-run average cost curves as given by  $SAC_2$  and  $SAC_3$  in Fig. 9.2(b). Thus, the firm has a series of short-run average cost curves, each having a minimum point showing the minimum average cost. For instance,  $C_1Q_1$  is the minimum  $SAC$  when the firm has only one plant. The  $SAC$

decreases to  $C_2Q_2$  due to economies of scale when the second plant is added. The  $SAC$  rises to  $C_3Q_3$  after the inclusion of the third plant.



**Fig. 9.2** Long-run Cost Curve

The long-run average cost curve ( $LAC$ ) can now be drawn by joining the bottom of  $STC_1$ ,  $STC_2$  and  $STC_3$  as shown in Fig. 9.2(b). The  $LAC$  curve is also called as 'Envelope Curve'. It is also called as "Planning Curve", as it serves as a guide to the entrepreneur in his planning to expand the production in future. The relationship between  $LTC$  and output, and between  $LAC$  and output can now be easily inferred. As is obvious from the  $LTC$  Fig. 9.2(a), the long-run cost-output relationship is similar to the short-run cost-output relation. With the subsequent increase in the output, the  $LTC$  first increases at a decreasing rate, and then, at an increasing rate. As a result,  $LAC$  initially decreases until the optimum utilisation of the second plant. The addition of the third plant makes the  $LAC$  move upward because  $SAC_3$  lies above the level of  $SAC_2$ .

From these relations are drawn the 'law of returns to scale'. The laws of returns to scale state that if a firm increases the quantity of all inputs simultaneously and proportionately, the total output initially increases more than proportionately but eventually increases less than proportionately. It implies that when the scale of the firm expands, per unit cost first decreases, but ultimately increases as shown in Fig. 9.2(b). The decrease in the per unit cost is attributed to the internal and external economies and the eventual increase in cost, to the internal and external diseconomies.

Long-Run Marginal Cost Curve

The long-run marginal cost curve (*LMC*) is derived from the short-run marginal cost curves (*SMC<sub>s</sub>*). The derivation of *LMC* is illustrated in Fig. 9.3 in which *SAC<sub>1</sub>* and *LAC* are the same as in Fig. 9.2(b). To derive the *LMC*, consider the points of tangency between *SAC<sub>1</sub>* and the *LAC*, i.e., points *A*, *B* and *C*. In the long-run production planning, these points determine the output levels at the different levels of production. For example, if we draw perpendiculars from points *A*, *B* and *C* to *X*-axis, the corresponding output levels will be *OQ<sub>1</sub>*, *OQ<sub>2</sub>* and *OQ<sub>3</sub>*. The perpendicular *AQ<sub>1</sub>* intersects the *SMC<sub>1</sub>* at point *M*. It means that at output *OQ<sub>1</sub>*, *LMC* is *MQ<sub>1</sub>*. When output increases to *OQ<sub>2</sub>*, marginal cost will be *BQ<sub>2</sub>*. Similarly, if perpendicular *CQ<sub>3</sub>* is extended upward, it intersects *SMC<sub>3</sub>* at point *N*. Thus, *NQ<sub>3</sub>* measures the *LMC* at output *OQ<sub>3</sub>*. If a curve is drawn through points *M*, *B* and *N*, as shown by the *LMC*, the curve will represent the behaviour of marginal cost in the long run. This curve is known **long-run marginal cost curve (*LMC*)**.

Some important inferences may be drawn from Fig. 9.3. The *LMC* must be equal to *SMC* for the output at which the corresponding *SAC* is tangent to the *LAC*: at the point of tangency *LAC = SAC*. For all other levels of output (considering each *SAC* separately), *SAC > LAC*. Similarly, for all the levels of output corresponding to *LAC = SAC*, the *LMC = SMC*. For all other levels of output, the *LMC* is either greater or less than the *SMC*. Another important point to notice is that *LMC* intersects *LAC* when the latter is at its minimum point *B*. There is one, and only one, short-run plant size whose minimum *SAC* coincides with the minimum *LAC*. This point is *B* where

$$SAC_2 = SMC_2 = LAC = LMC$$

The point *B* indicates also the optimum scale of the firm in the long-run, given the technology.

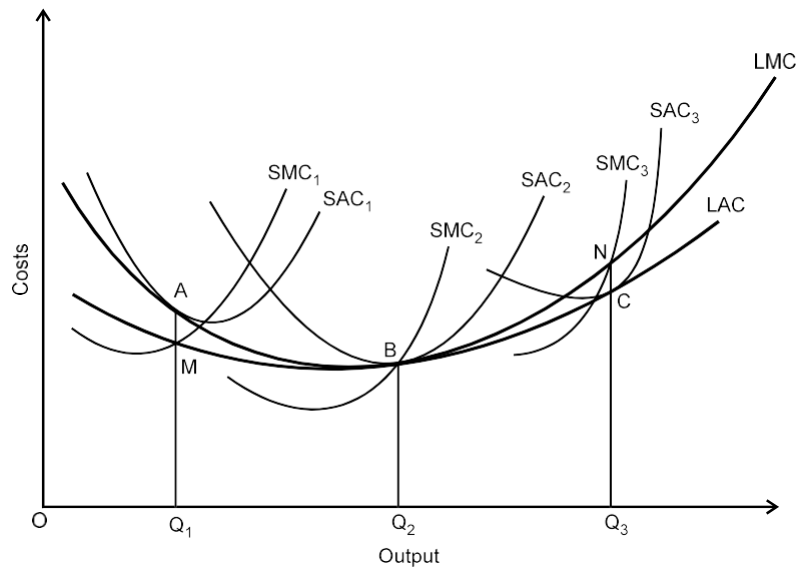


Fig. 9.3 Derivation of LMC

The short-run cost curves are helpful in showing how a firm can decide on the optimum utilisation of a plant (the fixed factor) or how it determines the least-cost output level, given the size of the plant. Long-run cost curves, on the other hand, can be used to show how a firm finds its the optimum size. An optimum size (or scale) of a plant is one which leads to the most efficient utilisation of resources. Given the state of technology over time, there is technically a unique size of the firm and level of output that determine the optimum size of the firm. This unique size of the firm can be obtained with the help of *LAC* and *LMC*. In Fig. 9.3, the unique size consists of two plants which together produce  $OQ_2$  units of a product at minimum long-run average cost (*LAC*) of  $BQ_2$ . The downtrend in the *LAC* indicates that until output reaches the level of  $OQ_2$ , the firm is of non-optimal size. Similarly, expansion of the firm beyond production capacity  $OQ_2$  leads to rise in *SMC* as well as *LAC*. It follows that, given the technology, a firm trying to minimise its average cost over time must choose a scale of production that minimises its *LAC* at a point where  $SAC = SMC = LAC = LMC$ . This assures the most efficient utilisation of the resources.

---

### 9.3 COST FUNCTIONS AND COST CURVES: MEANING

---

We have discussed and illustrated graphically the nature and shape of the short-run and long-run cost curves. In this section, we introduce the different types of cost functions of empirical nature and illustrate the derivation of the cost curves.

As mentioned above, the shape of the cost curves depends on the nature of the cost functions. Cost functions are derived from actual cost data of the firms. Given the cost data, cost functions may take a variety of forms, yielding different kinds of cost curves. The cost curves produced by **linear**, **quadratic** and **cubic cost functions** are illustrated below.

#### 9.3.1 Types of Cost Functions

In this section, we will study the different types of cost functions.

A linear cost function takes the following form

$$TC = a + bQ \quad \dots(9.1)$$

where  $TC$  = total cost,  $Q$  = output,  $a$  = fixed cost and  $b$  is a constant.

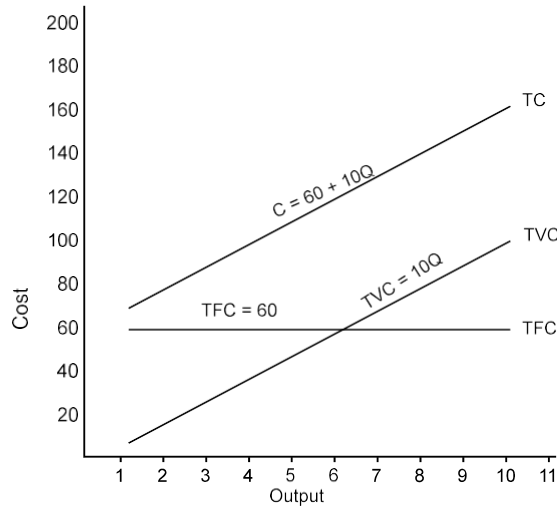
Given the cost function (Eq. 9.1), *AC* and *MC* can be obtained as follows:

$$AC = \frac{TC}{Q} = \frac{a + bQ}{Q} = \frac{a}{Q} + b$$

$$\text{and } MC = \frac{\partial TC}{\partial Q} = b$$



Note that since 'b' is a constant, *MC* remains constant throughout in case of a linear cost function. The cost curves (*TC*, *TVC* and *TFC*) are graphed in Figs. 9.4 and 9.5 assuming an actual cost function given as



**Fig. 9.4** Linear Cost Function

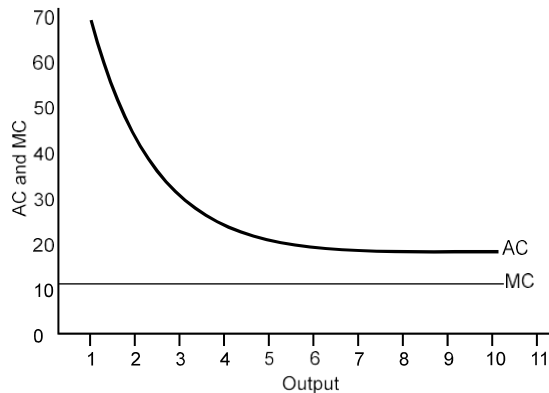
$$TC = 60 + 10Q$$

Given this function,

$$AC = 60/Q + 10$$

and  $MC = 10$

Figure 9.4 shows the behaviour of total cost curves. The straight horizontal line shows *TFC* and line marked  $TVC = 10Q$  shows the movement in *TVC*. The total cost function is shown by  $TC = 60 + 10Q$ .



**Fig. 9.5** AC and MC Curves Derived from Linear Cost Function

More important is to notice the behaviour of *AC* and *MC* curves in Fig. 9.5. Note that, in case of a linear cost function,  $MC = AVC$  and it remains constant, while *AC* continues to decline with the increase in output. This is so simply because of the logic of the linear cost function.

A quadratic cost function is of the following form:

$$TC = a + bQ + Q^2 \quad \dots(9.2)$$

where  $a$  and  $b$  are constants and  $TC$  and  $Q$  are total cost and total output, respectively. Given the cost function (9.2),  $AC$  and  $MC$  can be obtained as follows:

$$\begin{aligned}
 AC &= \frac{TC}{Q} \\
 &= \frac{a + bQ + Q^2}{Q} \quad \dots(9.3) \\
 &= \frac{a}{Q} + b + Q
 \end{aligned}$$

$$\text{and } MC = \frac{dTC}{dQ} = b + 2Q \quad \dots(9.4)$$

Let the actual (or estimated) cost function be given as

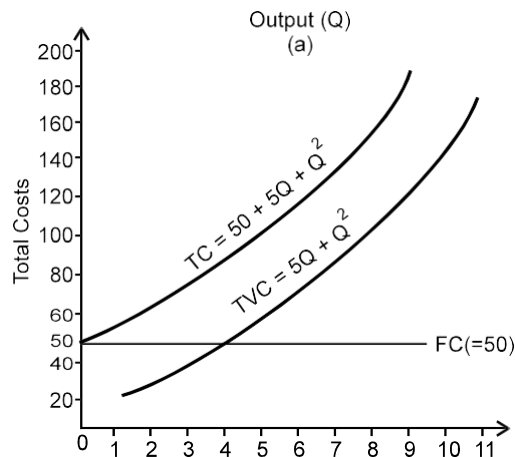
$$TC = 50 + 5Q + Q^2 \quad \dots(9.5)$$

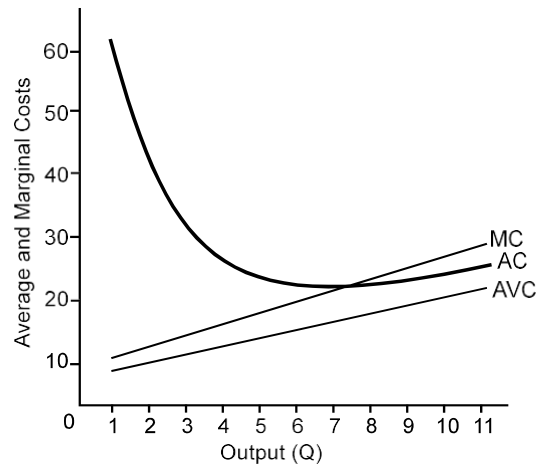
Given the cost function (9.5),

$$AC = \frac{50}{Q} + Q + 5$$

$$\text{and } MC = \frac{dTC}{dQ} = 5 + 2Q$$

The cost curves that emerge from the cost function (9.2) are graphed in Fig. 9.6(a) and (b). As shown in panel (a), while fixed cost remains constant at 50,  $TVC$  is increasing at an increasing rate. The rising  $TVC$  sets the trends in the total cost ( $TC$ ). Panel (b) shows the behaviour of  $AC$ ,  $MC$  and  $AVC$  in a **quadratic cost function**. Note that  $MC$  and  $AVC$  are rising at a constant rate whereas  $AC$  first declines and then tends to increase. Also compare these cost curves with those in Fig. 9.4(b).





(b)

**Fig. 9.6** Cost Curves Derived from a Quadratic Cost Function

### Cubic Cost Function

A cubic cost function is of the following form:

$$TC = a + bQ - cQ^2 + Q^3 \quad \dots(9.6)$$

where  $a$ ,  $b$  and  $c$  are the parametric constants. From the cost function (9.6),  $AC$  and  $MC$  can be derived as follows:

$$\begin{aligned} AC &= \frac{TC}{Q} \\ &= \frac{a + bQ - cQ^2 + Q^3}{Q} \\ &= \frac{a}{Q} + b - cQ + Q^2 \end{aligned}$$

$$\begin{aligned} MC &= \frac{\partial TC}{\partial Q} \\ &= b - 2cQ + 3Q^2 \end{aligned}$$

Assuming an estimated cubic cost function of the form

$$TC = 100 + 55Q - 10Q^2 + Q^3$$

we can generate the cost data as given in Table 9.2.

When this data is plotted on a graph paper, it will produce cost curves as shown in Fig. 9.1(a) and (b). (Plot the data on a graph paper and compare).

**Table 9.2** Cost Data Obtained from Cubic Cost Function

$Q$	$C$	$FC$	$TVC$	$AC$	$MC$	$AVC$
0	100	100	—	—	—	—
1	146	100	46	146	46	46
2	178	100	78	89	32	39
3	202	100	102	67	24	34
4	224	100	124	56	22	31
5	250	100	150	50	26	30
6	286	100	186	48	36	31
7	338	100	238	48	52	34
8	412	100	312	52	74	39
9	514	100	414	57	102	46
10	650	100	550	65	136	55

## 9.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Total cost represents the value of the total resources used in the production of goods and services.
2. The shape of the cost curves depends on the nature of the cost functions.
3. Cost functions are derived from actual cost data of the firms.

## 9.5 SUMMARY

- Actual costs are those which are actually incurred by the firm in payment for labour, material, plant, building, machinery, equipment, travelling and transport, etc.
- The total money expenses recorded in the books of accounts are, for all practical purposes, the actual costs. Actual cost concept comes under the accounting cost concept.
- Opportunity cost is another fundamental cost concept used in business decisions. The concept of ‘opportunity cost’ is related to scarcity concept. The opportunity cost is the return expected from the second-best use of the resources, which is foregone for availing the gains from the best use of the resources.

- Business costs include all the expenses which are incurred in carrying out a business. The concept of business cost is similar to the actual or real cost. Business costs “include all the payments and contractual obligations made by the firm together with the book cost of depreciation on plant and equipment.”
- Fixed costs are those which are fixed in volume for a certain given output. Fixed costs do not vary with the variation in the output between zero and a certain level of output.
- Short-run costs can be defined as the costs which vary with the variation of output, the size of the firm remaining the same. In other words, short-run costs are the same as variable costs. Long-run costs, on the other hand, can be defined as the costs which are incurred on the fixed assets, like plant, building machinery, land, etc.
- Cost functions depend on (i) production function; and (ii) market-supply function of inputs. Production function specifies the technical input-combination and its relation to the output.
- From cost theory point of view, long run is the period during which all the inputs become variable. In the short run, at least one input (especially capital) remains fixed.
- The variability of inputs is based on the assumption that, in the long-run, supply of all the inputs, including those held constant in the short-run, becomes elastic.
- With fast expansion of the production scale, personal contacts and communication between (i) owners and managers, and (ii) managers and labour, get rapidly reduced. Close control and supervision is replaced by remote control and management.

---

## 9.6 KEY WORDS

---

- **Economy:** It is the state of a country or region in terms of the production and consumption of goods and services and the supply of money.
- **Diseconomy:** It is an economic disadvantage such as an increase in cost arising from an increase in the size of an organization.
- **Firm:** It is an organization that employs productive resources to obtain products and/or services which are offered in the market with the aim of making a profit.

---

## 9.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

### *Short-Answer Questions*

1. Write a short note on short run and long run costs. Discuss their importance.
2. What is the relationship between cost and output in short run.
3. Write short notes on the following:  
(a) Marginal cost (b) Implicit cost (c) Actual cost

### *Long-Answer Questions*

1. Discuss the different types of costs.
2. Differentiate between (i) Fixed and variable cost and (ii) Private and social cost.
3. What are cost functions? Describe the various types of cost functions.

---

## 9.8 FURTHER READINGS

---

Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.

Keat Paul G. and K.Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.

Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition. Singapore: Pearson Education, Inc.

Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Publishers.

---

## UNIT 10 REVENUE FUNCTION

---

### Structure

- 10.0 Introduction
- 10.1 Objectives
- 10.2 Meaning and Characteristics of Revenue Function
  - 10.2.1 Relationship between Total Revenue Profit and Total Costs
  - 10.2.2 Profit and Revenue
  - 10.2.3 Short Run Profit Maximization
- 10.3 Baumol's Model of Revenue Maximization
- 10.4 Answers to Check Your Progress Questions
- 10.5 Summary
- 10.6 Key Words
- 10.7 Self Assessment Questions and Exercises
- 10.8 Further Readings

---

### 10.0 INTRODUCTION

---

In this unit, you will study the revenue function in detail. There are primarily three ways to measure revenue, namely total revenue, average revenue and marginal revenue. This unit highlights the various concepts of revenue and revenue maximization. Baumol's model of revenue maximization is discussed in detail in this unit.

---

### 10.1 OBJECTIVES

---

After going through this unit, you will be able to:

- Explain meaning and characteristics of revenue function
- Discuss the relationship between profit and revenue
- Describe short-run profit maximization
- Assess Baumol's model of revenue maximization

---

### 10.2 MEANING AND CHARACTERISTICS OF REVENUE FUNCTION

---

Revenue is the income a firm retains from selling its products once it has paid indirect tax. Revenue provides the income which a firm needs to enable it to cover its costs of production, and from which it can derive a profit. Profit can be distributed to the owners, or shareholders, or retained in the business to purchase new capital assets or upgrade the firm's technology.

Revenue is measured in three ways:

- **Total revenue:** Total revenue (TR), is the total flow of income to a firm from selling a given quantity of output at a given price, less tax going to the government. The value of TR is found by multiplying price of the product by the quantity sold.
- **Average revenue:** Average revenue (AR), is revenue *per unit*, and is found by dividing TR by the quantity sold, Q. AR is equivalent to the price of the product, where  $P \times Q/Q = P$ , hence AR is also *price*.
- **Marginal revenue:** Marginal revenue (MR) is the revenue generated from selling one extra unit of a good or service. It can be found by finding the change in TR following an increase in output of one unit. MR can be both positive and negative.

### Revenue Function

Revenue is equal to the number of units sold times the price per unit. To obtain the revenue function  $R(x)$ , multiply the output level by the price function  $P(x)$ .

$$R(x) = X \cdot P(x)$$

#### 10.2.1 Relationship between Total Revenue Profit and Total Costs

Total revenue is the income a business receives from the sale of all the goods produced. Total profit is determined by subtracting total costs from revenues. Total revenue is determined by multiplying the price received for each unit sold by the number of units sold. Total revenue profits are a product of subtracting total costs from total revenue.

Price elasticity measures consumer responsiveness in relationship to quantity demanded and price per unit purchased. If producers can increase total revenue by lowering price, demand is considered elastic. If producers can increase total revenue by increasing price, demand is considered inelastic. Businesses receive maximum total revenue at the point when the greatest number of units can be sold for the highest possible price. Economists plot demand and price data on a graph to determine at what point price and demand will yield the highest total revenues.

#### 10.2.2 Profit and Revenue

In classical economics, it is assumed that firms will seek to maximise their profits. This occurs when the difference between TR (total revenue) – TC (total cost) is the greatest. Profit maximisation will also occur at an output where MR (marginal revenue) = MC (marginal cost). When  $MR > MC$ , the firm is increasing its total profit. When  $MR < MC$  total profit starts to fall. Therefore, profit is maximised where  $MR = MC$  (Figure 10.1).



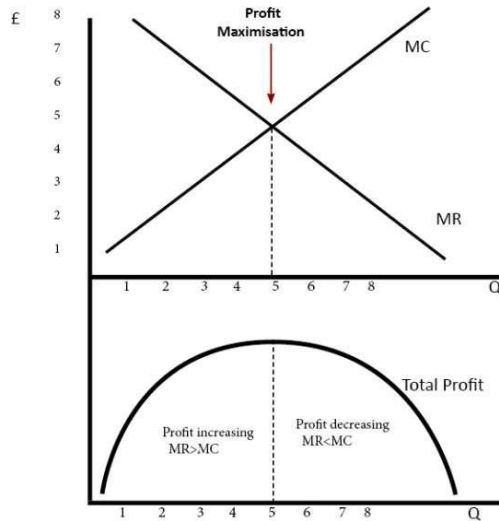


Fig. 10.1 Profit-Revenue Relationship

### 10.2.3 Short Run Profit Maximization

A firm maximizes its profits by choosing to supply the level of output where its marginal revenue equals its marginal cost. When marginal revenue exceeds marginal cost, the firm can earn greater profits by increasing its output. When marginal revenue is below marginal cost, the firm is losing money, and consequently, it must reduce its output. Profits are therefore maximized when the firm chooses the level of output where its marginal revenue equals its marginal cost.

---

## 10.3 BAUMOL'S MODEL OF REVENUE MAXIMIZATION

---

Baumol's theory of sales maximization is one of the most important alternative theories of firm's behaviour. The basic premise of Baumol's theory is that *sales maximization*, rather than *profit maximization*, is the plausible goal of the business firms. He argues that there is no reason to believe that all firms seek to maximize their profits. Business firms, as noted above, pursue a number of incompatible objectives and it is not easy to single out one as the most common objective pursued by the firms. However, research conducted at by Baumol revealed that

most managers seek to maximize sales revenue rather than profits. He argues that, in modern business, management is separated from ownership, and managers enjoy the discretion to pursue goals other than profit maximization. Their discretion eventually falls in favour of sales maximization.

According to research findings of Baumol, business managers pursue the goal of sales maximization for the following reasons.

**First**, financial institutions consider sales as an index of performance of the firm and tend to provide finance to the firm with growing sales.

**Secondly**, while profit figures are available only annually at the end of the final accounting year, sales figures can be obtained easily and more frequently to assess the performance of the management. Maximization of sales is more satisfying for the managers than the maximization of profits that go into the pockets of the shareholders.

**Thirdly**, salaries and slack earnings of the top managers are linked more closely to sales than to profit. Therefore, managers aim at maximizing sales revenue.

**Fourthly**, the routine personnel problems are more easily handled with growing sales. Higher payments may be offered to employees if sales figures indicate better performance.

Profits are generally known after a year. To rely on profit figures means, therefore, a longer waiting period for both the employees and the management for resolving labour problems.

**Fifthly**, where profit maximization is the goal and it rises in one period to an unusually high level, this becomes the standard profit target for the shareholders that managers find very difficult to maintain in the **long run**. Therefore, managers tend to aim at sales maximization rather than profit maximization.

**Finally**, sales growing at a rate higher than the rate of market expansion indicate growing market share, a greater competitive strength and better bargaining power of a firm in a collusive oligopoly. In a competitive market, therefore, sales maximization is found to be a more reasonable target.

To formulate his theory of sales maximization, Baumol<sup>3</sup> has developed two basic models: (i) Static Model and (ii) Dynamic Model—each with and without advertising. His static models with and without advertising are discussed below<sup>4</sup>.

### *Baumol's Model Without Advertising*

Baumol assumes cost and revenue curves to be given as in conventional theory of pricing. Suppose that the total cost ( $TC$ ) and the total revenue ( $TR$ ) curves are given as in Fig. 10.2. The total profit curve,  $TP$ , is obtained by plotting the difference between the  $TR$  and  $TC$  curves. Profits are zero where  $TR = TC$ . This point is indicated by the points of intersection of  $TR$  and  $TC$  curves.

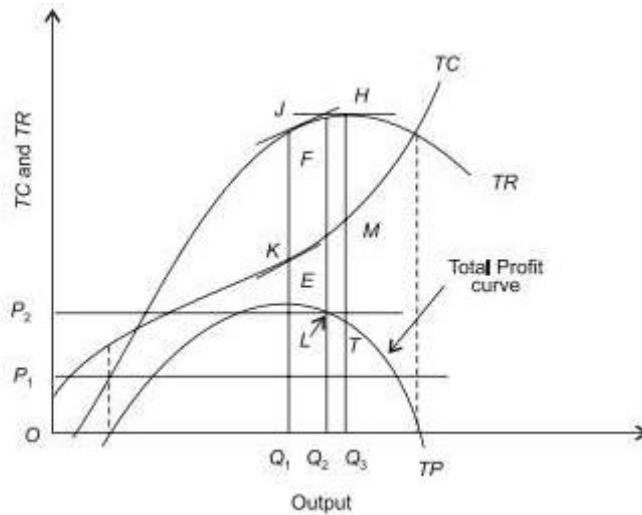


Fig. 10.2 Sales Revenue Maximization

Given the  $TR$  and  $TC$  curves, there is a unique level of output at which total sales revenue is maximum. The total sales revenue is maximum at the highest point of the  $TR$  curve. At this point, slope of the  $TR$  curve (i.e.,  $MR = "TR / "Q$ ) is equal to zero. The highest point on the  $TR$  curve can be obtained easily by drawing a line parallel to the horizontal axis and tangent to the  $TR$  curve. The point  $H$  on the  $TR$  curve in Fig. 10.2 represents the total maximum sales revenue. A line drawn from point  $H$  to output axis shows that sales revenue is maximized at output  $OQ_3$ . It implies that a sales revenue maximizing firm will produce output  $OQ_3$  and its price equals  $HQ_3/OQ_3$ .

### Profit Constraint and Revenue Maximization

At output  $OQ_3$ , the firm maximizes its total revenue. At this output, the firm makes a total profit equal to  $HM = \text{Total Revenue } HQ_3 \text{ less Total cost } MQ_3$ . Since total  $TP$  curve gives the measure of total profit at different levels of output, profit  $HM = TQ_3$ . If this profit is enough or more than enough to satisfy the stockholders, the firm will produce output  $OQ_3$  and charge a price =  $HQ_3/OQ_3$ . But, if profit at output  $OQ_3$  is not enough to satisfy the stockholders, then the firm's output must be changed to a level at which it makes a satisfactory profit, say  $OQ_2$ , which yields a profit  $LQ_2 > TQ_3$ .

Thus, there are two types of probable equilibrium: **one** in which the profit constraint does not provide an effective barrier to sales maximization,<sup>5</sup> and **second** in which profit constraint does provide an effective barrier to sales maximization. In the second type of equilibrium, the firm will produce an output that yields a satisfactory or target profit. It may be any output between  $OQ_1$  and  $OQ_2$ . For example, if minimum required profit is  $OP_1$ , then the firm will stick to its sales maximization goal and produce output  $OQ_3$  which yields a profit much greater than the required minimum. Since actual profit ( $TQ_3$ ) is much greater than the minimum required, the minimum profit constraint is not operative.

However, if required minimum profit level is  $OP_2$ , output  $OQ_3$  will not yield sufficient profit to meet the target profit. The firm will, therefore, produce an output which yields the required minimum level of profit  $OP_2 (=LQ_2)$ . Given the profit target  $OP_2$ , the firm will produce  $OQ_2$  where its profit is just sufficient to meet requirement of minimum profit. As can be seen in Fig. 10.2, output ( $OQ_2$ ) is less than the sales maximization output  $OQ_3$ . Evidently, the profit maximization output,  $OQ_1$  is less than the sales maximization output  $OQ_2$  (with profit constraint).

### Baumol's Model with Advertising

We have shown above how price and output are determined in a static model without advertising. In an oligopolistic market structure, however, determination of price and output is subject to non-price competition. Baumol has reconstructed his model with *advertising* as the typical form of *non-price competition* and suggests that the various forms of non-price competition may be analyzed on similar lines.

In his analysis of advertising, Baumol makes the following assumptions.

- Firm's objective is to maximize sales, subject to a minimum profit constraint;
- Advertising causes a shift in the demand curve and hence the total sales revenue ( $TR$ ) rises with an increase in advertisement expenditure ( $A$ ) i.e., " $TR/A > 0$ ";
- Price remains constant—a simplifying assumption, and
- Production costs are independent of advertising. This is rather an unrealistic assumption since increase in sales may put output at a different cost structure.

Baumol's model with advertising is presented in Fig. 10.3. The  $TR$  and  $TC$  are measured on the  $Y$ -axis and total advertisement outlay on the  $X$ -axis. The  $TR$  curve is drawn on the assumption that advertising increases total sales in the same manner as price reduction.

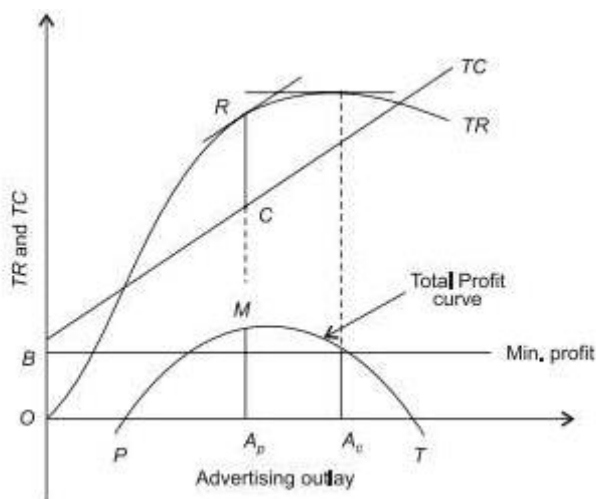


Fig. 10.3 Sales Revenue Maximisation

The  $TC$  curve includes both production and advertisement costs. The total profit curve is drawn by subtracting  $TC$  from  $TR$ . The profit so estimated is shown by the curve  $PT$ . As shown in Fig. 10.3 profit maximizing advertisement expenditure is  $OAp$  which maximizes profit at  $MAp$ . Note that  $MAp = RC$ . Assuming that minimum profit required is  $OB$ , the sales maximizing advertisement outlay would be  $OAc$ . This implies that a firm increases its advertisement outlay until it reaches the target profit level which is lower than the maximum profit. This also means that sales maximizers advertise more than the profit maximizers to capture a large market share.

### Criticism of Baumol's Model

Although Baumol's sales maximization model is found to be theoretically sound and empirically practicable, economists have pointed out the following shortcomings in his model.

**First**, it has been argued that in the **long-run**, Baumol's sales maximization hypothesis and the conventional hypothesis would yield identical results, because the *minimum required* level of profits would coincide with the *normal* level of profits.

**Second**, Baumol's theory does not distinguish between firm's equilibrium and industry equilibrium. Nor does it establish industry's equilibrium when all the firms are sales maximizers.

**Third**, it does not clearly bring out the implications of interdependence of the firm's price and output decisions. Thus, Baumol's theory ignores not only actual competition between the firms but also the threat of potential competition in an oligopolistic market.

**Fourth**, Baumol's claim that his solution is preferable to the solutions offered by the conventional theory, from a social welfare point of view, is not necessarily valid.

---

## 10.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

---

1. Total revenue is determined by multiplying the price received for each unit sold by the number of units sold. Total revenue profits are a product of subtracting total costs from total revenue.
2. A firm maximizes its profits by choosing to supply the level of output where its marginal revenue equals its marginal cost.

3. The basic premise of Baumol's theory is that sales maximization, rather than profit maximization, is the plausible goal of the business firms.
4. It has been argued that in the long-run, Baumol's sales maximization hypothesis and the conventional hypothesis would yield identical results, because the minimum required level of profits would coincide with the normal level of profits.

---

## 10.5 SUMMARY

---

- Revenue is the income a firm retains from selling its products once it has paid indirect tax. Revenue provides the income which a firm needs to enable it to cover its costs of production, and from which it can derive a profit.
- Total revenue (TR), is the total flow of income to a firm from selling a given quantity of output at a given price, less tax going to the government. The value of TR is found by multiplying price of the product by the quantity sold.
- Marginal revenue (MR) is the revenue generated from selling one extra unit of a good or service. It can be found by finding the change in TR following an increase in output of one unit. MR can be both positive and negative.
- Revenue is equal to the number of units sold times the price per unit. To obtain the revenue function  $R(x)$ , multiply the output level by the price function  $P(x)$ .
- Total revenue is the income a business receives from the sale of all the goods produced. Total profit is determined by subtracting total costs from revenues.
- Price elasticity measures consumer responsiveness in relationship to quantity demanded and price per unit purchased.
- In classical economics, it is assumed that firms will seek to maximise their profits.
- This occurs when the difference between TR (total revenue) – TC (total cost) is the greatest. Profit maximisation will also occur at an output where  $MR$  (marginal revenue) =  $MC$  (marginal cost).
- A firm maximizes its profits by choosing to supply the level of output where its marginal revenue equals its marginal cost. When marginal revenue exceeds marginal cost, the firm can earn greater profits by increasing its output.
- Baumol's theory of sales maximization is one of the most important alternative theories of firm's behaviour. The basic premise of Baumol's theory is that *sales maximization*, rather than *profit maximization*, is the plausible goal of the business firms.
- Baumol assumes cost and revenue curves to be given as in conventional theory of pricing.

- Although Baumol's sales maximization model is found to be theoretically sound and empirically practicable, economists have pointed out the following shortcomings in his model.

---

## 10.6 KEY WORDS

---

- **Total Revenue:** It is the income a business receives from the sale of all the goods produced.
- **Maximizer:** It means to increase to the greatest possible amount or degree.
- **Advertisement:** It is a notice or announcement in a public medium promoting a product, service, or event or publicizing a job vacancy.

---

## 10.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

### *Short-Answer Questions*

1. What are the three methods of measuring revenue?
2. What is short-run profit maximization?
3. Briefly mention the criticism of Baumol's model.

### *Long-Answer Questions*

1. Describe the relationship between total revenue profit and total costs. How are profit and revenue related?
2. Give a detailed description of Baumol's model of revenue maximization.
3. Explain Baumol's model with and without advertising.

---

## 10.8 FURTHER READINGS

---

Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.

Keat Paul G. and K.Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.

Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition. Singapore: Pearson Education, Inc.

Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Publishers.

---

**BLOCK - IV**

---

**DIFFERENT MARKET STRUCTURES**

---

---

**UNIT 11 MARKET STRUCTURE**

---

*Structure*

- 11.0 Introduction
- 11.1 Objectives
- 11.2 Characteristics of Market Structure
  - 11.2.1 Perfect Competition and Imperfect Competition
- 11.3 Features of Perfect Competition
- 11.4 Answers to Check Your Progress Questions
- 11.5 Summary
- 11.6 Key Words
- 11.7 Self Assessment Questions and Exercises
- 11.8 Further Readings

---

**11.0 INTRODUCTION**

---

The term “market” refers to a place where sellers and buyers meet and facilitate the selling and buying of goods and services. However, in economics, the term has a wider notion.

The market structure refers to the characteristics of the market, either organizational or competitive, that describes the nature of competition and the pricing policy followed in the market. We can define market structure as the number of firms producing the identical goods and services in the market and whose structure is determined on the basis of the competition prevailing in that market. In this unit, you will study market structure in detail.

---

**11.1 OBJECTIVES**

---

After going through this unit, you will be able to:

- Describe the characteristics of market structure
- Explain perfect competition and imperfect competition
- Discuss the features of perfect competition

---

**11.2 CHARACTERISTICS OF MARKET STRUCTURE**

---

The term market structure refers to the organisational features of an industry that influence the firm’s behaviour in its choice of price and output. Market structure is



an economically significant feature of the market. It affects the behaviour of firms in respect of their production and pricing behaviour.

Market structure is classified on the basis of organisational features of the industry, more specifically, on the basis of degree of competition among the firms. In general, the organisational features include the number of firms, distinctiveness of their products, elasticity of demand and the degree of control over the price of the product.

In this section, we present a brief description of the market structure, the playing field of the firms. This will give also an idea of the coverage and subject matter of this part of the book. The nature and the characteristics of different kinds of market will be discussed in detail in the subsequent chapters along with price and output determination. Here, we present only an overview of the market structure.

The market structure is generally classified on the basis of the degree of competition as follows:

- (i) Perfect Competition
- (ii) Imperfect Competition
  - (a) Monopolistic Competition
  - (b) Oligopoly with and without product differentiation
  - (c) Duopoly
- (iii) Monopoly

The basic features of these kinds of market are summarised in Table 11.1. However, a brief description of each kind of market is given below.

**Table 11.1** *Kinds of Market Structure*

<i>Type of Market</i>	<i>No. of Firms</i>	<i>Nature of Product</i>	<i>Firm's Control Over Price</i>
(i) <b>Perfect Competition</b>	Very large	Homogeneous (wheat sugar, vegetables....)	None
(ii) <b>Imperfect Competition</b>			
(a) <b>Monopolistic Competition</b>	Many (most retail trade)	Real or perceived difference in product	Some
(b) <b>Oligopoly</b>	Few	(i) Product without differentiation, e.g., bread, steel, and chemicals, sugar, etc.  (ii) Differentiated products (tea, toothpastes, soaps, detergents, automobiles)	Some
<b>Monopoly</b>	Single	Products without close substitutes, like gas, electricity and telephones	Full but usually regulated

**(i) Perfect Competition**

Perfect competition is a market situation in which a large number of producers offer a homogeneous product to a very large number of buyers of the product. The number of sellers is so large that each seller offers a very small fraction of the total supply, and therefore, has no control over the market price. Likewise, the number of buyers is so large that each buyer buys an insignificant part of the total supply and has no control over the market price. Both buyers and sellers are “price takers”, not “price makers”. The price of a commodity is determined in this kind of markets by the market demand and market supply. Each seller faces a horizontal demand curve (with  $e = \infty$ ), which implies that a seller can sell any quantity at the market determined price. **Each firm is in competition with so large a number of firms that there is virtually no competition.** This kind of market is however more of a hypothetical nature rather than being a common or realistic one. Some examples of a perfectly competitive market include stock markets, vegetable markets, wheat and rice mandis where goods are sold by auction.

*(ii) Imperfect Competition*

Perfect competition, in strict sense of the term, is a rare phenomenon. In reality, markets for most goods and services have imperfect competition. Imperfect competition is said to exist when a number of firms sell identical or differentiated products with some control over the price of their product. Barring a few goods like shares and vegetable markets, you name any commodity, its market is imperfect. In spite of a large number of dealers (*arhatias*) in wheat market, the Food Corporation of India is the biggest buyer and seller of wheat in India, with a great degree of control over wheat prices.

Imperfect competition creates two different forms of markets with different number of producers and with different degrees of competition, classified as (a) monopolistic competition, and (b) oligopoly.

**(a) Monopolistic Competitions**

Monopolistic competition is a kind of market in which a large number of firms supply differentiated products. The number of sellers is so large that each firm can act independently of others, without its activities being watched and countervailed by others. Besides, it is not only extremely difficult to keep track of competitors' strategy, but also it is not of any avail. In this respect, it is similar to perfect competition. It differs from perfect competition in that the products under monopolistic competition are somewhat differentiated whereas they are identical under perfect competition. There is free entry and free exit.

**(b) Oligopoly**

Oligopoly is an organisational structure of an industry in which a small number of firms supply the entire market, each seller having a considerable market share and

control over the price. Most industries in our country are oligopolistic. A small number of companies supply the entire sugar, tea, soaps, medicines, cosmetics, refrigerators, TV and VCRs, cars, trucks, jeeps, salt, vegetable oils (vanaspati), and so on. The producers of all these goods have some control over the price of their products. Their products are somewhat differentiated, at least made to look different in the consumers' perception. Products of different firms in industry are treated as close substitutes for one another, for example, Britannia and Modern breads. Therefore, demand curve for their product has high elasticity, but less than infinity, unlike under perfect competition.

### (c) Monopoly

Monopoly is the market of a single seller with control over his price and output. Monopoly is antithesis of perfect competition. Absolute monopolies are rare these days. They are found mostly in the form of government monopolies in public utility goods, e.g., electricity, telephone, radio broadcasting, water, gas, petrol and petroleum products, rail and postal services.

### Why Markets Are Imperfect?

Imperfect competition arises mainly from the barriers to entry. Barriers to entry are created by several factors.

**One**, the large size firms which enjoy economies of scale can cut down their prices to the extent that can eliminate new firms or prevent their entry to the industry, if they so decide.

**Two**, in some countries, like India, licencing policy of the government creates barrier for the new firms to enter an industry.

**Three**, patenting of rights to produce a well-established product or a new brand of a commodity prevents new firms from producing that commodity.

**Four**, sometimes entry of new firms to an industry is prevented by a law with a view to enabling the existing ones to have economies of scale so that prices are low.

In economics, a **market demand schedule** is a tabulation of the quantity of a good that all consumers in a market will purchase at a given price. **Market supply schedule** is the tabular statement showing the quantities that all the producers are willing to supply at given prices. To put it in simple words, it is the summation of various individual supply schedules.

---

## 11.3 FEATURES OF PERFECT COMPETITION

---

A perfectly competitive market is characterised by complete absence of rivalry among the individual firms. In fact, under perfect competition as conceived by the economists, competition among the individual firms is so widely dispersed that it amounts to no competition. Perfect competition is characterised by the following assumptions.

1. **Large Number of Buyers and Sellers.** Under perfect competition, the number of sellers is assumed to be so large that the share of each seller in the total supply of a product is so small that no single firm can influence the market price by changing its supply. Therefore, firms are *price-takers* not *price-makers*. Similarly, the number of buyers is so large that the share of each buyer in the total demand is so small that no single buyer or a group of buyers can influence the market price by changing their individual or group demand for a product.
2. **Homogeneous Product.** The commodities supplied by all the firms of an industry are assumed to be *homogeneous* or approximately identical. Homogeneity of the product implies that buyers do not distinguish between products supplied by the various firms of an industry. Product of each firm is regarded as a perfect substitute for the products of other firms. Hence, no firm can gain any competitive advantage over the other firms. This assumption limits the power of any firm to charge a price which is even slightly higher than the market price.
3. **Perfect Mobility of Factors of Production.** Another important characteristic of perfect competition is that the factors of production (especially, labour and capital) are freely mobile between the firms. Labour can freely change the firms as there is no barrier on labour mobility—legal, language, climate, skill, distance or otherwise. There is no trade union. Capital can also move freely from one firm to another. No firm has any kind of monopoly over any industrial input. This assumption guarantees that factors of production—labour, capital, and entrepreneurship—can enter or quit a firm or the industry whenever it is found desirable.
4. **Free Entry and Free Exit.** There is no legal or market barrier on entry of new firms to the industry. Nor is there any restriction on exit of the firms from the industry. That is, a firm may enter the industry and quit it at its will. Thus, when normal profit of the industry increases, new firms enter the industry and if profits decrease and better opportunities are available, firms leave the industry.
5. **Perfect Knowledge about the Market Conditions.** There is perfect knowledge about the market conditions. All the buyers and sellers have full information regarding the prevailing and future prices and availability of the commodity. As Marshall put it, "... though everyone acts for himself, his knowledge of what others are doing is supposed to be generally sufficient to prevent him from taking a lower or paying a higher price than others are doing." Information regarding market conditions is available free of cost. There is no uncertainty.
6. **No Government Interference.** Government does not interfere in any way with the functioning of the market. There are no taxes or subsidies; no licencing system, no allocation of inputs by the government, or any kind of other

direct control. That is, the government follows the *free enterprise* policy. Where there is intervention by the government, it is intended to correct the market imperfections.

7. **Absence of Collusion and Independent Decision-Making.** Perfect competition assumes that there is no collusion between the firms, i.e., they are not in league with one another in the form of *guild* or *cartel*. Nor are the buyers in collusion between themselves. There are no consumers' associations, etc. This condition implies that buyers and sellers take their decisions independently and they act independently.

### *Perfect vs. Pure Competition*

Sometimes a distinction is made between '**perfect competition**' and '**Pure Competition**'. The difference between the two is a matter of degree. While 'perfect competition' has all the features mentioned above, 'pure competition' does not assume *perfect mobility* of factors and *perfect knowledge*. That is, *perfect competition less perfect mobility and knowledge* is pure competition. 'Pure competition' is 'pure' in the sense that it has absolutely no element of monopoly.

The perfect competition, as characterised above, is considered as a rare phenomenon in the real business world. However, the actual markets that approximate the conditions of perfectly competitive market include the security markets for stocks and bonds, and agricultural markets like local vegetable markets. Despite its limited scope, perfect competition model has been the most popular model used in economic theories due to its analytical value.

---

## **11.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS**

---

1. Market structure is classified on the basis of organisational features of the industry, more specifically, on the basis of degree of competition among the firms.
2. Perfect competition is a market situation in which a large number of producers offer a homogeneous product to a very large number of buyers of the product.
3. The difference between 'perfect competition' and 'pure competition' is a matter of degree.

---

## 11.5 SUMMARY

---

- The term Market Structure refers to the organisational features of an industry that influence the firm's behaviour in its choice of price and output.
- Market structure is an economically significant feature of the market. It affects the behaviour of firms in respect of their production and pricing behaviour.
- Market structure is classified on the basis of organisational features of the industry, more specifically, on the basis of degree of competition among the firms.
- Perfect competition is a market situation in which a large number of producers offer a homogeneous product to a very large number of buyers of the product.
- The number of sellers is so large that each seller offers a very small fraction of the total supply, and therefore, has no control over the market price.
- Each firm is in competition with so large a number of firms that there is virtually no competition. This kind of market is however more of a hypothetical nature rather than being a common or realistic one.
- Perfect competition, in strict sense of the term, is a rare phenomenon. In reality, markets for most goods and services have imperfect competition. Imperfect competition is said to exist when a number of firms sell identical or differentiated products with some control over the price of their product.
- Monopolistic competition is a kind of market in which a large number of firms supply differentiated products. The number of sellers is so large that each firm can act independently of others, without its activities being watched and countervailed by others.
- Oligopoly is an organisational structure of an industry in which a small number of firms supply the entire market, each seller having a considerable market share and control over the price.
- Imperfect competition arises mainly from the barriers to entry.
- A perfectly competitive market is characterised by complete absence of rivalry among the individual firms.
- Under perfect competition, the number of sellers is assumed to be so large that the share of each seller in the total supply of a product is so small that no single firm can influence the market price by changing its supply. Therefore, firms are price-takers not price-makers.
- The commodities supplied by all the firms of an industry are assumed to be homogeneous or approximately identical.
- Sometimes a distinction is made between 'perfect competition' and 'Pure Competition'. The difference between the two is a matter of degree.

- While ‘perfect competition’ has all the features mentioned above, ‘pure competition’ does not assume *perfect mobility* of factors and *perfect knowledge*. That is, *perfect competition* less *perfect mobility and knowledge* is pure competition. ‘Pure competition’ is ‘pure’ in the sense that it has absolutely no element of monopoly.

---

## 11.6 KEY WORDS

---

- **Oligopoly:** It is a state of limited competition, in which a market is shared by a small number of producers or sellers.
- **Duopoly:** It is a situation in which two suppliers dominate the market for a commodity or service.
- **Homogenous:** It describes things that are all of the similar kind.

---

## 11.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

### *Short-Answer Questions*

1. Why are markets imperfect?
2. What are the features of perfect competition?
3. Define a homogenous product in your own words.

### *Long-Answer Questions*

1. What is market structure? Discuss the characteristics of market structure.
2. Give a detailed description of perfect competition and imperfect competition.
3. Distinguish between perfect and pure competition.

---

## 11.8 FURTHER READINGS

---

Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.

Keat Paul G. and K. Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.

Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition. Singapore: Pearson Education, Inc.

Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Publishers.

---

*Structure*

- 12.0 Introduction
- 12.1 Objectives
- 12.2 Pricing Under Different Market Structures
  - 12.2.1 Equilibrium and Supply Curve of the Firm
- 12.3 Price and Output Determination Under Perfect Competition
  - 12.3.1 Price and Output Determination in the Long Run
  - 12.3.2 Long-Run Supply Curve of a Competitive Industry
- 12.4 Answers to Check Your Progress Questions
- 12.5 Summary
- 12.6 Key Words
- 12.7 Self Assessment Questions and Exercises
- 12.8 Further Readings

---

**12.0 INTRODUCTION**

---

In this unit, we will study pricing under perfect competition. The other market structures will be dealt with in the next unit. Before learning about the pricing under different market structures, it is important that we learn about the concept of equilibrium and supply curve of the firm.

According to the traditional theory of firm, a firm is in equilibrium when its profit is maximum. Maximization of profits depends on the revenue and cost conditions. Revenue and cost conditions vary according to whether the period under reference is short or long. The equilibrium of the firm under short-run and long-run conditions are further discussed in this unit.

---

**12.1 OBJECTIVES**

---

After going through this unit, you will be able to:

- Explain pricing under different market structures
- Describe equilibrium and supply curve of a firm
- Discuss price and output determination under perfect competition

---

**12.2 PRICING UNDER DIFFERENT MARKET STRUCTURES**

---

The theory of firms came into existence during the 1930s with Joan Robinson's *The Economics of Imperfect Competition* and Edwin H. Chamberlin's *The Theory of Monopolistic Competition*, both written independently in 1933.

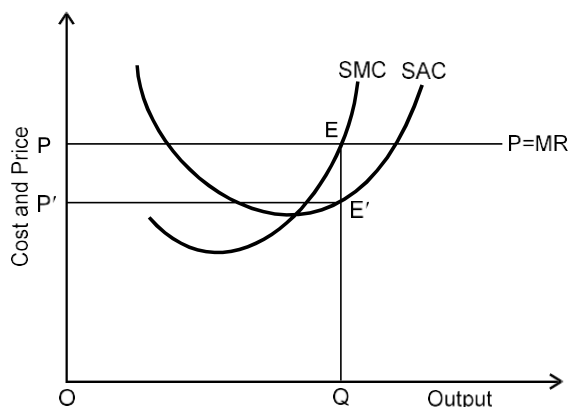


Earlier, the theory related to price determination was in the form of the ‘Theory of Value’ attributed to Alfred Marshall and his *Principles of Economics*. The theory of value propounded by Marshall, on the assumptions of perfect competition and a static equilibrium system, was regarded to provide answer to all questions regarding price and output determination. The existence of perfect competition was however challenged by Piero Sraffa. He showed that perfect competition was not logically consistent with partial equilibrium analysis. This led to the abandonment of the assumption of the perfect competition. Robinson and Chamberlin developed independently the theory of imperfect competition and theory of monopolistic competition, respectively. Joan Robinson and Chamberlin have demonstrated that price and output are determined by individual decisions under the condition of imperfect competition. They had however retained the earlier assumption of profit maximisation. This assumption was later challenged, and many new theories of firms were suggested. None of the theories has however received a universal acceptance.

Let us we begin our discussion on the theory of firm with the analysis of price and output determination in a perfectly competitive market.

### 12.2.1 Equilibrium and Supply Curve of the Firm

Here, short run refers to a period of time during which (i) price of the product is given in the market and the firm can sell any quantity at the prevailing price; (ii) plant-size of the firm is given; and (iii) the firm is faced with given short-run cost curves.



**Fig. 12.1** Equilibrium of the Firm

The firm’s equilibrium in the short run is illustrated in Fig. 12.1. Price of a commodity is fixed by the market forces in a perfectly competitive market. The firms, therefore, face a straight-line, horizontal demand curve, as shown by the line  $P = MR$ . It implies that price equals marginal revenue, i.e.,  $AR = MR$ . Firms short run average and marginal cost curves are shown by  $SAC$  and  $SMC$ , respectively.

It can be seen from Fig. 12.1 that  $SMC$  curve intersects the  $P = MR$  line at point  $E$ , from below, where  $SMC = MR$ . A perpendicular drawn from point  $E$  to the output axis, determines the equilibrium output at  $OQ$ . It means that output  $OQ$  meets both the first and second order conditions of profit maximization. At output  $OQ$ , therefore, profit is maximum. The output  $OQ$  is thus the equilibrium output. At this output, the firm is in equilibrium and is making maximum profit. Firm's maximum pure profit is shown by the area  $PEE'P$  which equals  $PP' \times OQ$  where  $PP'$  is the per unit abnormal profit at output  $OQ$ .

### Does a Firm Always Make Profit in the Short Run?

In the short-run equilibrium, a firm may not always make profits. In the short run, it may earn just a normal profit or even make losses. Whether a firm makes abnormal profits, normal profits, or losses depends on its cost and revenue conditions. If its short-run average cost ( $SAC$ ) is below the price ( $P = MR$ ) at equilibrium as shown in Fig. 12.1, the firm makes abnormal or pure profits. If its  $SAC$  is tangent to  $P = MR$  [Fig. 12.2(a)], the firm makes only normal profit as it covers only its  $SAC$  which includes normal profit. But, if its  $SAC$  falls above the price ( $P = MR$ ), the firm makes losses [Fig. 12.2(b)]. The total loss is shown by the area  $PP'EE'$  ( $= PP' \times OQ$ ), while per unit loss is  $PP' = EE'$ .

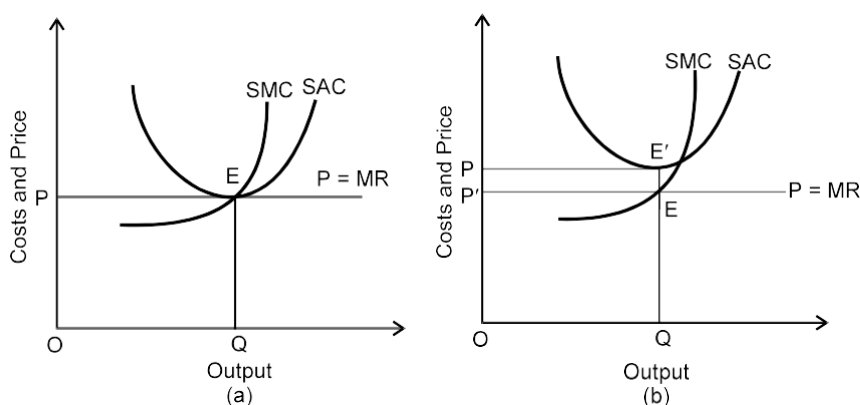
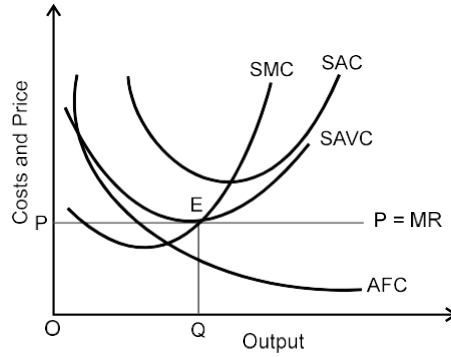


Fig. 12.2 Short-run Equilibrium of Firm

### Shut-down or Close-down Point

In case a firm is making loss in the short-run, it must minimise its losses. In order to minimise its losses, it must cover its short-run average variable cost ( $SAVC$ ). The behaviour of **short-run average variable cost** is shown by the curve  $SAVC$  in Fig. 12.3. A firm unable to recover its minimum  $SAVC$  will have to close down. The  $SAVC$  is minimum, where it equals the  $MC$ . The  $SMC$  intersects  $SAVC$  at its minimum level as shown in Fig. 12.3.

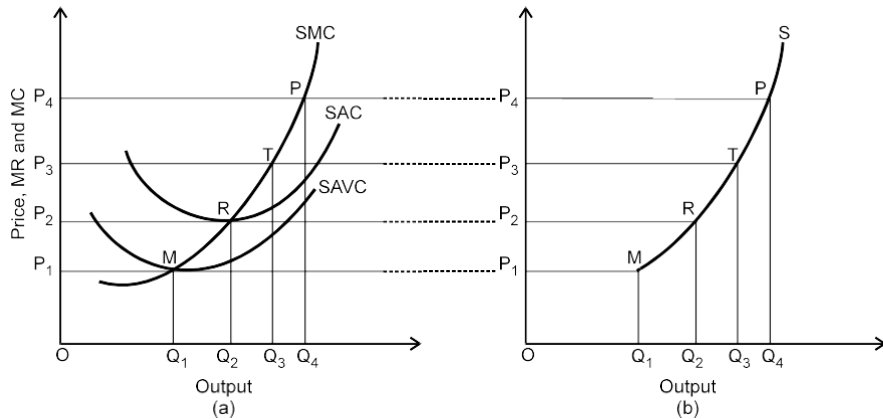


**Fig. 12.3** Shut-down Point

Another condition which must be fulfilled is  $P = MR = SMC$ . That is, for loss to be minimum,  $P = MR = SMC = SAVC$ . This condition is fulfilled at point *E* in Fig. 12.3. Point *E* denotes the ‘shut-down point’ or ‘breakdown point’ because at any price below  $OP$ , it pays firms to close down as it minimises its losses.

*Derivation of Supply Curve of the Firm*

The supply curve of an individual firm is derived on the basis of its equilibrium output at different levels of the price. The equilibrium output is determined by the intersection of *MR* and *MC* curves. The derivation of supply curve of a firm is shown in Fig. 12.4(a) and (b). The equilibrium level of output in the short-run is given at point *M*—the point of intersection between *MC* and *SAVC*. This is ‘breakdown point’ which gives the minimum supply of the firm in the short-run. The equilibrium level of output at this point is  $OQ_1$ . Let us suppose that price increases to  $OP_2$ . As a result, the equilibrium point shifts to *R* and output increases to  $OQ_2$ . Let the price further increase to  $OP_3$  so that equilibrium output increases to  $OQ_3$ . When price rises to  $OP_4$ , the equilibrium output rises to  $OQ_4$ . It may thus be concluded that as price increases, firm’s supply goes on increasing. That is, there is positive relation between price and supply. The price and output information contained in Fig. 12.4(a) is presented in the form of a supply curve, *MS*, in Fig. 12.4(b).



**Fig. 12.4** Derivation of Firm’s Supply Curve

Like market demand curve is a horizontal summation of individual demand curves, the industry supply curve or market supply curve is the horizontal summation of the supply curves of the individual firms. If cost curves of the individual firms of an industry have identical shape, their individual supply curves would also be identical. In that case, industry supply curve can be obtained by multiplying the individual supply at various prices by the number of firms. In the short-run, however, the individual supply curves may not be identical. If so, the market supply curve can be obtained by summing horizontally the individual supply curves. Let us suppose that there are only two firms having their individual supply curves  $S_1$  and  $S_2$  as shown in Fig. 12.5(a). At price  $OP_1$ , the industry supply equals  $P_1A + P_1B$  which equals  $P_1M$  in Fig. 12.5(b).

Similarly, at price  $OP_2$ , the industry supply equals  $P_2C + P_2C$  or  $2(P_2C)$  which equals  $P_2N$  in Fig. 12.5(b). In the same way, point  $T$  is located. By joining the points  $M$ ,  $N$  and  $T$ , we get the market or industry supply curve,  $SS'$ . The market supply curve so derived is used to show the determination of market price.

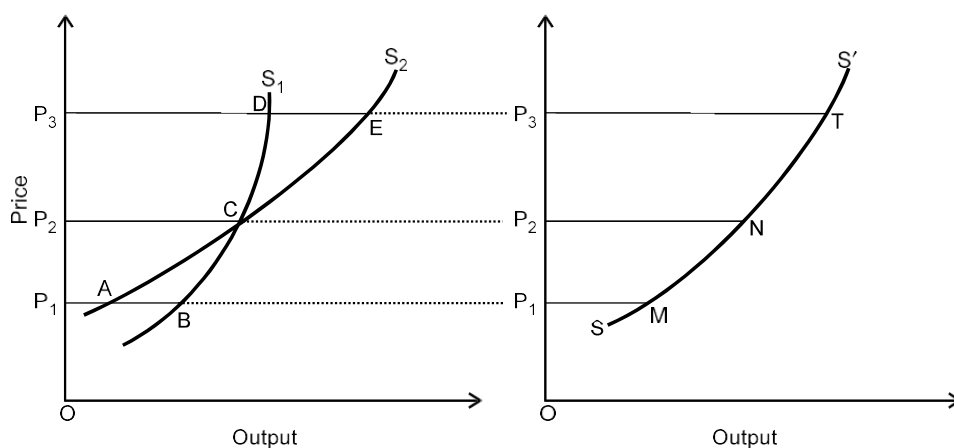


Fig. 12.5 Derivation of Industry Supply Curve

Now that we have derived the market supply curve, we may explain the price determination in a perfectly competitive market.

### 12.3 PRICE AND OUTPUT DETERMINATION UNDER PERFECT COMPETITION

Under perfect competition, market price in a perfectly competitive market is determined by the market forces, viz., demand and supply. Here, market demand

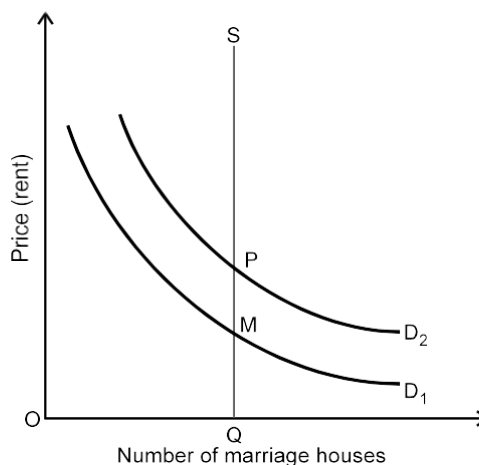
refers to the demand for the industry as a whole. It is equal to the sum of the quantity demanded by the individuals at different prices. Similarly, market supply is the sum of quantity supplied by the individual firms in the industry at a given price. The market price is therefore determined for the industry as a whole and is given for each individual firm and for each buyer. Thus, every seller in a perfectly competitive market is a 'price-taker', not a 'price-maker'.

In a perfectly competitive market, therefore, the main problem of a firm is not to determine the price of its product but to find its output at the given price so that profit is maximised.

The role of market forces and the mode of price determination depends on the time taken by supply position to adjust itself to the changing demand conditions. Price determination is analysed under three different time periods: (i) Market period or very short-run; (ii) short-run; and (iii) long-run. We will discuss below the price determination in the three periods.

### Price Determination in Very Short-Run

The market period or very short run refers to a time period in which quantity supplied is absolutely fixed or, in other words, supply response to change in price is nil. In the market period, therefore, the total output of the product is fixed. Each firm has a given quantity of commodity to sell. The aggregate supply of all the firms makes the market



**Fig. 12.6** Determination of Market Price

supply. The supply curve is *perfectly inelastic*, as shown by line  $SQ$  in Fig. 12.6. In this situation, price is determined entirely by the demand conditions. For instance, suppose that the number of marriage-houses (or tents) available per month in a city is given at  $OQ$  (Fig. 12.6), so that the supply curve takes the shape of a vertical straight line  $SQ$ . Let us also suppose that the monthly demand curve for marriage-houses is given by the demand curve,  $D_1$ . Demand and supply curves intersect each other at point  $M$ , determining the rental at  $MQ$ . Let us now suppose

that during a particular month demand for marriage-houses suddenly increases because a relatively large number of parents decide to celebrate the marriage of their daughters and sons due to, say, non-availability of auspicious dates for some time to come. Consequently, the demand curve shifts upward to  $D_2$ . The demand

curve  $D_2$  intersects the supply curve at point  $P$ . The equilibrium rate of rental is thus determined at  $PQ$ . This becomes parametric price for all the buyers. Note that the rise in the rental from  $MQ$  to  $PQ$  is caused by the upward shift in the demand curve and that market supply curve remains perfectly inelastic in the market period. The other example of very short-run markets may be of perishable commodities like fish, milk, vegetable, etc. and of non-perishable commodities like shares and bonds.

### Princing in the Short-Run

While in market period (or very short-run), supply is absolutely fixed, in the short-run it is possible to increase (or decrease) the supply by increasing (or decreasing) the variable inputs. In the short-run, therefore, supply curve is elastic, unlike a straight vertical line in the market period. Supply curve in the short-run approximates the  $SMC$  curve.

Under competitive conditions the process of price determination and output adjustment in the short-run is given in Fig. 12.7(a) and 12.7(b). Figure 12.7(a) shows demand curve  $DD$  and supply curve  $SS$  intersect at point  $P$  determining the price at  $OP_1$ . This price is fixed for all the firms in the industry.

Given the price  $PQ (= OP_1)$ , in Fig. 12.7(a), an individual firm can produce and sell any quantity at this price. But any quantity will however not yield maximum profit. The firms will have to adjust their output to the price  $OP_1$ . The process of output determination is presented through Fig. 12.7(b).

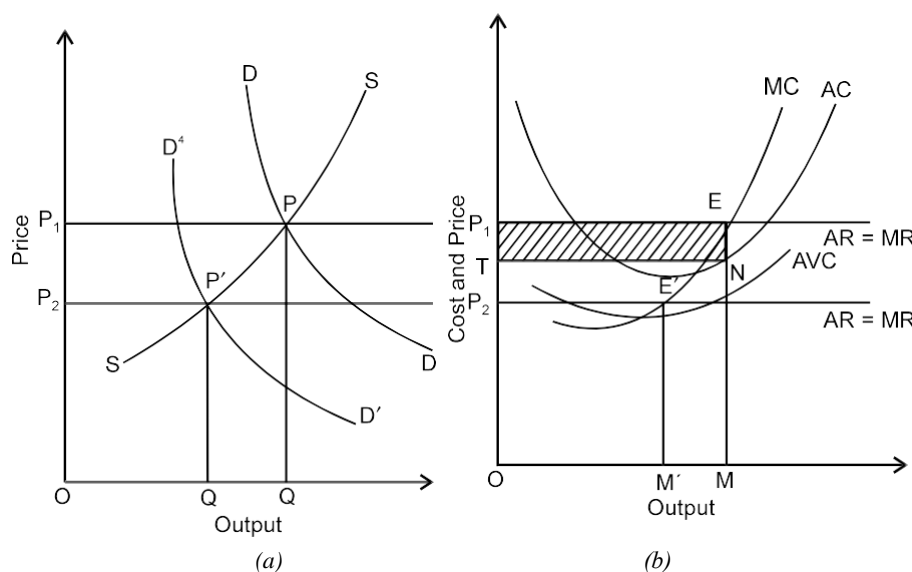


Fig. 12.7 Pricing under Perfect Competition: Short-run

Since a firm can sell any quantity at price  $OP_1$ , the demand for the firm's product is given by a horizontal straight line,  $AR = MR$ . Price being constant, its average revenue ( $AR$ ) and marginal revenue ( $MR$ ) are equal. Firm's upward sloping  $MC$  curve beyond its  $AVC$  curve represents its supply curve. Firm's  $MR$  and  $MC$  curves intersect each other at point  $E$ . This is firm's equilibrium point. The perpendicular  $EM$  determines the profit-maximising output at  $OM$ . At this output, firm's  $MR = MC$ , which satisfies both the first order and the second order conditions of maximum profit. The total maximum profit is shown by the area  $P_1TNE$ . The total profit ( $\pi$ ) may be calculated as

$$\pi = (AR - AC)Q$$

In Fig. 12.7(b),

$$AR = EM;$$

$$AC = NM;$$

and  $Q = OM$ .

By substituting the values from Fig. 12.7(b), we get

$$\pi = (EM - NM) OM$$

Since  $EM - NM = EN$ ,

$$\pi = EN \cdot OM$$

This is the maximum profit that a firm can make, given the cost and revenue conditions as presented in Fig. 12.7(b).

Now, if price falls to  $OP_2$  due to downward shift in the demand curve to  $D'D'$ , the firm will be in equilibrium at point  $E'$ . Here again firm's  $AR' = MR' = MC$ . But its  $AR < AC$ . Therefore, the firm incurs loss. But, in the short-run, it may not be desirable to close down so long as it covers its  $MC$ .

### Short-Run Equilibrium of the Industry

We have discussed above the equilibrium of the firm in the short run. To complete the discussion on short-run price and output determination, we discuss now the short-run equilibrium of the industry.

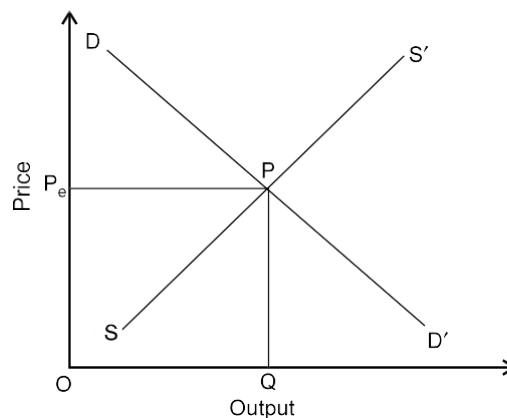


Fig. 12.8 Equilibrium of the Industry

An industry is in equilibrium in the short-run when market is cleared at a given price, i.e., when the total supply of the industry equals the total demand for its product. The price at which market is cleared is equilibrium price. The industry being in equilibrium, there is no tendency to expand or to contract the output. The equilibrium of industry is shown at point  $P$  in Fig. 12.8. The industry demand and supply curves intersect at point  $P$ , determining equilibrium price  $OP_e$ . The industry is supplying as much as consumers demand. In the short run equilibrium of the industry, individual firms may make pure profits, normal profits or losses, depending on their cost conditions.

### 12.3.1 Price and Output Determination in the Long Run

Unlike in the short-run, the supply curve in the long run is supposed to be more elastic. Long-run brings in two additional factors in operation which make the supply curve more elastic. **First**, in the long run, it becomes possible for the existing firms to increase their output by increasing the size of their plant. **Second**, and what is more important, new firms may enter and some existing ones may leave the industry. Entry and exit of firms bring about the long-run variation in the output. If cost and revenue conditions in the long run are such that some firms are making losses and are not able to adjust their plant-size and cost structure to the market price, such firms leave the industry. This makes the market supply curve shift leftward causing a rise in the price. The increase in market price increases the excess profit of the profit-making firms. Under the conditions of the perfect competition (i.e., free entry and exit), the pure profit would invite many new firms to the industry. This will make supply curve shift rightward, causing a decrease in the price, which will eventually take away the excess or pure profits. All firms earn only *normal profit*. Let us now explain the price and output determination in the long run and also the equilibrium of the firm and of the industry.

#### Price Determination in the Long Run

As in the short-run, market price is determined in the long-run by the market forces of demand and supply. Let us suppose that the market demand curve is given by  $DD_1$  which is relevant for both short-run and long-run, and short-run

supply curve is given by  $SS_1$  in Fig. 12.9(a). The market demand curve  $DD_1$  and market supply curve  $SS_1$  intersect each other at point  $P_1$  and the short-run market price is determined at  $OP_0$ . At this price, the firms find their short-run equilibrium at point  $E_1$  and each of them produces output  $OQ_1$ . The total market supply equals  $OQ_1 \times \text{No. of firms} = ON_1$  [in panel (a) of Fig. 12.9] and the industry is in short-run equilibrium.

Given the cost and revenue conditions in Fig. 12.9(b), the firms are making super normal profit of  $E_1M$  per unit. The existence of super normal profit in the short run leads to increase in the market supply on two accounts: *one*, new firms will enter the industry attracted by the super normal profits, and *two*, the existing firms would expand their plant-size because returns to scale would increase as



shown by the *LAC*. As a result, the market supply would increase so that supply curve shifts rightward to *SS*<sub>2</sub> [Fig. 12.9(a)]. The shift in supply curve brings down the market price to *OP* which is the long-run equilibrium price. Thus, equilibrium price is once again determined in the market.

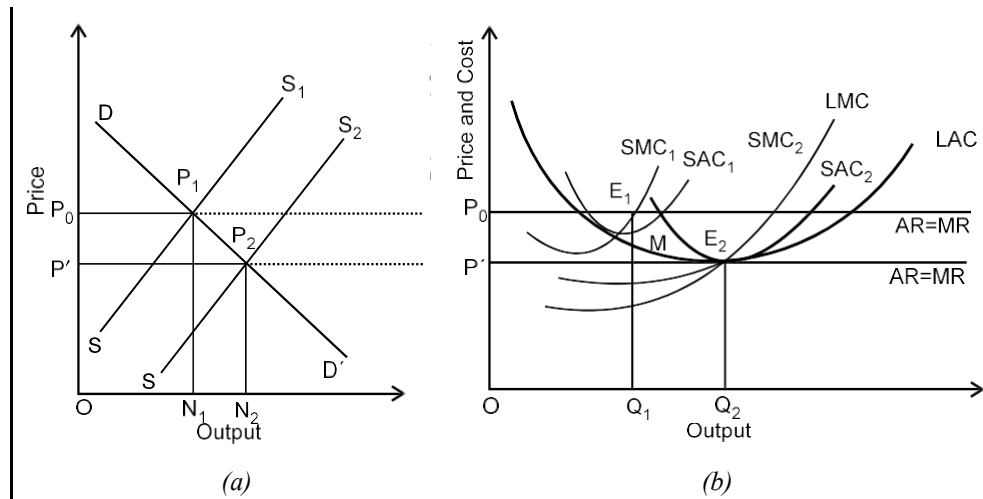


Fig. 12.9 Long-run Equilibrium of the Firm

### Equilibrium of the Firm in the Long Run

The firms are in equilibrium in the long-run when their

$$AR = MR = LMC = LAC = SMC = SAC$$

That is, the firms of an industry reach their equilibrium position in the long-run where both a short-run and long-run equilibrium conditions coincide. In a perfectly competitive market, the cost and revenue conditions are given for the firms. What the firms can do, therefore, is to adjust their output to the given revenue and cost conditions in order to maximise their profit. Let us now illustrate the process of adjustment of output so as to reach the equilibrium in the long run.

Suppose that the firms are in equilibrium at point *E*<sub>1</sub> in Fig. 12.9(a) where they make excess profits  $AR - SAC_1 = EM$  per unit. This gives incentives to the firms to expand their scale of production, i.e., they add more plants to the existing ones. As a result, market supply increases. Besides, supply increases also because new firms enter the industry. Therefore, the market supply curve *SS*<sub>1</sub> tends to shift rightward causing a fall in price to *OP*'. On the other hand, due to increase in demand for inputs, cost tends to rise. But so long as economies of scale are greater than the diseconomies of scale, the *LAC* tends to decrease and it pays firms to expand their plant-size. When a stage is arrives where  $P < LAC$ , firms incur losses. The firms which are not able to make adjustment in the plant-size or scale of production leave the industry. This works in two directions. On the one hand, supply decreases and price increases, and on the other, demand for inputs decreases which causes a decrease in the input prices. This process of adjustment continues until  $LAC$  is tangent to  $P = AR = MR$  for each firm in the industry.

This position is shown at point  $E_2$  in Fig. 12.10(b). Eventually, at point  $E_2$ , i.e., at the point of equilibrium,

$$P = MR = LMC = LAC = SMC = SAC$$

Since  $P = LAC$ , the firms make only normal profits in the long-run. If firms deviate from point  $E_2$ , due to some short-run disturbances, the market forces will make them return to this point only.

### Equilibrium of the Industry

An industry is in equilibrium when its market demand equals its market supply. When an industry is in equilibrium, all its firms are supposed to be in equilibrium [as shown in Fig. 12.9(b)]. When an industry is in equilibrium, all its firms earn only normal profits, because under the conditions of perfect competition all the firms are assumed to achieve the same level of efficiency in the long run. Since industry yields only normal profits, there is no incentive for new firms to enter the industry. These conditions are fulfilled at price  $OP_1$  in Fig. 12.9(a) and (b).

$$LMC = LMR = SMC = SAC = P = LAC$$

Since  $P = LAC$ , all the firms are earning only normal profits. At industry's equilibrium output  $ON_1$ , market demand equals market supply [Fig. 12.9(a)]. At price  $OP_1$ , therefore, market is cleared. The output  $ON_1$  has a fair chance to remain stable in the long run. For, there is no incentive for new firms to enter the industry and for existing ones to leave the industry. The industry is therefore in equilibrium.

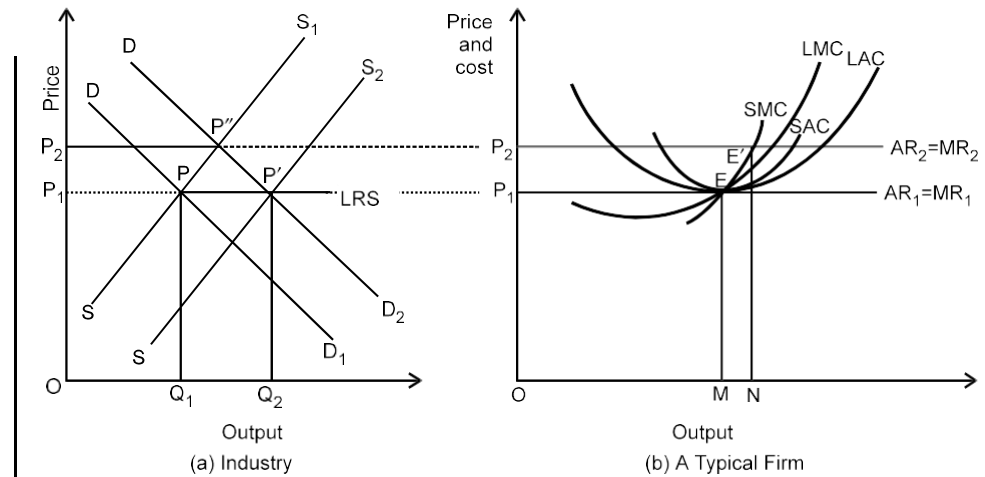
### 12.3.2 Long-Run Supply Curve of a Competitive Industry

We have earlier derived the short-run supply curve of the industry by summing horizontally the individual supply curves of the firms, which are based on their short-run  $MC$  curves. The long-run supply curve of a competitive industry, however, has nothing to do with the  $LMC$  curves. The shape of the long-run supply curve of an industry, under perfect competition, depends on whether factor prices remain *constant* or *increase* when demand for inputs increases in the long-run as a result of expansion in the output of the industry. According to whether cost is constant, increasing or decreasing, an industry is termed as *constant cost*, *increasing cost* or *decreasing cost* industry. Let us now derive the long-run supply curve of each of such industries.

#### Constant Cost Industry

An industry in which factor prices are independent of the rate of increase in factor demand is referred to as *constant cost industry*. In other words, when the expansion of output in an industry does not entail an increase in factor prices, the industry is said to be a constant-cost industry. The derivation of the supply curve of such an industry is illustrated in Fig. 12.10(a) and (b). Let us suppose that the industry is in equilibrium at  $P$  where demand curve  $DD_1$  and supply curve  $SS_1$

intersect each other. The industry is in equilibrium at price  $OP_1$  and output  $OQ_1$ . At price  $OP_1$ , all firms are in equilibrium as their  $LMC = P = MR = SMC = SAC$ .



**Fig. 12.10** Long-run Supply Curve of the Constant Cost Industry.

Suppose now that demand curve shifts to  $DD_2$  due to increase in consumers' income or increase in population or due to both. As a result market price increases to  $OP_2$ . In the short-run, this increase in price causes increase in supply by the firms from  $OM$  to  $ON$  [Fig. 12.10(b)] determined by the point of intersection of firm's  $SMC$  and new price line through  $P_2$ . The firms enjoy abnormal profits in the short-run.

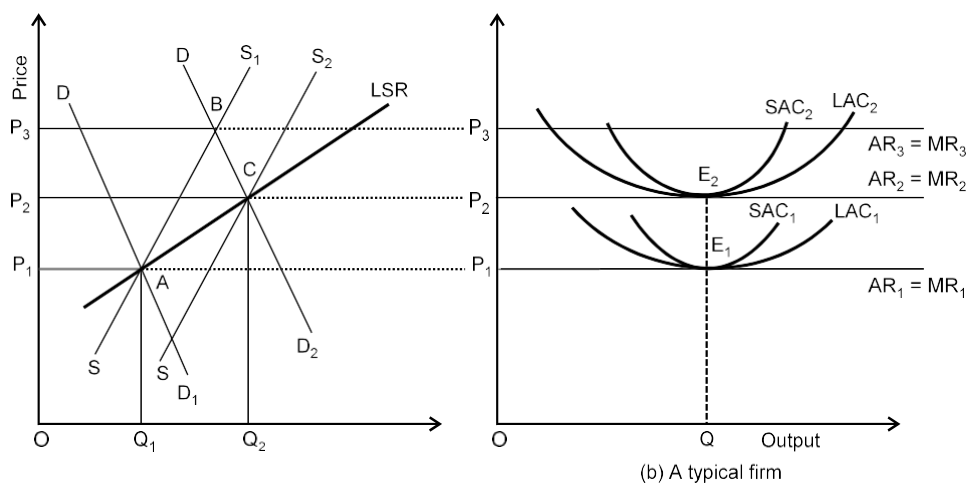
The abnormal profits attract new firms into the industry. The entry of new firms leads to increase in demand for factors. The industry being a constant-cost industry, factor prices do not increase. Cost of production for both new and old firms remain constant at the previous level. But, due to the entry of new firms, market supply increases and market supply curve shifts to  $SS_2$  [Fig. 12.10(a)]. Consequently, in the long-run, market price falls to its previous level,  $OP_1$ , and individual firms return to their previous equilibrium point  $E$ . But the industry output increases from  $OQ_1$  to  $OQ_2$ , since industry moves to a new equilibrium  $P'$ . By joining the two industry-equilibrium points,  $P$  and  $P'$ , we get long-run supply curve ( $LRS$ ) of the constant cost industry. Obviously, the long-run supply curves ( $LRS$ ) of a constant cost industry is a horizontal straight line, as given by the line  $LRS$ .

### Increasing Cost Industry

An industry is referred to as an *increasing cost industry* if factor prices increase due to increase in demand for inputs. The long-run supply curve of an increasing cost industry has a positive slope.

The derivation of long-run market supply curve under increasing cost condition is demonstrated in Fig. 12.11. Let the original demand and supply curves of the industry be represented, by  $DD_1$  and  $SS_1$ , respectively, and industry be in equilibrium at point  $A$ . Let us now suppose that for some reasons, demand curve  $DD_1$  shifts rightward to  $DD_2$ . As a result, short-run market price increases from

$OP_1$  to  $OP_3$ . With this increase in price, the demand curve for the individual firms shifts upward to  $AR_3 = MR_3$  [Fig. 12.11(b)]. The firms, therefore, enjoy super normal or economic profits. This profit attracts new firms into the industry and demand for inputs increases. Since in an increasing cost industry, the supply of inputs is assumed to be less than infinitely elastic, the entry of new firms causes an increase in the input prices. Consequently, cost curves, both short-run and long-run, shift upward from  $LAC_1$  to  $LAC_2$ . In this process of adjustments, however, industry-supply increases so that market supply curve  $SS_1$  shifts rightward to  $SS_2$ . With this shift in supply curve, the industry reaches another equilibrium position at point  $C$  where new demand and supply curves intersect each other. A new market price  $OP_2$  is determined. At price  $OP_2$  [Fig. 12.11(b)], the long-run and short-run cost curves ( $LAC_2$ , etc.) are tangent to the price line ( $OP_2$ ). The firms shift to a new long-run equilibrium,  $E_2$ , its output remaining the same. Whether equilibrium output of the firms remains constant, increases or decreases, depends, respectively, on whether cost curves shift upward vertically, upward to the right or upward to the left.



(a) Industry  
**Fig. 12.11** Long-run Supply Curve of an Increasing Cost Industry

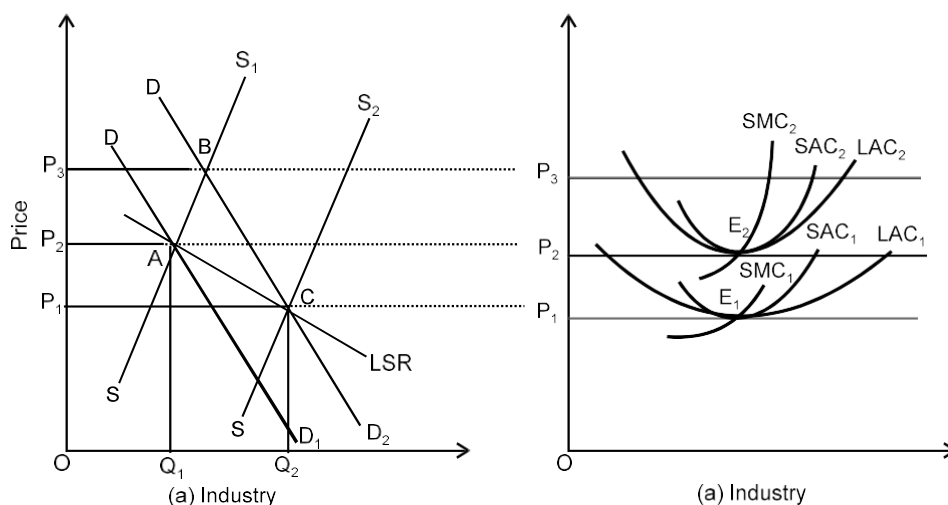
Note that at price  $OP_2$ , both industry and individual firms are in equilibrium. In the absence of any further disturbance, the equilibrium of both firms and industry has a fair chance to remain stable. Thus, at the new equilibrium price  $OP_2$ , the industry-output increases from  $OQ_1$  to  $OQ_2$  and corresponding equilibrium points are  $A$  and  $C$ , respectively. By joining the long-run equilibrium points  $A$  and  $C$ , we get the **long-run supply curve** for the industry, as shown by the curve  $LSR$ . Obviously, the  $LSR$  has a positive slope in an increasing cost industry.

### Decreasing Cost Industry

If expansion of output of an industry is associated with or leads to decrease in the input prices, the industry is referred to as a decreasing cost industry. A decreasing cost industry has a long-run industry supply curve with a negative slope, since input prices decrease with the expansion of the industry output.

The derivation of long-run industry supply curve (*LRS*) for a decreasing cost industry is illustrated in Fig. 12.12(a) and (b). Let the industry be initially in equilibrium at point *A* [Fig. 12.12(a)] and firms at *E*<sub>2</sub> [Fig. 12.12(b)]. Now suppose that demand curve shifts from *DD*<sub>1</sub> to *DD*<sub>2</sub> and, consequently, price rises from *OP*<sub>1</sub> to *OP*<sub>2</sub>. The short-run equilibrium of firms at price *OP*<sub>2</sub> [Fig. 12.12(b)] moves

upward on the *SMC*<sub>2</sub> where the firms make abnormal profits. The abnormal profits attract new firms to the industry causing increase in demand for inputs. If input industries are enjoying increasing returns to scale due to economies of scale, the increase in demand for inputs would encourage increased supply of inputs. Increase in the supply of inputs causes input prices to fall. The industry therefore, enjoys the external economies to scale. As a result, their long-run and short-run cost curves shift downward, from *LAC*<sub>2</sub> to *LAC*<sub>1</sub> [Fig. 12.12(b)].



**Fig. 12.12** Long-run Supply Curve of a Decreasing Cost Industry

From the industry’s point of view, industry supply increases due to the entry of the new firms, even if the existing firms maintain their old level of output. Therefore, the industry supply curve shifts from *SS*<sub>1</sub> to *SS*<sub>2</sub> which intersects the new demand curve *DD*<sub>2</sub> at point *C*. Thus, equilibrium of the industry shifts from point *A* to point *C*. Industry output increases from *OQ*<sub>1</sub> to *OQ*<sub>2</sub>. In the absence of any external disturbance, the industry equilibrium point *C* would tend to stabilise. By joining the two equilibrium points *A* and *C*, we get the long-run supply curve of the industry, *LRS*. The *LRS* has a negative slope.

**Is Decreasing Cost a Reality?**

Some authors argue that the ‘phenomenon of decreasing cost... is not consistent with all the requirements of the perfect competition.’ However, the possibility of a decreasing cost industry cannot be ruled out in a very long period. *One* reason for this is the likelihood of existence of large *external economies* of sale, particularly in case of young industries in undeveloped areas. An increase in the number of industries and the consequent growth of transportation, marketing facilities and financial institutions may reduce the industry’s cost of production. Nevertheless, it

depends on how substantial are the external economies of scale. R.G. Lipsey has cited the car industry of England as an example of decreasing cost industry. In his own words, 'As the output of cars increased, the industry's demand for types grew greatly. This... would have increased the demand for rubber and tended to raise its price, but it also provided the opportunity for tyre manufacturers to build large modern plants and reap the benefits of increasing returns in tyre production. At first these economies were large enough to offset any factor-price increases and tyre-price charged to car manufacturers fall. Thus car costs fell because of lower prices of an important input.' Whether industry-costs are constant or decreasing depends also on the proportion of total input supplies which is used by the industries. For example, output of pencil industry can be increased without substantially affecting the lumber prices, as pencil industry uses a small proportion of the total lumber output. But increase in output of furniture industry will affect the lumber price. Similarly, output of pin-industry can be substantially increased without affecting the steel price. But, output of car industry cannot be substantially increased without affecting steel prices. Another factor which may cause rise in input-prices is whether or not input industries enjoy economies of scale.

Moreover, the most common cases are of the constant and increasing cost industries. Decreasing cost industries are most unlikely to exist. The constant and decreasing cost industries tend over time to become increasing cost industries. Because external economies have a limit to which cost can be decreased, external diseconomies would exceed the external economies in the very long period.

## 12.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. In case a firm is making loss in the short-run, it must minimise its losses. In order to minimise its losses, it must cover its short-run average variable cost (SAVC).
2. The supply curve of an individual firm is derived on the basis of its equilibrium output at different levels of the price. The equilibrium output is determined by the intersection of MR and MC curves.
3. The role of market forces and the mode of price determination depends on the time taken by supply position to adjust itself to the changing demand conditions.
4. Price determination is analysed under three different time periods: (i) Market period or very short-run; (ii) short-run; and (iii) long-run.

---

## 12.5 SUMMARY

---

- The theory of firms came into existence during the 1930s with Joan Robinson's *The Economics of Imperfect Competition* and Edwin H. Chamberlin's *The Theory of Monopolistic Competition*, both written independently in 1933.
- Earlier, the theory related to price determination was in the form of the 'Theory of Value' attributed to Alfred Marshall and his *Principles of Economics*.
- In case a firm is making loss in the short-run, it must minimise its losses. In order to minimise its losses, it must cover its short-run average variable cost (SAVC).
- The supply curve of an individual firm is derived on the basis of its equilibrium output at different levels of the price. The equilibrium output is determined by the intersection of MR and MC curves.
- Like market demand curve is a horizontal summation of individual demand curves, the industry supply curve or market supply curve is the horizontal summation of the supply curves of the individual firms.
- Under perfect competition, market price in a perfectly competitive market is determined by the market forces, viz., demand and supply. Here, market demand refers to the demand for the industry as a whole.
- The role of market forces and the mode of price determination depends on the time taken by supply position to adjust itself to the changing demand conditions.
- Price determination is analysed under three different time periods: (i) Market period or very short-run; (ii) short-run; and (iii) long-run. We will discuss below the price determination in the three periods.
- An industry is in equilibrium in the short-run when market is cleared at a given price, i.e., when the total supply of the industry equals the total demand for its product.
- Unlike in the short-run, the supply curve in the long run is supposed to be more elastic. Long-run brings in two additional factors in operation which make the supply curve more elastic.
- An industry is in equilibrium when its market demand equals its market supply. When an industry is in equilibrium, all its firms are supposed to be in equilibrium.
- An industry in which factor prices are independent of the rate of increase in factor demand is referred to as constant cost industry.
- An industry is referred to as an increasing cost industry if factor prices increase due to increase in demand for inputs. The long-run supply curve of an increasing cost industry has a positive slope.

---

## 12.6 KEY WORDS

---

- **Profit:** It is a financial gain, especially the difference between the amount earned and the amount spent in buying, operating, or producing something.
- **Industry:** It is an economic activity concerned with the processing of raw materials and manufacture of goods in factories.

---

## 12.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

### *Short-Answer Questions*

1. Briefly state equilibrium of the firm in the short run.
2. Does a firm always make profit in the short run? Give reasons for your answer.
3. What is short-run equilibrium of the industry?
4. Write a short note on decreasing cost industry.

### *Long-Answer Questions*

1. Explain price determination in very short-run.
2. Discuss in detail the price and output determination in the long run.
3. Describe the equilibrium of the firm and the industry in the long run.
4. What do you understand by constant cost industry and increasing cost industry?

---

## 12.8 FURTHER READINGS

---

Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.

Keat Paul G. and K.Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.

Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition. Singapore: Pearson Education, Inc.

Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Publishers.



---

## UNIT 13 MONOPOLY AND

# MONOPOLISTIC COMPETITION

---

### Structure

- 13.0 Introduction
- 13.1 Objectives
- 13.2 Monopoly: Meaning and Features
  - 13.2.1 Profit Maximization and Equilibrium under Monopoly
- 13.3 Price Discrimination by Monopoly
- 13.4 Monopolistic Competition: Meaning and Features
  - 13.4.1 Foundations of Monopolistic Power
- 13.5 Price and Output Determination Under Monopolistic Competition
- 13.6 Answers to Check Your Progress Questions
- 13.7 Summary
- 13.8 Key Words
- 13.9 Self Assessment Questions and Exercises
- 13.10 Further Readings

---

## 13.0 INTRODUCTION

---

The term *pure monopoly* signifies an absolute power to produce and sell a product which has no close substitute. In other words, a monopoly market is one in which there is only one seller of a product having no close substitute. The cross-elasticity of demand for a monopolised product is either zero or negative. In a monopolised market structure, the industry is a single-firm-industry. Firm and industry are identical in a monopoly setting.

Moreover, the precise definition of monopoly has been a matter of opinion and purpose. For instance, in the opinion of Joel Dean, a monopoly market is one in which 'a product of lasting distinctiveness is sold.' The monopolised product has distinct physical properties recognised by its buyers and the distinctiveness lasts over many years. Such a definition is of practical importance if one recognises the fact that most of the commodities have their substitutes varying in degree and it is entirely for the consumers or users to distinguish between them and to accept or reject a commodity as the substitute. Another concept of pure monopoly has been advanced by D.H. Chamberlin who envisages the control of all goods and services by the monopolist. But such a monopoly has hardly ever existed, hence his definition is unrealistic. In the opinion of some others, any firm facing a sloping demand curve is a monopolist. This definition however includes all kinds of firms except those under perfect competition. We will, however, adopt for our purpose

here a general definition of a pure monopoly: a pure monopoly means an absolute power to produce and sell a commodity which has no close substitute.

This unit discusses monopoly and monopolistic competition in detail.

---

## 13.1 OBJECTIVES

---

After going through this unit, you will be able to:

- Explain the meaning and features of monopoly
- Evaluate demand and revenue curves under monopoly

---

## 13.2 MONOPOLY: MEANING AND FEATURES

---

The emergence and survival of a monopoly is attributed to the factors which prevent the entry of other firm into the industry. The barriers to entry are therefore the sources of monopoly power. The major sources of barriers to entry to a monopolised market are described here briefly.

- Legal Restrictions.** Some monopolies are created by the law in the public interest. Most of the state monopolies in the public utility sector, including postal, telegraph and telephone services, radio and TV services, generation and distribution of electricity, railways, airlines and state roadways, etc. are public monopolies that are created by the public law. The state may create monopolies in the private sector also by restricting entry of other firms by law or by granting patent rights. Such monopolies are intended to reduce cost of production to the minimum by enlarging the size and investing in technological innovations. Such monopolies are known as *franchise monopolies*.
- Control over Key Raw Materials.** Some firms acquire monopoly power because of their traditional control over certain scarce and key raw materials, which are essential for the production of certain other goods, e.g., bauxite, graphite, diamond, etc. For instance, Aluminium Company of America had monopolised the aluminium industry before World War II because it had acquired control over almost all sources of bauxite supply. Such monopolies are often called '*raw material monopolies*'. The monopolies of this kind emerge also because of monopoly over certain specific knowledge or technique of production.
- Efficiency.** A primary and technical reason for growth of monopolies is the economies of scale. In some industries, long-run minimum cost of production or the most efficient scale of production almost coincides with the size of the market. Under this condition, the large-size firm finds it profitable in the long-run to eliminate the competition by cutting down its price for a short

period. Once monopoly is established, it becomes almost impossible for the new firms to enter the industry and survive. Monopolies existing on account of this factor are known as **natural monopolies**. A natural monopoly emerges either due to technical efficiency or is created by the law on efficiency grounds.

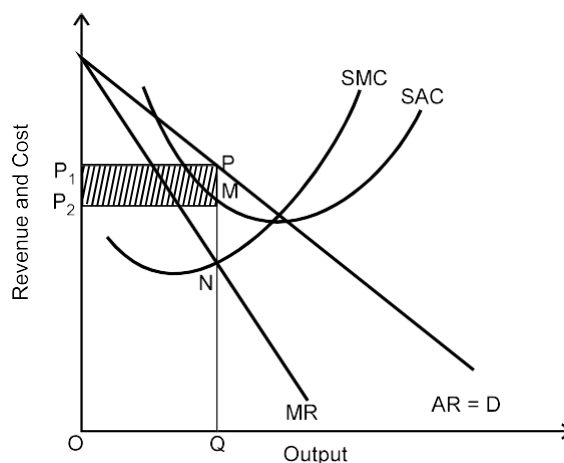
- (iv) **Patent Rights.** Another source of monopoly is the patent right of the firm for a product or for a production process. Patent rights are granted by the government to a firm to produce a commodity of specified quality and character or to use a specified technique of production. Patent rights gives a firm exclusive rights to produce the specified commodity or to use the specified technique of production. Such monopolies are called *patent monopolies*.

### 13.2.1 Profit Maximization and Equilibrium under Monopoly

The objective of a monopoly firm, like all other firms, is assumed to be profit maximisation. Profit maximisation is however not necessarily the sole objective of the firm. The monopoly firm may seek to maximise its utility function, particularly where management of the firm is divorced from its ownership. But, as mentioned earlier, most common objectives of business firm assumed in traditional theory of pricing is profit maximisation. We will therefore explain the equilibrium of monopoly firm in short run and long run under profit maximisation hypothesis.

#### Monopoly Equilibrium in the Short Run

Like any other firm, a monopoly firm reaches its equilibrium where it maximises its total profits. As noted earlier, profits are maximum where the two following conditions are fulfilled: (i) that  $MC = MR$ —the necessary condition, and (ii) that the  $MC$  curve must intersect the  $MR$  curve from below under increasing cost condition—the supplementary condition. The monopoly firm fixes its price and output in accordance with the these conditions.



**Fig. 13.1** Price Determination under Monopoly: Short-run

The price and output determination under monopoly, and also the firm's equilibrium, are demonstrated in Fig. 13.1. The  $AR = D$  and  $MR$  curves show the revenue conditions, while  $SMC$  and  $SAC$  present the short-run cost conditions faced by the monopoly firm. Given the revenue and cost curves, the decision rule for selecting profit maximising output and price is the same as for a firm in the competitive industry, i.e., firm's  $MR = MC$  and slope of  $MC >$  the slope of  $MR$ . Therefore, the monopoly firm chooses a price-output combination for which  $MR = SMC$ . The  $MR$  and  $SMC$  curves intersect each other at point  $N$ . Thus, the profit maximising output for the firm is  $OQ$ , since at this output firm's  $MR = SMC$ . Given the demand curve  $AR = D$ , the output  $OQ$  can be sold per time unit at only one price, i.e.,  $PQ (= OP_1)$ . Thus, the determination of equilibrium output simultaneously determines the price for the monopoly firm. Once price and output are determined, the total profits are also simultaneously determined.

At output  $OQ$  and price  $PQ$ , the monopoly firm maximises its profit. Its per unit monopoly or super-normal profit (i.e.,  $AR - SAC$ ) is  $(PQ - MQ) = PM$ . Its total profit  $\square = OQ \times PM$ . Since  $OQ = P M$ ,  $\square = P M \times PM$ , as shown by the

2 2

shaded area. Since in the short-run cost and revenue conditions are not expected to change, the equilibrium of the monopoly firm will remain stable.

### Two Common Misconceptions

There are two common misconceptions about monopoly firm which must be cleared before we proceed. One of the misconceptions is that a monopoly firm necessarily makes super normal profits. There is however no guarantee that monopoly firm will always make profits in the short run. In fact, whether a monopoly makes profits or losses in the short run depends on its revenue and cost conditions. It is quite likely that its  $SAC$  lies above its  $AR$  as shown in Fig. 13.2. The monopoly firm then makes losses to the extent of  $PM \times OQ = P_2MPP_1$ . The firm may yet continue to produce and sell in the hope of making profits in the long-run. The monopoly firm, like a competitive firm, will however stick to the maximisation rules (i.e.,  $MR = MC$ ) in order to minimise its losses.

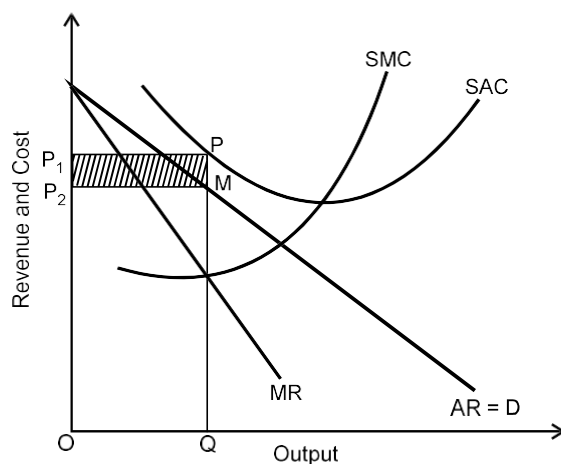


Fig. 13.2 Monopoly Equilibrium in the Short-run: Losses

Another common misconception about monopoly is that the demand curve faced by a monopoly firm is inelastic so that it can charge any price it likes. In fact, the demand curve faced by a monopolist is both firm's and industry's demand curve. And, most market demand curves are negatively sloped being highly elastic towards their upper end and highly inelastic towards their lower end. The equilibrium output of the monopolist that maximises his profits will always be within the elastic region of the demand curve, if his  $MC > 0$ .

### Monopoly Equilibrium in the Long Run

The long-run conditions faced by a monopolist are different from those faced by competitive firms in an important respect, i.e., the entry of new firms into the industry. While in a competitive industry, there is free entry of new firms to the industry, a monopoly firm is protected from competition by the *barriers to entry*.

Protected by barriers to entry, a monopoly firm gets an opportunity to expand the size of its plant with a view to maximising its long-run profits. The expansion of the plant-size may however be subject to such conditions as (a) size of the market; (b) expected economic profits; and (c) risk of inviting legal restrictions. Assuming none of these conditions limits the expansion of monopoly firm, **the general case** of monopoly equilibrium in the long-run is illustrated in Fig. 13.3. The *AR* and *MR* curves show the market demand and marginal revenue conditions faced by the monopoly firm. The *LAC* and *LMC* curves show the long-run cost conditions. The profit maximising monopoly firm equalises its *LMC* and *MR* at output  $OQ_2$ . The price at which the total output  $OQ_2$  can be sold is  $P_2Q_2$ . Thus, in the long run equilibrium, price is  $P_2Q_2$  and equilibrium output is  $OQ_2$ . This output-price combination maximises the monopolist's long-run profits. The total monopoly profit is shown by the area  $LP_2SM$ .

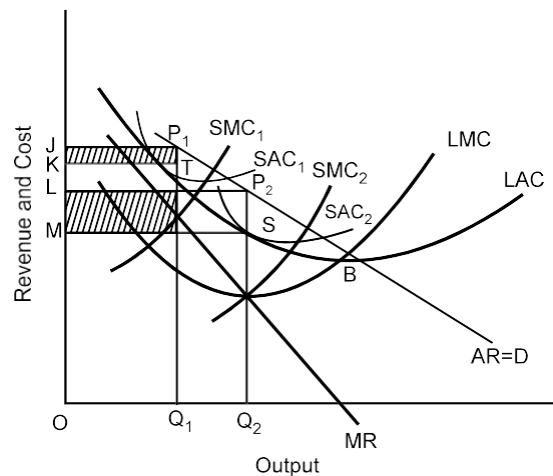
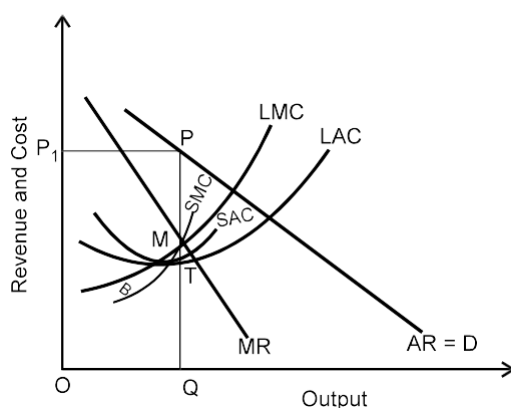


Fig. 13.3 Monopoly Equilibrium in the Long-run

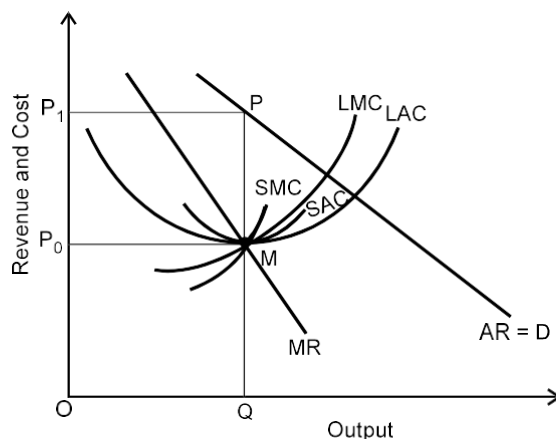
It may be noted at the end that if there are barriers to entry, the monopoly firm would not reach the optimal scale of production in the long-run, nor will make full use of its existing capacity. This case can be verified from Fig. 13.3. The

optimum size of the plant is given by point *B*, i.e., at the minimum *LAC*. But the monopoly firm settles at less than optimal output because optimum size of the plant will not yield the maximum profit.



**Fig. 13.4** Monopoly Equilibrium: Overutilization of Plant Size

Also, if the size of the market and the cost conditions permit, a profit maximising monopoly firm may even exceed the optimum size of the plant and overutilise its long-run capacity. Figure 13.4 depicts the more-than-optimal size of the plant and its overutilisation. The optimum size of the plant is given at point *B*, the point of intersection between *LAC* and *LMC*, whereas the monopoly firm chooses output at *M* where his profit is maximum. Alternatively, the monopoly firm may find its equilibrium just at the optimum size of the plant. This is possible only when the market-size is just large enough to permit optimisation and full utilisation of the plant size. This possibility has been illustrated in Fig. 13.5.



**Fig. 13.5** Monopoly Equilibrium at Optimal Size of the Plant

### 13.3 PRICE DISCRIMINATION BY MONOPOLY

The theory of pricing under monopoly, as discussed above, gives the impression that once a monopoly firm fixes up the price of its product, the same price is charged from all the consumers. This however may not be the case. A monopolist,

simply by virtue of its monopoly power, is capable of charging different prices from different consumers or groups of consumers. When the same (or slightly differentiated) product is sold at different prices to different consumers, it is called **price discrimination**. When a monopolist sells the same product at different prices to different buyers, the monopoly is called a **discriminatory monopoly**.

Consumers are discriminated in respect of prices on the basis of their incomes or purchasing powers, geographical location, age, sex, quantity they purchase, their association with the sellers, frequency of visits to the shop, the purpose of the use of the commodity or service, and on other grounds which the seller may find suitable.

A common example of consumers being discriminated on the basis of their incomes is found in medical and legal professions. Consulting physicians and lawyers (having excess capacity) charge different fees from different clients on the basis of their paying capacity. Delhi Vidyut Board charges different rates of tariffs for different grades and purpose of units of electricity consumed. Price discrimination on the basis of age is found in railways, roadways and airways: children below 15 years are charged only-half the adult-rates. Price discrimination on the basis of quantity purchased is very common. It is generally found that private businessmen charge lower price (or give discount) when bulk-purchase is made. In case of public utility services, however, lower rates are charged when commodity or service is consumed in smaller quantity, for example, lower rates on the first few calls by the telephone owners, and no surcharge on electricity upto certain level of consumption. The most common practice of price discrimination is found in cinema shows, musical concerts, game-shows, etc.

For the purpose of price discrimination, the product or service in question may be identical or slightly modified. For example, services of consulting physicians and lawyers are identical. The services of railways, roadways and entertainment shows may be slightly modified by providing more comfortable seats for the purpose of price discrimination. The modification in service may involve some additional cost. But price differentials are much more than is justified by cost differentials.

Although price-discrimination is most common practice under monopoly, it should not mean that this practice exists only under monopoly. Price discrimination is quite common also in other kinds of market structures, particularly where market imperfection exists. Most business firms discriminate between their customers on the basis of personal relationship, quantity purchased, duration of their association with the firm as buyers, and so on.

### *Necessary Conditions for Price Discrimination*

**First**, the market for different class of consumers must be separable so that buyers of low-price market are not in a position to resell the commodity in the high-price market for such reason as (i) geographical distance involving high cost of transportation, e.g., domestic versus foreign markets; (ii) exclusive use of the

commodity, e.g., doctor's services, entertainment shows, etc.; and (iii) lack of distribution channels, e.g., transfer of electricity and gas.

**Second**, if market is divided into submarkets, the elasticity of demand must be different in each submarket. The purpose of price-discrimination is to maximise the profit by exploiting the markets with different price elasticities. It is the difference in price-elasticities that provides opportunity for price discrimination. If price-elasticities of demand in different markets are the same, price discrimination would not be serve the objective of profit maximization.

**Third**, the seller must possess some monopoly over the supply of the product to be able to distinguish between different classes of consumers, and to charge different prices.

### Degrees of Price Discrimination

The degree of price discrimination refers to the extent to which a seller can divide the market and can take advantage of it in extracting the consumer's surplus. According to Pigou, there are three degrees of price-discrimination practiced by the monopolists: (i) *first degree* price discrimination; (ii) *second degree* price discrimination; and (iii) *third degree* price discrimination.

(a) **First Degree Price Discrimination.** The discriminatory pricing that attempts to take away the entire consumers' surplus is called *first degree discrimination*. First degree discrimination is possible only when a seller is in a position to know the price each buyer is willing to pay. That is, he knows his buyer's demand curve for his product. Under perfect price discrimination, the seller sets the price at the highest possible level at which all those who are willing to buy the product at that price buy at least one unit each. When the consumer's surplus of this section of consumers is exhausted, he gradually lowers down the prices so that the consumer's surplus of the users of the subsequent units can be extracted. This method of pricing is continued until the whole consumer's surplus available at the price where  $MR = MC$  is extracted. Also consider the case of services of exclusive use, e.g., medical services. A doctor who knows or can guess the paying capacity of his patients can charge the highest possible fee from presumably the richest patient and the lowest fee from the poorest one. The first degree of price discrimination is the limit of discriminatory pricing.

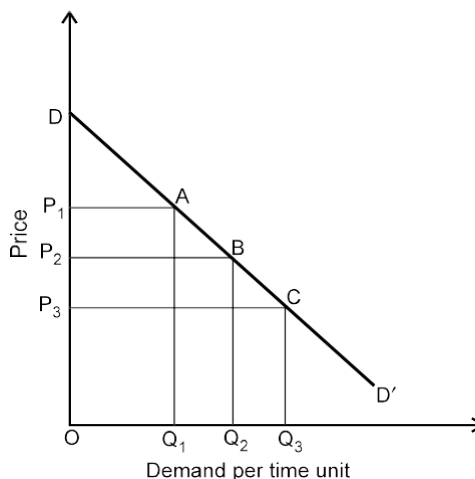
(b) **Second Degree Price Discrimination.** Under the second degree of discriminatory pricing, the firm charges different prices from different class of consumers—high, middle and low income consumers. The monopolist adopting the *second degree* price discrimination intends to siphon off only the major part of the consumer's surplus, rather than the entire of it. The second degree price discrimination is feasible where (i) the number of consumers is large and price rationing can be effective, as in case of utility services like telephones and natural gas; (ii) demand curves of all the consumers are identical; and (iii) a single rate is applicable for a large number of buyers. As shown in Fig. 13.6, a monopolist using



a second degree price discrimination charges price  $OP_1$  for the first few units,  $OQ_1$  and price  $OP_2$  for the next  $O_1Q_2$ , units, and price  $OP_3$  for the next additional

purchase of  $Q_2Q_3$  units. Thus, by adopting a block-pricing system, the monopolist maximises his total revenue ( $TR$ ) as

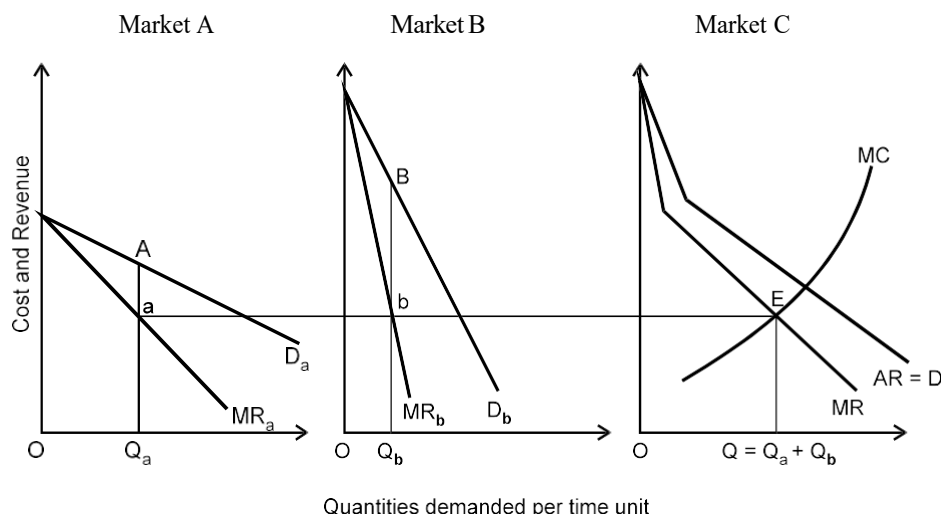
$$TR = (OQ_1 \cdot AQ_1) + (Q_1Q_2 \cdot BQ_2) + (Q_2Q_3 \cdot CQ_3)$$



**Fig. 13.6** Second Degree Price Discrimination

If a monopolist is restrained from price discrimination and is forced to choose any one of the three prices,  $OP_1$ ,  $OP_2$ , or  $OP_3$ , his total revenue will be much less.

(c) **Third Degree Price Discrimination.** When a profit maximising monopoly sets different prices in different markets having demand curves with different elasticities, it is using third degree price discrimination. When a monopolist is faced with two or more markets, completely separated from each other—each having a demand curve with different elasticity—a uniform price cannot be set for all the markets without losing profits. The monopolist is therefore required to allocate total output between the different markets so that profit can be maximised in all the markets. The profit in each market would be the maximum only when the  $MR = MC$  in each market. The monopolist therefore divides total output between the markets so that in all the markets  $MR = MC$ . Suppose that a monopolist has to sell goods in only two markets, A and B. The two markets are so separated that resale of commodity is not possible. The demand curve ( $D_a$ ) and marginal revenue curve ( $MR_a$ ) given in Fig. 13.7(a) represent the  $AR$  and  $MR$  curves in market A and curves  $D_b$  and  $MR_b$ , in Fig. 13.7(b) represent  $AR$  and  $MR$  curves, respectively, in market B. The horizontal summation of demand curves  $D_a$  and  $D_b$  gives the total demand curve for the two markets, as shown by the curve  $AR = D$ , and horizontal summation of  $MR_a$  and  $MR_b$  is given by the curve  $MR$  (Fig. 13.7). The firm's marginal cost is shown by the curve  $MC$  which intersects  $MR$  at point E. Thus, optimum level of output for the firm is determined at  $OQ$ . At this level of output,  $MR = MC$ . Since the whole of  $OQ$  cannot be profitably sold in any one market because of their limited size, the firm has to allocate the output between the two markets.



**Fig. 13.7** Third Degree Price Discrimination

The monopolist allocates output  $OQ$  between the two markets in such proportions that the necessary condition of profit maximisation is satisfied in both the markets. That is, in both the markets  $MC = MR$ . The profit maximising output for each market can be obtained by drawing a line from point  $E$  and parallel to X-axis, through  $MR_b$  and  $MR_a$ . The points of intersection on curves  $MR_a$  and  $MR_b$  at points  $a$  and  $b$ , respectively, determine the optimum share for each market. As shown in the Fig. 13.7, the monopoly firm maximises its revenue in market A by selling  $OQ_a$  units at price  $AQ_a$ , and by selling  $OQ_b$  units in market B at price  $BQ_b$ .

The firm's total equilibrium output  $OQ = OQ_a + OQ_b$ . Since at  $OQ_a$ ,  $MR_a = MC$  in market A, and at  $OQ_b$ ,  $MR_b = MC$  in market B,

$$MC = EQ = MR_a = MR_b$$

Thus, the equilibrium condition is satisfied in both market segments, and the monopoly firm adopting the *third degree* method of price discrimination maximises its profits.

The third degree method of price discrimination is most suitable where the total market is divided between the home and foreign markets. However, it need not be limited only to domestic and foreign markets. It may be suitably practised between any two or more markets separated from each other by any or more of such factors as geographical distance, transport barriers or cost of transportation, legal restrictions on the inter-regional or interstate transportation of commodities by individuals, etc.

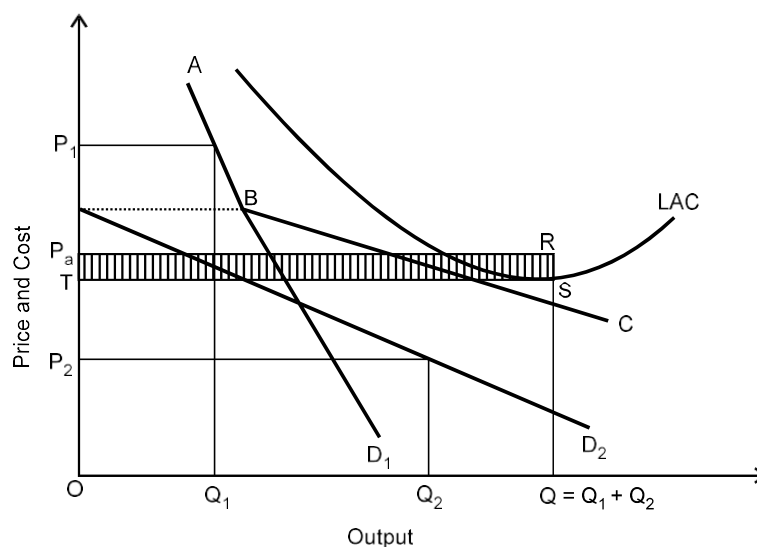
### Whether Price Discrimination is Justified

Price discrimination has been condemned as illegal and immoral. The objection is: why charge higher price from some and lower price from others while there is no extra advantage to those who pay higher price or why benefit some at the cost of some others? In the United Kingdom and the United States, railways were prohibited to charge discriminatory rates. Discriminatory pricing has also been

criticised as a destructive tool in the hands of a monopoly. For, in the past, large corporations had sought to use price discrimination to prevent the growth of competition. Besides, price discrimination may cause malallocation of resources and, hence, may be deterrent to social welfare. This is however not the case always. In some cases price discriminations is socially advantageous. In fact as Lipsey has observed, “whether an individual judges price discrimination to be good or bad is likely to depend upon the details of the case as well as upon his own personal value judgements.” He adds, “Certainly there is nothing in economic theory to suggest that price discrimination is always in some sense worse than non-discrimination under conditions of monopoly or oligopoly.”

Price discrimination is however considered to be desirable in certain specific cases on the following grounds:

**First**, is the case of goods and services which are essential for the society as a whole but their production is uneconomic in the sense that long-run average cost curve (*LAC*) lies much above the aggregated market demand curve as shown is Fig. 13.8. Such goods and services cannot be produced. But, production of such goods and services can be possible if price discrimination is permitted. Price discrimination thus becomes essential for the survival of the industry.



**Fig. 13.8** Price Discrimination for Industry's Survival

Suppose, for example, (i) that there are two markets I and II, (ii) that individual demand curves for the two markets, I and II, are given as  $D_1$  and  $D_2$ , (iii) market demand curve is given by  $ABC$ , and (iv) the long-run average cost curve is given by  $LAC$  (Fig. 13.8). Note that  $LAC$  lies throughout above the total demand curve  $ABC$ . Therefore, production is not possible if one price is to be charged. But, if price discrimination is adopted and prices are so charged in the two markets that the total revenue exceeds  $LAC$  at some level of output, then monopoly may profitably survive to the advantage of the society. Let us suppose that the monopolist sets price  $OP_1$  in the market I in which demand is less elastic

and  $OP_2$  in market II in which demand is highly elastic. He would sell  $OQ_1$  units at price  $OP_1$  in market I and  $OQ_2$  at price  $OP_2$  in market II. His total output would

then be at  $OQ = OQ_1 + OQ_2$ . His total revenue ( $TR$ ) would be

$$TR = (OP_1 \times OQ_1) + (OP_2 \times OQ_2)$$

and suppose

$$AR = (OP_1 \times OQ_1 + OP_2 \times OQ_2) / OQ = OP_a$$

At output  $OQ$ , the  $LAC$  is  $OT$ . Thus his total cost,

$$TC = OQ \times OT = OQST$$

and his total revenue,

$$TR = OQ \times OP_a = OQRP_a$$

Since  $OQRP_a > OQST$ , the monopoly firms not only covers its cost but also makes excess profit. Its total profit,

$$\square = OQRP_a - OQST = P_aRST$$

This kind of situation arises mostly in public utility services, like railways roadways, post and telegraph services, etc., in which high paying sector of the market subsidises the low paying sector. But, if low-paying sector is not subsidised, no production would be possible.

**Secondly**, discriminatory pricing can be adopted with justification where a uniform, single profitable price is likely to restrict the output and deprive many (particularly the people of lower income groups) of the essential goods or service. For example, if doctors in private practice, who often charge discriminatory price for their services, are asked to charge a uniform fee from all the patients, they would charge a fee high enough to maintain the level of their income. The high fee may deprive the poor of the doctor's service and may force them to opt for inferior or inadequate treatment. The result of the uniform high fee will be that the rich patients who can pay a still higher fee gain as they pay a price lower than what they could afford, and on the other hand, poor patients are deprived of proper medical service.

**Thirdly**, there may be cases where a section of consumers gains more than the people of other sections from the use of the same product. For example, from the use of electricity factory-owners gain more than the households. In such cases, uniform price would be unjustified from a normative point of view, provided the objectives is not to restrain the domestic consumption of electricity and spare it for productive purposes. There is, on the other hand, full justification for discriminatory pricing of electricity.

### Government Regulation of Monopoly Prices

The existence of monopolies in a market economy is criticised on the grounds that they restrict production and consumption, widen income and wealth disparities, exploit consumers and employees, cause distortions in allocation of resources, reduce the prospect of employment, and cause loss of social welfare. In most

countries, therefore, there is general apathy towards the monopolies. Consequently, governments in the market economies attempt to control and regulate monopolies to the advantage of the society. There are various measures—direct, indirect, legal and otherwise—to control and regulate the monopolies. However, we discuss below only the price regulation of natural monopolies.

Price regulation is a common feature in case of natural monopolies. When the size of the market is small relative to the optimum size of the firm, market size cannot support more than one firm of optimal size. The monopoly in such a market is a natural monopoly. The natural monopoly is thus protected by market size itself. The government may either nationalise such monopolies or regulate their prices so as to eliminate the excess profits. If the government intends to regulate the monopoly price, the question arises: what price should be fixed for the monopolist to charge? The two alternative prices that have been suggested are: **one** that allows some excess profit to the monopolist, and the **second** that allows only normal profit to the monopolist. An unregulated monopoly would produce  $OQ_1$  units, charge price  $OP_3$ , and make excess profit of  $MT = MQ_1 - TQ_1$  per unit. If monopoly price is regulated, one possible price is given at point  $P$  where  $LMC = AR$ , the price being  $OP_2 (= PQ_2)$ . Alternatively, price may be fixed at point  $C$  where  $AR = LAC$  and price =  $OP_1 (= CQ_3)$ . When  $OP_1$  is the price set for the monopolist, only a normal profit is allowed to the firm, but output is maximum possible under the given cost and revenue conditions. If price is fixed at  $OP_2$ , the monopolist gets some excess profit, but the output is less than that at price  $OP_1$ . In both the cases, however, the total output under regulated monopoly is much higher than that under unregulated monopoly. Which of the two alternative price ( $OP_1$  and  $OP_2$ ) is more appropriate is a matter of debate.

---

### 13.4 MONOPOLISTIC COMPETITION: MEANING AND FEATURES

---

The model of monopolistic competition developed by Edward H. Chamberlin presents a more realistic picture of the actual market structure and the nature of competition. In this unit, we will discuss briefly the nature of the market structure and monopolistic competition among the firms.

**Monopolistic Competition.** *Monopolistic Competition* is a market structure in which a *large number* of sellers sell *differentiated* products which are close, but not perfect, substitutes for one another. Monopolistic competition combines the characteristics of perfect competition and monopoly.

The assumptions of the monopolistic competition are the same as those of pure competition, with an exception of homogeneity of products. While pure competition model assumes that products are *homogeneous* in every possible dimension, monopolistic competition model assumes that products are *differentiated*. The product of each firm is so differentiated from those of other

firms that consumers are able to distinguish the product of a firm from those of others. For example, consumers know for sure the difference between different brands of mobile phones—Nokia, Sony, Samsung, Reliance, etc. Since each firm produces a product distinguishable from that of other firms, each firm holds a monopoly power over its own products.

Although products are differentiated, they remain a close substitute for one another. This creates condition for competition among the firms which are monopolists in their own rights. This kind of competition is the genesis of monopolistic competition.

### 13.4.1 Foundations of Monopolistic Power

#### (i) Assumptions

Chamberlin's model of monopolistic competition is based on the following assumptions:

1. There is a large number of buyers and sellers in the market.
2. Each seller sells a product differentiated from that of others.
3. The differentiated products are close, not perfect, substitute for one another.
4. There is free entry and free exit of firms.
5. The firms seek to maximise their profits in both short and long runs.
6. Technology and factor prices are given and the firms are aware of revenue and cost curves.

#### (ii) Product Differentiation and the Demand Curve

Chamberlin has defined *product differentiation* in the following words: "A general class of product is differentiated if any significant basis exists for distinguishing the goods (or services) of one seller from those of others. Such a basis may be real or fancied, so long as it is of any importance whatever to buyers, and leads to a preference for one variety of the product over another. Differentiation may be based upon certain characteristics of the product itself, such as exclusive patented features, trade marks, trade names, peculiarities of the package or container, if any, or singularity in quality, design, colour or style. It may also exist with respect to the conditions surrounding its sales. In retail trade, these conditions include such factors as the convenience of the seller's location, the general tone or character of his establishment, his way of doing business, his reputation for fair dealing, courtesy, efficiency, and all the personal links which attach his customers either to himself or to those employed by him." So far as these and other tangible and intangible factors create consumers' preference for one product over the others, the products are virtually differentiated.

Thus, product differentiation is primarily intended to make consumers distinguish the product of one producer from that of the other producers in the industry. When the consumers are able to distinguish one product from the others,

they may develop a preference or brand loyalty for one product over the others. Once preference for a product is created, it alters the course of demand curve for the product. In ultimate analysis, product differentiation leads to a change in demand curve for the product from a horizontal demand line (as under pure competition) to a downward sloping demand curve. The downward sloping demand curve enables the sellers to exercise some discretion in determining the price of his product.

### (iii) Cost Curves and Selling Cost

In his model of monopolistic competition, Chamberlin has assumed the traditional U-shaped cost curves— $AC$ ,  $AVC$  and  $MC$ . In addition, he has introduced a new cost, i.e., *selling cost*. “Selling costs are defined as costs incurred in order to alter the position or the shape of the demand curve for a product.” Selling costs include all the expenses that are intended to promote the sales, including cost of advertisement, salesmen’s salaries, expenses of sales department, margins granted to dealers—wholesalers and retailers—and on window displays and demonstration of new goods. Selling costs affect demand curve in two ways.

**First**, selling costs make the demand curve for the product shift upward by informing consumers about the availability of the product and by increasing consumer’s preference for the product.

**Second**, selling cost makes the demand curve less elastic by strengthening the consumers’ preference for the product.

Chamberlin assumes average selling cost to be U-shaped, that is, selling cost per unit of sales initially decreases but eventually increases. Thus, the average selling-cost curve has a shape similar to the  $AC$  curve.

### (iv) The Concept of Industry and Product Groups

Under monopolistic competition, products are so differentiated that each product is distinguishable from others, and each firm is, in a sense, an industry in itself, exactly as a monopoly firm is an industry in itself. The heterogeneity of the products, therefore, causes a problem in analytical treatment of the industry. It may be recalled that, in case of homogeneous products, demand curve for an industry can be obtained by adding individual demand curve. But in case of heterogeneous products, the demand for individual products cannot be added to obtain market demand and supply curves.

For this reason, Chamberlin attempted to redefine the industry for his analytical purpose. He defined the monopolistically competitive industry as a ‘group’ of firms producing a ‘closely related’ commodity, referred to as *product group*. The product of the ‘group’ must be *close, technological and economic substitutes*. The two products are *technological substitutes* for each other if they technically satisfy the same want, e.g., personal computers, soaps, toothpastes,

automobiles, TV sets, etc. The two products are considered as *economic substitutes* for each other if they satisfy the same want and have *more or less* the same price. For example, all brands of TV sets are economic substitutes for one another. But flat TV sets are not economic substitutes for ordinary ones since their prices are widely different, though they remain technological substitutes. Operationally, the product group may be defined as the group of firms whose products have between themselves high *price* and *cross elasticities*. This definition, although theoretical plausible, involves the problems of measuring cross-elasticities and of determining the degree of cross-elasticities that can make a commodity admissible to the group. Determining the product group would therefore involve subjective judgement.

---

### 13.5 PRICE AND OUTPUT DETERMINATION UNDER MONOPOLISTIC COMPETITION

---

Chamberlin's theory of price and output determination under monopolistic competition is basically the same as that under monopoly with a difference, of course. While under monopoly demand and cost curves are both assumed to be given, under monopolistic competition, firms are assumed to indulge in competition to change the slope of the demand curve or to make it shift rightward, given the cost curves. They seek to make these changes in the demand curve by any or all of the following measures:

- (i) change in the price of the product;
- (ii) change in the nature of the product; and
- (iii) change in the advertisement outlays.

As to price change, since a monopolistically competitive firm faces an elastic demand curve similar to one faced by a monopoly firm, it has the option to raise the price and sell less or to lower the price and sell more. But, it fixes a price that maximises its profits. As to change in product, the sales can be promoted by change in the quality of product through technical changes, introduction of a new design, use of better material, use of new package or containers, prompt and courteous services, credit facilities, etc. Also, the firm may influence its volume of sales by increasing advertisement expenditure so that more consumers are attracted to the product. Increase in advertisement expenditure also increases the selling price. The firm is therefore required to so adjust its price and output that its profits is maximum.

While adjustment between price and output for profit maximisation is a short-run phenomenon, changes in the quality of the product and advertisement expenses are long-run phenomena. We will therefore explain Chamberlin's theory of price and output determination under monopolistic competition under both short-run and long-run conditions.



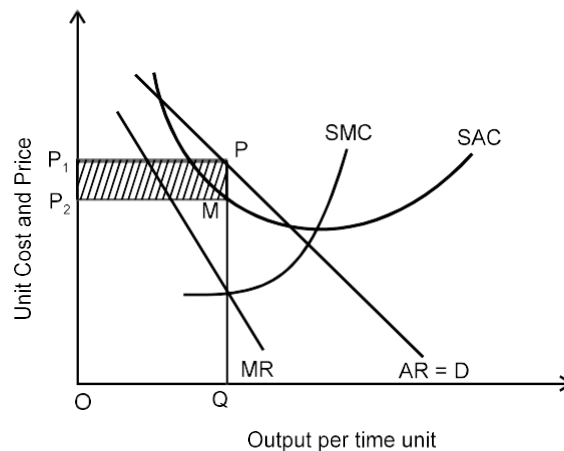
### Firm's Short-run Equilibrium

While monopolistic competition is characteristically closer to perfect competition, it is closer to monopoly in regard to pricing and output determination. Like a monopolist, a monopolistic competitor faces a downward sloping demand curve having a smaller slope. This demand curve is the product of (i) strong preference of a section of consumers for a particular product; and (ii) the quasi-monopoly of the seller over the supply. The strong preference or loyalty of the consumers gives the seller an opportunity to raise the price and yet retain some customers. And, since each product is a close substitute for another, they attract the consumers of other products by lowering down their prices.

As mentioned above, short-term analysis of pricing and output determination under monopolistic competition is similar to price and output determination under monopoly. The short-term equilibrium analysis is primarily the adjustment of price and output to the given cost and revenue conditions. The short-run price and output adjustment is illustrated in Fig. 13.9. The  $AR = D$  and  $MR$  curves show the revenue conditions and  $SAC$  and  $SMC$  curves show the cost conditions faced by the firm in the short-run.

As shown in Fig. 13.9, the necessary condition of profit maximisation, i.e.,  $MR$  must be equal to  $MC$ , is fulfilled at output  $OQ$ . This output can be sold at price  $PQ$ , so the price is also determined. At this output and price, the firm earns a maximum economic profit, shown by the rectangle  $P_1PMP_2$ .

The economic profit per unit ( $PM$ ) exists in the short-run because new firms cannot enter the industry. But the rate of profit would not be the same for all the firms under monopolistic competition because of difference in the elasticity of demand. For the same reason, product price will be different for the different products, though price differentials will be only marginal.



**Fig. 13.9** Short-run Equilibrium Under Monopolistic Competition

In the short run, the firms may attempt to maximise their profits by changing the nature of the product and by increasing advertisement expenditure. But, since

there are many close substitutes, neither of the strategies would be of much avail in the short run. If the firms do adopt these strategies, they would do so only to maximise their profits.

### *Long-run Equilibrium under Monopolistic Competition*

The conditions faced by a firm of the 'product-group' under monopolistic competitions are different from those in the short run. In the long run, a firm can expand its plant-size and if there are no barriers to entry, new firms will enter the product group, especially when existing firms are making pure profits. Besides, the firms get the opportunity to change the nature and position of the demand curve for their product by (i) changing the quality of the product; and (ii) incurring a large amount of advertisement expenditure. Since conditions change in the long run, the firms can maneuver their price and output in order to maximise their profits in the long run too.

### Basic Assumptions

Chamberlin has made the following explicit and implicit assumptions in order to develop his theory of monopolistic competition under long-run conditions.

**First**, the basic assumption is that there is a large number of firms producing differentiated products which are close substitutes for one another.

**Second**, the number of firms in the product group is so large that each firm expects its maneuvering of prices and output to go unnoticed by the rival firms.

**Third**, one of the heroic assumptions of Chamberlin is that both demand and cost curves for all the products are uniform throughout the group. That is, all firms have identical revenue and cost curves.

**Fourth**, his second heroic assumption is that consumer's preferences are evenly distributed among the different varieties, and that differences between products are not such as to give rise to differences in cost.

Under these assumptions, Chamberlin develops **three models** of equilibrium:

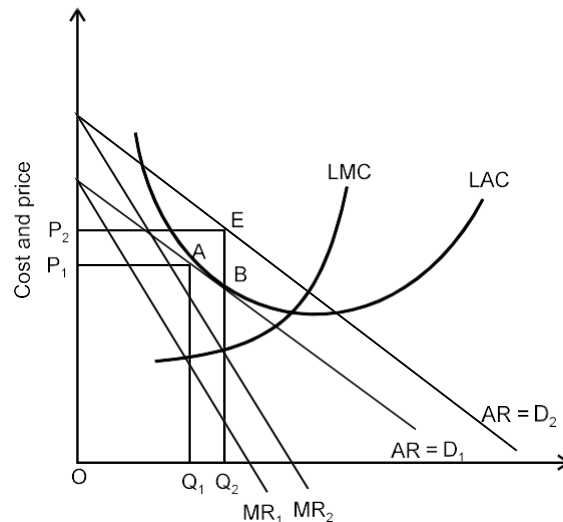
- (i) to analyse equilibrium with free entry of new firms to the industry with non-price competition;
- (ii) to analyse equilibrium under price competition, assuming no entry or exit; and
- (iii) to present a combined analysis of the first and the second models.

We will explain briefly the three models in the subsequent subsections.

#### *(i) Long-run Equilibrium with Free Entry of New Firms*

The long-run equilibrium of firms under the assumption of free entry of new firms is illustrated in Fig. 13.10. The *LAC* and *LMC* are the cost curves faced by the firms. The initial *AR* and *MR* curves (i.e., prior to the entry of new firms) are given by  $AR = D_2$  and  $MR_2$ . Given the cost and revenue curves, the firms will be in

short-run equilibrium at point  $E$  where price is  $OP_2$  and output  $OQ_2$ . The firms are making super-normal profits to the extent of  $EB$  per unit of output.



**Fig. 13.10** Introduction of Second Demand Curve

The existence of super-normal profits attracts new firms to the product group. With the entry of new firms, the sale of each firm in the group decreases. Consequently, the demand curve (or  $AR$  curve) for the firm shifts leftward and so does the  $MR$  curve. This forces the firms to adjust their price and output to the new  $AR$  and  $MR$  conditions, cost conditions remaining the same. If still there exists super-normal profit more new firms join the product group, and existing firms will be forced to readjust their price and output to another equilibrium position. This process will continue until the demand curve becomes tangent to the  $LAC$ , and all the firms earn only normal profits. The final equilibrium position of the firms in the long-run under monopolistic competition is shown at point  $A$ , where price is  $OP_1$  and output  $OQ_1$ . At this price and output, all firms make only normal profit. Therefore, there is no incentive for the new firms to enter the industry. The equilibrium will therefore be stable at point  $A$ .

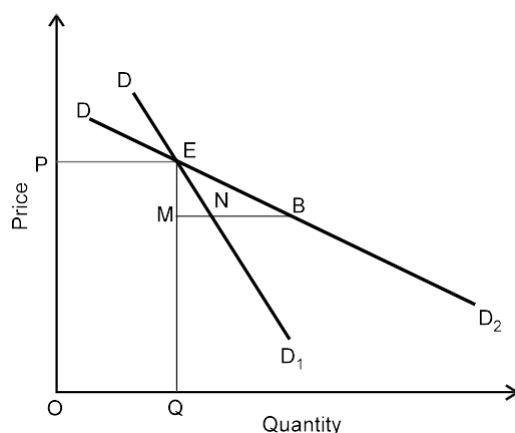
*(ii) Long-run Equilibrium with Price Competition*

In his model of long-run equilibrium with price competition Chamberlin assumes that the number of existing firms in the product group is optimal. That is, the number of firms is compatible with long-run equilibrium of the industry. There is no entry or exit of the firms. In this case, the equilibrium analysis has been accomplished in two stages. In the first stage, a second demand curve is introduced to the model to incorporate the effects of competitive change in prices. In the second stage, the long-run equilibrium under the condition of optimal number of firms has been analysed.

To commence the analysis of stage one, let us assume that demand curve for the product of a firm is given by  $DD_1$ , as shown in Fig. 13.11, and the firm is in

equilibrium at point  $E$ , with price  $OP$  and output  $OQ$ . Now suppose, that the firm contemplates a price reduction and assumes that the rival firms will not react to its price cut. By reducing its price, the firm expects to expand its sales on account of two factors: *first*, the demand for its product is elastic, and *second*, the consumers of other products will switch over to the product of this firm if other firms do not simultaneously reduce their prices. Thus, the firm can expect a substantial increase in its total sales.

For example, if the firm *reduces* its price by  $EM$ , the demand for its product increases by  $MB$ . Of this,  $MN$  is due to the elasticity of its original demand  $DD_1$  and  $NB$  is due to the fact that some customers of other firms switch over to the product of this firm. On the other hand, if the firm *raises* its price, and the rival firms do not raise their prices it loses its market partly because its demand curve is elastic and partly because its customers switch over to other products which become automatically cheaper. Thus, another demand curve  $DD_2$  emerges. This is called as the *second demand curve*. But, if all the firms change their prices simultaneously this advantage to an individual firm is lost and all the firms return to the original demand curve  $DD_1$ .



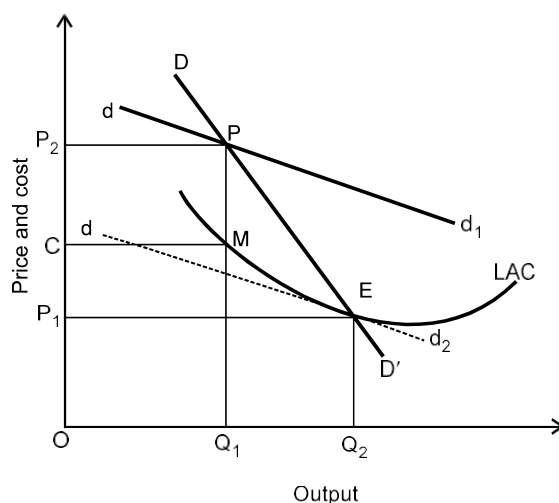
**Fig. 13.11** Introduction of Second Demand Curve

Having introduced the second demand curve, we can proceed to present Chamberlin's long-run equilibrium with price competition, i.e., *second stage* of his analysis. The long-run equilibrium with price competition is presented in Fig. 13.12. The curves  $DD_1$  and  $dd_1$  are the two demand curves and  $LAC$  is the long-

run average cost curve of a 'typical' firm of the group. Let the initial short-run equilibrium of the firms of the group be at point  $P$  with price  $OP_2$  and output  $OQ_1$ . At this price and output, the firm makes the total abnormal profit represented by the area  $P_2PMC$ .

Although in equilibrium, each firm regards  $dd_1$  to be its demand curve and believes that it can increase its profits by reducing the price and expanding the output. The basis of this belief is the elasticity of their assumed demand curve  $dd_1$ . In an attempt to increase profits, each firm reduces its price expecting to move

along the demand curve  $dd_1$ . However, instead of moving along  $dd_1$ , the firms move along the market demand curve  $DD_1$ , because all of them reduce their prices simultaneously. However, according to Chamberlin, the firms do not learn from their past experience and each firm sticks to its own belief that the demand curve ( $dd$ ) for its product is more elastic than the market demand curve ( $DD$ ). Therefore, the firms go on reducing their prices independently and their assumed demand curve ( $dd_1$ ) shifts downward. This process continues until the downward shift in  $dd_1$  makes it tangent to the  $LAC$  curve, as shown by  $dd_2$ . A further reduction in price will make firms incur loss. Therefore, reduction in price below  $OP_1$  is not desirable. Thus, the long-run equilibrium of firms takes place at  $E$ , where each firm produces  $OQ_2$  and fixes its price at  $OP_1$ .



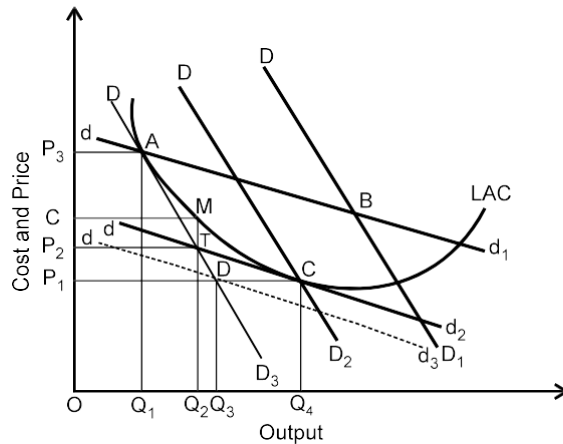
**Fig. 13.12** Long-run Equilibrium Price Competition

(iii) Long-run Equilibrium with Free Entry and Price Competition

We have explained above the equilibrium of the firm with free entry and with price competition separately. In this section, we bring together the two equilibrium analyses, and explain Chamberlin's third and the final model of firms' equilibrium under monopolistic competition. According to Chamberlin, the ultimate equilibrium under monopolistic competition can be achieved through an integrated analysis of effects of free entry and price adjustments. The integrated analysis of equilibrium is presented in Fig. 13.13.

In Fig. 13.13,  $DD_1$  is assumed to represent the initial demand curve and  $LAC$  to represent the long-run average cost curve. Let us suppose that the firms are initially in equilibrium at point  $B$ , and they make abnormal profits to the extent of vertical distance between  $DD_1$  and the  $LAC$ . Since entry to the 'product group' is free, new firms are attracted by the industry. When new firms with slightly differentiated products enter the 'product group', the market share of each existing firm is reduced. Hence, their demand curve  $DD_1$ , shifts leftward. Given the  $LAC$ , the leftward shift in the demand curve will continue until it becomes tangent to

LAC, as shown by  $DD_3$  in Fig. 13.13, because till this point of tangency is reached, firms make abnormal profits and new firms continue to enter the 'product group'.



**Fig. 13.13** Long-run Equilibrium with Free Entry and Price Competition

Thus, it might seem that the long-run equilibrium is attained at point A with output  $QQ_1$  and price  $OP_3$ . This is however not the case. This is only half of the story, i.e., the influence of free entry. Let us now consider the competitive maneuvering of price and its role in determining to the long-run equilibrium.

Once the firms reach point A, each firm thinks that its demand curve is  $dd_1$ , not  $DD_3$ . Each firm believes that it can increase its profit by reducing the price and thereby increasing the sales. Therefore, in their attempt to increase their profits, they reduce prices of their product simultaneously because each firm has the same incentive to do so. As a result, their subjective demand curve ( $dd_1$ ) slides downward to  $dd_2$  and they incur losses. For example, if price is reduced to  $OP_2$ , the total loss equals the rectangle  $CMTP_2$ . It might seem that the firms could eliminate their loss by reducing the price to  $OP_1$ . But when all the firms reduce their price to  $OP_1$ —and they will do so under the assumption—their subjective demand curve  $dd_2$  slides further down to  $dd_3$ , the dotted line that lies below the  $LAC$ . As a result, the firms make increasing losses. A temporary equilibrium will be attained at point D with output  $OQ_3$ , where all firms incur heavy losses. Consequently, the firms which are unable to sustain losses will eventually leave the industry. The remaining firms find their share in the market increasing. Therefore,  $DD_3$  and  $dd_3$  move to the right until  $DD_3$  shifts to  $DD_2$  and  $dd_3$  shifts upward to the position of  $dd_2$ . Note that  $dd_2$  intersects  $DD_2$  at point C the point where  $dd_2$  is tangent to  $LAC$ . Thus, the long-run equilibrium is attained at point C where it is stable because all firms earn only normal profits and, therefore, there is no entry or exit of the firms.

### Critical Appraisal of Chamberlin's Theory of Monopolistic Competition

Chamberlin's theory of monopolistic competition has been criticised on both theoretical and empirical grounds. Let us first look into its theoretical or methodological weaknesses.

**First,** Chamberlin assumes that monopolistic competitors act independently and their price maneuvering goes unnoticed by the rival firms. This assumption has been questioned on the ground that the sales of other firms are bound to be affected by the decisions of rival firms since their products are close substitutes for one another and, therefore, they are bound to react.

**Second,** Chamberlin's model implicitly assumes that monopolistically competitive firms do not learn from their past experience. They continue to commit the mistake of reducing their prices even if successive price reductions lead to increase in their losses. Such an assumption can hardly be accepted.

**Third,** Chamberlin's concept of industry as a 'product group' is ambiguous. It is also incompatible with product differentiation. In fact, each firm is an industry by virtue of their specialised and unique product.

**Fourth,** his 'heroic assumptions' of identical cost and revenue curves are questionable. Since each firm is an industry in itself, there is a greater possibility of variations in the costs and revenue conditions of the various firms.

**Finally,** Chamberlin's assumption of free entry is also considered to be incompatible with product differentiation. Even if there are no legal barriers, product differentiation and brand loyalties are in themselves barriers to entry.

**Empirical validity.** So far as empirical validity of Chamberlin's concept of monopolistic competition is concerned, it is claimed that it is difficult to find any example in the real world to which his model of monopolistic competition is relevant. Most markets frequently available in the real world may be classified under perfect competition, oligopoly or monopoly. It is therefore alleged that Chamberlin's model of monopolistic competition analyses an unrealistic market. Some economists, e.g., Cohen and Cyert, hold the position that the model of monopolistic competition is not a useful addition to economic theory because it does not describe any market in the real world.

Despite above criticism, Chamberlin's contribution to the theory of price cannot be denied. Chamberlin is first to introduce the concept of *differentiated product* and *selling costs* as a decision variable and to offer a systematic analysis of these factors. Another important contribution of Chamberlin is the introduction of the concept of demand curve based on market share as tool of analysing behaviour of firms, which later became the basis of the *kinked-demand* curve analysis.

---

## **13.6 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS**

---

1. The discriminatory pricing that attempts to take away the entire consumers' surplus is called first degree discrimination.
2. Monopolistic competition is a market structure in which a large number of sellers sell differentiated products which are close, but not perfect, substitutes for one another.
3. The assumptions of the monopolistic competition are the same as those of pure competition, with an exception of homogeneity of products.

---

## **13.7 SUMMARY**

---

- The term pure monopoly signifies an absolute power to produce and sell a product which has no close substitute. In other words, a monopoly market is one in which there is only one seller of a product having no close substitute.
- The emergence and survival of a monopoly is attributed to the factors which prevent the entry of other firm into the industry. The barriers to entry are therefore the sources of monopoly power.
- The nature of revenue curves under monopoly depends on the nature of demand curve a monopoly firm faces. We have noted earlier that in a perfectly competitive market, firms face a horizontal, straight-line demand curve. It signifies that an individual firm of an industry can sell any quantity at the prevailing price.
- In the short-run, cost conditions faced by a monopoly firm are, for all practical purposes, identical to those faced by a firm under perfect competitions, particularly when a monopoly firm is a competitive buyer in the input market.
- The long-run conditions faced by a monopolist are different from those faced by competitive firms in an important respect, i.e., the entry of new firms into the industry. While in a competitive industry, there is free entry of new firms to the industry, a monopoly firm is protected from competition by the barriers to entry.
- Protected by barriers to entry, a monopoly firm gets an opportunity to expand the size of its plant with a view to maximising its long-run profits. The expansion of the plant-size may however be subject to such conditions as (a) size of the market; (b) expected economic profits; and (c) risk of inviting legal restrictions.
- The theory of pricing under monopoly, as discussed above, gives the impression that once a monopoly firm fixes up the price of its product, the same price is charged from all the consumers.



- The discriminatory pricing that attempts to take away the entire consumers' surplus is called first degree discrimination. First degree discrimination is possible only when a seller is in a position to know the price each buyer is willing to pay.
- Price regulation is a common feature in case of natural monopolies. When the size of the market is small relative to the optimum size of the firm, market size cannot support more than one firm of optimal size.
- Monopolistic competition is a market structure in which a large number of sellers sell differentiated products which are close, but not perfect, substitutes for one another. Monopolistic competition combines the characteristics of perfect competition and monopoly.
- The assumptions of the monopolistic competition are the same as those of pure competition, with an exception of homogeneity of products.

---

### 13.8 KEY WORDS

---

- **Barrier:** It refers to a fence or other obstacle that prevents movement or access.
- **Franchise:** It is an authorization granted by a government or company to an individual or group enabling them to carry out specified commercial activities, for example acting as an agent for a company's products.
- **Patent:** It is a government authority or license conferring a right or title for a set period, especially the sole right to exclude others from making, using, or selling an invention.

---

### 13.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

#### *Short-Answer Questions*

1. State the misconceptions about a monopoly firm.
2. Briefly discuss the degrees of price discrimination.
3. Write a short note on government regulation of monopoly prices.

#### *Long-Answer Questions*

1. Describe the major sources of barriers to enter a monopolised market.
2. Discuss the concept of price discrimination under monopoly. What are the necessary conditions for price discrimination?

3. What do you understand by monopolistic competition? What are the foundations of monopolistic power?
4. Discuss price and output determination under monopolistic competition.

---

### **13.10 FURTHER READINGS**

---

Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.

Keat Paul G. and K.Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.

Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition. Singapore: Pearson Education, Inc.

Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Publishers.

---

# OLIGOPOLY

---

## *Structure*

- 14.0 Introduction
- 14.1 Objectives
- 14.2 Oligopoly: Meaning and Features
  - 14.2.1 Characteristics of Oligopoly
  - 14.2.2 Price Determination under Non-collusive Models of Oligopoly
  - 14.2.3 Price Determination under Collusive Models of Oligopoly
- 14.3 Duopoly
- 14.4 Answers to Check Your Progress Questions
- 14.5 Summary
- 14.6 Key Words
- 14.7 Self Assessment Questions and Exercises
- 14.8 Further Readings

---

## 14.0 INTRODUCTION

---

In the previous units, you studied about monopoly and monopolistic competition.

A firm's ability to decide prices of a commodity in a market gives rise to monopoly power. This ability is dictated by the shape of the demand curve facing that firm. If the firm faces a downward sloping demand curve, it is no longer a price taker but rather a price setter. In our perfect competition model, we assume there exist multiple participants, and because there are so many participants, the slope of the demand curve each firm sees is but a flat line. These firms are price takers. There is a medium between monopoly and perfect competition in which only a few firms exist in a market. None of these firms faces the entire demand curve in the way a monopolist would, but each does have some power to set prices.

A small collection of firms who dominate a market is called an oligopoly. A duopoly is a special case of an oligopoly, in which only two firms exist. In this unit, you will study the concepts and models of oligopoly and duopoly in detail.

---

## 14.1 OBJECTIVES

---

After going through this unit, you will be able to:

- Describe the concepts and features of oligopoly and duopoly
- Assess price determination under oligopoly and duopoly
- Explain various models under oligopoly and duopoly

---

## 14.2 OLIGOPOLY: MEANING AND FEATURES

---

Oligopoly is a form of market structure in which a few sellers sell differentiated or homogeneous products. 'How few are the sellers' is not easy to define numerically in the oligopolistic market structure. The economists are not specified about a definite number of sellers for the market to be oligopolistic in its form. It may be two, three, four, five or more. In fact, the number of sellers depends on the size of the market. Given the size of the market, if number of sellers is such that each seller has command over a sizeable proportion of the total market supply then there exists oligopoly in the market.

The products traded by the oligopolists may be differentiated or homogeneous. Accordingly, the market may be characterised by **heterogeneous oligopoly** or **homogeneous (or pure) oligopoly**. In automobile industry, Maruti Zen, Hyundai's Santro, Daewoo's Matis, Fiat's Palio and Tata's Indica, etc., are the outstanding examples of differentiated oligopoly. Similarly, cooking gas of Indane and of Burshane are the examples of homogeneous oligopoly. Differentiated oligopolies include automobiles, cigarettes, refrigerators, TV industries. Pure oligopoly includes such industries as cooking gas, cement, baby food, vegetable oils, cable wires, dry batteries, etc. Other examples of oligopolistic industries are aluminium, paints, tractors, steel, tyres and tubes, etc.

### 14.2.1 Characteristics of Oligopoly

The basic characteristics of oligopolistic market structure are following:

1. **Intensive Competition.** The characteristic fewness of their number brings oligopolist in intensive competition with one another. Let us compare oligopoly with other markets structures. Under perfect competition, competition is non-existent because the number of sellers is so large that no seller is strong enough to make any impact on market conditions. Under monopoly, there is a single seller and, therefore there is absolutely no competition. Under monopolistic competition, number of sellers is so large that degree of competition is considerably reduced. But, under oligopoly, the number of sellers is so small that any move by one seller immediately affects the rival sellers. As a result, each firm keeps a close watch on the activities of the rival firms and prepares itself with a number of aggressive and defensive marketing strategies. To an oligopolist, business is a 'life' of constant struggle as market conditions necessitate making moves and counter-moves. This kind of competition is not found in other kinds of market. ***Oligopoly is the highest form of competition.***
2. **Interdependence of Business Decisions.** The nature and degree of competition among the oligopolists makes them interdependent in respect of decision-making. The reason for interdependence between the oligopolists is that a major policy change made by one of the firms affects

the rival firms seriously and immediately, and forces them to make counter-moves to protect their interest. Therefore, each oligopolist, while making a change in his price, advertisement, product characteristics, etc. takes it for granted that his actions will cause reaction by the rival firms. Thus, interdependence is the source of action and reaction, moves and counter-moves by the competing firms. An illuminating example of strategic maneuvering by the oligopoly firm has been given by Robert A. Meyer. To quote the example, one of the US automobile companies announces in September an increase of \$ 180 in the list price of its new car model. Following it, a few days later, a second company announces an increase of only \$80 and a third announces increase of \$91. The first company makes a counter-move: it suddenly reduces the increase in list price to \$71 from \$180 announced earlier. One can now expect that other firms will follow the first in price-cutting. Obviously, there is a good deal of uncertainty in the behaviour of firms.

3. **Barrier to Entry.** An oligopolistic market structure is also characterised, in the long run, by strong barriers to entry of new firms to the industry. If entry is free, new firms attracted by the super-normal profits, if it exists, enter the industry and the market eventually becomes competitive. Usually barriers to entry do exist in an oligopolistic market. Some common barriers to entry are economies of scale, absolute cost advantage to old firms, price-cutting, control over important inputs, patent rights and licencing, preventive price and existence of excess capacity. Such factors prevent the entry of new firms and preserve the oligopoly.

#### *Oligopoly Models: An Overview*

The uncertainty in respect of behaviour pattern of a oligopoly firms arising out of their unpredictable action and reaction makes systematic analysis of oligopoly extremely difficult. Under the circumstances, a wide variety of behaviour pattern has been observed: they may come in collusion with each other or 'may try to fight each other to the death'. The agreement may last or may breakdown soon. Indeterminateness of price and output therefore becomes the basic feature of oligopolistic markets. In accordance with the variety of behaviours, economists have developed a variety of analytical models based on different behavioural assumptions. Among notable models are Cournot's Duopoly model (1838), Bertrand's model (1883), Edgeworth's model (1897), Stackelberg's leadership model (1930), Hotelling's model (1930s), Chamberlin's model (1933), Sweezy's kinked-demand curve model (1939), Neumann and Morgenstern's game theory model (1944), and Baumol's sales maximisation model. None of these models, however provide a universally acceptable analysis of oligopoly, though these models do provide insight into the behavioural pattern of oligopolists. Moreover, these models are studied for their pedagogic importance.

The analytical models of oligopoly, suggested by the economists, may be classified under two broad categories:

- (i) duopoly models, and
- (ii) general oligopoly models.

The oligopoly models may be further sub-classified as (a) Non-collusive models, and (b) Collusive models.

#### *14.2.2 Price Determination under Non-collusive Models of Oligopoly*

In this section, we move to explain some important theories of price and output determination under oligopoly in which the number of sellers is more than two but only a few. As mentioned above, there are two kinds of oligopoly models: (i) non-collusive models and (ii) collusive models. We will first discuss the non-collusive models and then the collusive models. The non-collusive models of oligopoly explain the price and output determination in a market structure in which oligopolists recognise their interdependence. Chamberlin's non-collusive model of oligopoly, i.e., 'small group' model, is considered a major contribution to the theory of oligopoly. Another famous model of this category is Sweezy's kinked demand curve model. We begin our discussion oligopoly models with Chamberlin's non-collusive model.

#### *Non-Collusive Models of Oligopoly*

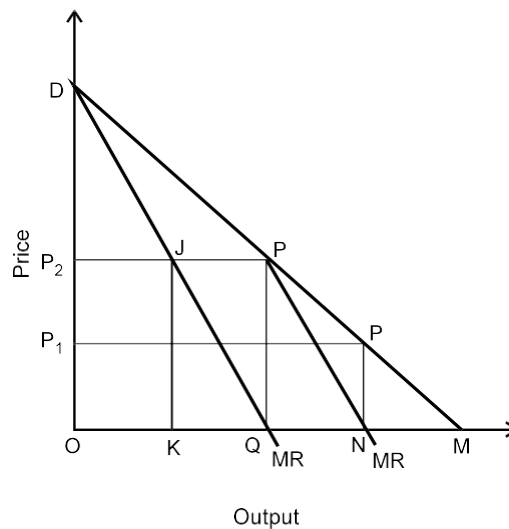
##### **(i) Chamberlin's Model of Oligopoly: The 'Small Group' Model**

The *classical* models of duopoly assumed independent action by the rival firms in their attempt to maximise their profits. Chamberlin rejected the assumption of independent action by the competing firms. He developed his own model of oligopoly assuming *interdependence* between the competitors. He argued that firms do not act independently. They do recognise their mutual interdependence. Firms are not as 'stupid' as assumed in the models of Cournot, Edgeworth and Bertrand. In his own words, "When a move by one seller evidently forces the other to make a counter-move, he is very stupidly refusing to look further than his nose if he proceeds on the assumption that it will not." Chamberlin suggests that each seller seeking to maximise his profit reflects well and looks into the consequences of his move. The total consequence of a seller's move consists of both its direct and indirect effects. The *direct effects* are those which result from a seller's own action, rival sellers not reacting to his action. The *indirect effects* are those which result from the reaction of the rival sellers to the moves made by a seller.

Chamberlin suggests in his model that, if rival firms are assumed to recognise their interdependence and act accordingly, a stable equilibrium can be reached where each firm charges monopoly price. When all firms are in equilibrium, industry profit is maximised. Chamberlin's oligopoly model of 'small group' can be best

understood if presented in the framework of Cournot's duopoly model since Chamberlin follows Cournot to develop his own model.

Cournot's model is reproduced in Fig. 14.1, except the ordinate  $JK$ . Assuming there are two firms,  $A$  and  $B$ , let  $A$  first enter the market as a monopolist. Following the profit maximisation rule, firm  $A$  will produce  $OQ$  and charge monopoly price  $OP_2 (= PQ)$ . When firm  $B$  enters the market, it considers that  $PM$  is its demand curve. Under Cournot's assumption, firm  $B$  will sell output  $QN$  at price  $OP_1$ . As a result, market price falls from  $OP_2$  to  $OP_1$ . It is now  $A$ 's turn to appraise the situation. At this point, Chamberlin deviates from Cournot's model. According to Cournot's model, firm  $A$  does not recognise their interdependence and acts independently. Chamberlin however assumes that firm  $A$  does recognise the interdependence between them and it does recognise the fact that  $B$  will react to its decisions. Therefore, firm  $A$  decides to compromise with the existence of firm  $B$ , and decides to reduce its output to  $OK$  which is half of the monopoly output,  $OQ$ . Its output  $OK$  equals  $B$ 's output  $QN (= KQ)$ . In its turn, firm  $B$  also recognises their interdependence. It realises that  $KQ$  is the most profitable output for it. Thus, the industry output is  $OQ$  which is the same as monopoly output, and market price is  $OP_2 (= PQ)$  which equals monopoly price. Thus, according to Chamberlin, by recognising their interdependence, the firms reach an equilibrium which is the same as monopoly equilibrium and share the market equally. One of the firms supplies  $OK$  and the other supplies  $KQ$  where  $OK + KQ = OQ$ , the profit maximising monopoly output. This equilibrium is stable because under the condition of interdependence, firms do not gain by changing their price and output.



**Fig. 14.1** Chamberlin's Model of Stable Oligopoly Equilibrium

Chamberlin's model is regarded as an improvement over the earlier models, at least in respect of its behavioural assumption of interdependence. His model

has however been criticised on the grounds that his idea of joint profit maximisation is beset with problems of estimating demand and cost functions. Unless demand and cost functions are fully known to the competitors, joint profit maximisation is doubtful.

(ii) *Sweezy's Kinked-Demand Curve Model of Oligopoly*

The origin of kinked-demand curve can be traced in Chamberlin's theory of monopolistic competition. Later, Hall and Hitch used *kinked-demand curve* to explain rigidity of prices in oligopolistic market. However, neither Chamberlin nor Hall and Hitch used kinked-demand curve as a tool of analysis in their respective theories. It was Paul M. Sweezy who used the kinked-demand curve in his model of price stability in oligopolistic market. In this section, we will briefly describe Sweezy's Model.

The kinked-demand curve model developed by Paul M. Sweezy has features common to most oligopoly pricing models. This is the best known model to explain, relatively more satisfactorily, the behaviour of the oligopolistic firms. The *kinked-demand curve* analysis does not deal with price and output determination. Instead, it seeks to establish that once a price-quantity combination is determined, an oligopoly firm will not find it profitable to change its price even in response to the small changes in the cost of production. The logic behind this proposition is as follows. An oligopoly firm believes that if it reduces the price of its product, the rival firms would follow and neutralise the expected gain from price reduction. But, if it raises the price, the firms would either maintain their prices or even indulge in price-cutting, so that the price-raising firms stand to lose, at least, a part of its market share. This behaviour is true for all the firms. The oligopoly firms would therefore find it more desirable to maintain the prevailing price and output.

To look more closely at the kinked-demand curve analysis, let us look into the possible actions and reactions of the rival firms to the price changes made by one of the firms.

There are three possible ways in which rival firms may react to change in price by one of the firms: (i) the rival firms follow the price changes, both cut and hike; (ii) the rival firms do not follow the price changes; (iii) rival firms do not react to price-hikes but they do follow the price-cutting. If rival firms react in manners (i) an oligopoly firm taking lead in changing prices will face demand curve  $dd'$  in Fig. 14.2. If rival firms react in manner (ii), the firm faces demand curve  $DD'$ . The demand curve  $dd'$  which is based on reaction (i) is less elastic than the demand curve  $DD'$  which is based on reaction (ii). Demand curve  $dd'$  is less elastic because changes in demand in response to changes in price are restrained by the counter-moves by the rival firms.



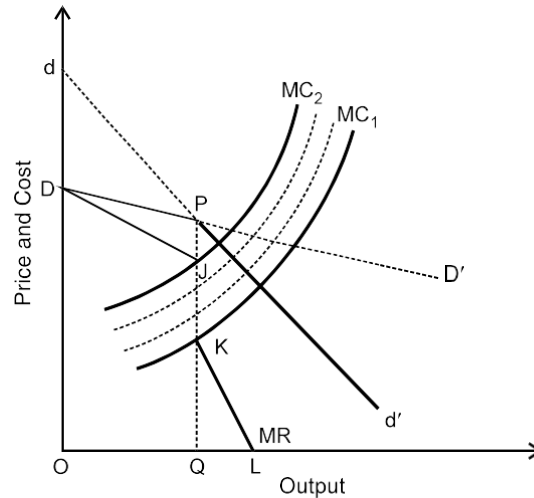


Fig. 14.2 Kinked-demand Analysis

Given the two demand curves, let point  $P$  represent the equilibrium price-quantity combination of an oligopolist. Let us now introduce reaction (iii), i.e., rival firms follow the oligopolist leading in price-cutting when he reduces his price but do not follow him when he increases his price. This asymmetrical behaviour of the rival firms, makes only a part of each of the two demand curves relevant for the oligopolist. This can be established by allowing an oligopolist to alternatively increase and decrease his price. If an oligopolist increases his price and his rivals do not follow him, he loses a part of his market to his rivals. The demand for his product decreases considerably indicating a greater elasticity. The oligopolist is therefore forced down from demand curve  $dP$  to  $DP$ . Thus, the relevant segment of demand curve for the oligopolist is  $DP$ .

On the other hand, if an oligopolist decreases his price, the rival firms, react by cutting down their prices by an equal amount or even more. This counter move by the competitors prevents the oligopolist from taking full advantage of price-cut along the demand curve  $DD'$ . Therefore, his demand curve below point  $P$  rotates down. Thus, the relevant segment of demand curve for the oligopolist (below point  $P$ ) is  $Pd'$ . If the two relevant segments of the two demand curves are put together, the relevant demand curve for the oligopolist is  $DPd'$  which has a *kink* at point  $P$ . Therefore, it is called a 'Kinked-demand curve'.

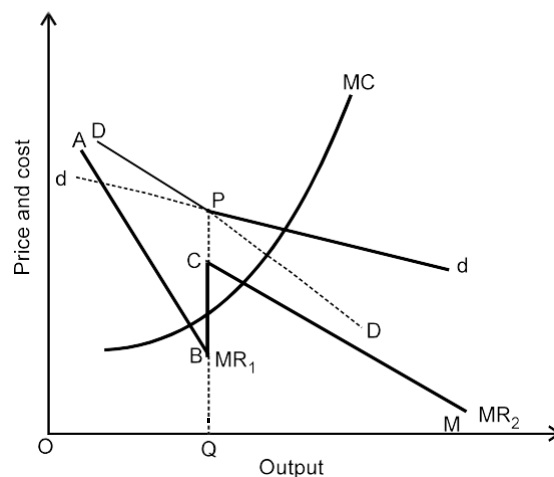
Consider now the relationship between  $AR (=D)$  and  $MR$ . We know that  $MR = AR - AR/e$ . The  $MR$  curve, drawn on the basis of this relationship, will take a shape as shown by  $DJKL$  in Fig. 14.2. It is discontinuous between point  $J$  and  $K$ , at output  $OQ$ . Suppose that the original marginal cost curve resembles  $MC_1$  which intersects  $MR$  at point  $K$ . Since at output  $OQ$ , the necessary condition of maximum profit ( $MR = MC$ ) is satisfied, the oligopolist is earning maximum profit. Now, if marginal cost curve shifts upwards to  $MC_2$  or to any level between points  $J$  and  $K$ , his profit would not be affected. Therefore, he has no motivation for increasing or decreasing his price. It is always beneficial to stick to the price and output. Thus, both price and output are stable.

The oligopolists will think of changing their price and output only if  $MC$  rises beyond point  $J$  or decreases below point  $K$  (Fig. 14.2). But, even if it so happens, price and output would tend to stabilise. Suppose that the general level of costs rises for the industry so that  $MC$  moves above point  $J$ . The oligopolists will ultimately find it profitable to raise the price. When one of the oligopolists raises his price, his competitors match the price increase. As a result, the kinked-demand curve shifts upward to a new position and the point of kink shifts rightward and horizontally. Again, at the new price there is no incentive for any oligopolist to raise his price. Therefore, price tends to stabilise.

Alternatively, if  $MC$  moves down below point  $K$ , firms get incentive to reduce their price. When one firm cuts its price, others follow with matching price-reduction. There is a possibility of competitors reducing their prices by a greater margin. The only way to prevent this situation is that the oligopolist must keep his costs as low as possible, at least lower than that of his competitors. This is the reason why there is keen technological competition in an oligopolistic market. In other words, there is incentive for oligopolists to use new and efficient techniques of production, to introduce new products, to make innovations, to increase their productivity or to reduce their cost of production to the possible minimum. They find it safe to concentrate on efficiency rather than to indulge in price-war.

#### Oligopoly Equilibrium in a Buoyant Market

The kinked-demand curve analysis presented above is found relevant mainly to those oligopolistic industries which are faced with normal market conditions. A similar analysis may be performed for the oligopolistic industries enjoying buoyant business conditions. The relevant kinked-demand curve under buoyant business conditions is given by the curve  $DPd$  in Fig. 14.3.



**Fig. 14.3** Equilibrium of Oligopolistic Firms under Buoyant Conditions

The curve  $DPd$  (in Fig. 14.3) is based on the assumption that when an oligopoly firm raises its price above  $P$ , other firms also raise their prices by an

equal amount, but when it reduces its price below  $P$ , the rival firms do not cut their prices because their sales are not seriously effected if market is buoyant. Since rival firms follow the price rise, only  $DP$  segment of the demand curve  $DD_1$  is relevant. But, rival firms do not follow price-cut because they are able to sell their output in the buoyant market in spite of the lower price set by one of the firms. Therefore, the price-raising firm retains its individual demand curve  $Pd_1$ . The net result is that the demand curve for the product of an oligopolist takes a shape as shown by  $DPd_1$ . Despite the difference in prices, the market price would tend to stabilise. It is obvious from the fact that  $MC$  intersects both  $MR$  curves,  $AB$  and  $CM$ . It means that some firms equalise their  $MC$  and  $MR$  at  $MR_1$  and some firms at  $MR_2$ . So both categories of firms maximise their profits under buoyant business conditions. This very condition may however create instability in price and output. Since  $MC$  intersects both the  $MR$  curves, it implies that profits increase irrespective of whether price rises or falls. Therefore, while some firms would like to raise the price, others would prefer to reduce their price. This makes the market conditions unstable, and prices and outputs become uncertain. It is therefore not possible to make any generalisations about oligopoly prices and output under buoyant conditions.

#### *Some Implications of Sweezy's Model*

According to Sweezy, his model for price stability in an oligopolistic market has the following implications:

**First**, since elasticity of the demand curve below point  $P$  is assumed to be less than unity and  $MR$  beyond a point is negative, the conditions of short-run equilibrium are not precise. That is, profit maximisation rule,  $MC = MR$ , cannot be applied to the short-run conditions.

**Second**, since  $MC$  can shift up and down between the finite points  $J$  and  $K$  (Fig. 14.7),  $MR$  remaining the same, his model deviates from the marginal productivity theory, i.e., factor prices do not equal their marginal revenue productivity.

**Third**, any short-term disturbance in  $MC$  will not affect the equilibrium price or output and the total profits. Thus, the general belief that a successful strike by the trade unions reduces profits gets little theoretical support from Sweezy's model.

#### *Criticism*

The major criticism against this model is that it explains only the stabilisation of output and price. It does not tell, why and how the initial price is fixed at a certain level. The Sweezy's thesis must therefore be regarded as an *ex-post* rationalisation rather than as an *ex-ante* explanation of market equilibrium.

Besides, Sweezy's claim of price stability does not stand the test of empirical verification: there is a surprising lack of price rigidity. Monopoly prices have been

found more stable than oligopoly prices. However, economists are divided on the issue of price rigidity. While Stigler doubts the existence of kinked-demand curve and price rigidity, Liebhafsky finds considerable evidence of price rigidity in the US. Cohen and Cyert argue that kink in the demand curve and price rigidity may exist for a brief period for lack of inter-firm information, particularly when new and unknown rivals enter the market. They are of the opinion that kink is clearly not a stable long-run equilibrium.

### 14.2.3 Price Determination under Collusive Models of Oligopoly

From the non-collusive models, we now turn to the collusive models of the oligopoly theory. In the non-collusive models, oligopoly firms are assumed to act independently. In the *collusive models*, however, firms are assumed to act in unison, i.e., in collusion with one another. This assumption is based on empirical facts, rather than being conjectural.

#### Why Collusion?

There are at least three major factors which bring collusion between the oligopolistic firms. **First**, collusion reduces the degree of competition between the firms and helps them act mono-polistically in their effort of profit maximisation. **Second**, collusion reduces the oligopolistic uncertainty surrounding the market since cartel members are not supposed to act independently and in the manner that is detrimental to the interest of other firms. **Third**, collusion forms a kind of barrier to the entry of new firms.

Collusion between oligopoly firms may take many forms depending on their relative strength, their objective and legal status of collusion. There are however two main types of collusion (i) Cartels; and (ii) Price leadership.

#### 1. Cartels under Oligopoly

A **cartel** is a formal organisation of the oligopoly firms in an industry. Cartels are the perfect form of collusion. A general purpose of cartels is to centralise certain managerial decisions and functions of individual firm in the industry with a view to promoting common benefits. Cartels may be in the form of *open collusion* or *secret collusion*. Whether open or secret, cartel agreements are *explicit* and *formal* in the sense that agreements are enforceable on member firms trying to pursue an independent pricing policy. Cartels are therefore regarded as the perfect form of collusion. Cartels and cartel type agreements between the firms in manufacturing and trade are illegal in most countries. Yet, cartels in the broader sense of the term exist in the form of trade associations, professional organisations and the like.

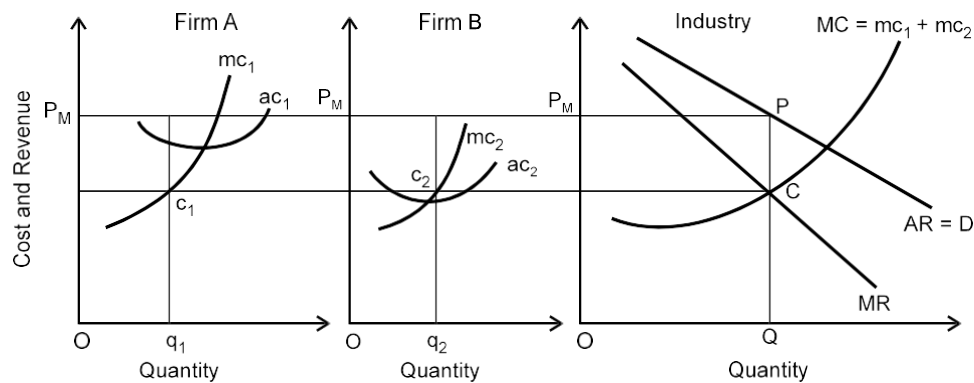
A cartel performs a variety of services for its members. The two typical services of central importance are (i) fixing price for joint maximisation of industry profits; and (ii) market-sharing between its members. In this section, we will examine, these two activities of cartels.

Let us suppose that a group of firms producing a homogeneous commodity form a cartel aiming at joint profit maximisation. The firms appoint a central management board with powers to decide the following aspects:

- (i) the total quantity to be produced;
- (ii) the price at which the product has to be sold; and
- (iii) share or each firm in the total output.

The central management board is provided with cost figures of individual firms. Besides, it is supposed to obtain the necessary data required to formulate the market demand (*AR*) curve. The management board calculates the marginal cost (*MC*) and marginal revenue (*MR*) for the industry. Furthermore, the management board holds the position of a multiplant monopoly. It determines the price and output for each firm in the manner a multiplant monopoly determines the price and output for each plant.

The model of price and output determination for each is presented in Fig. 14.4. It is assumed for the sake of convenience that there are only two firms, *A* and *B*, in the cartel. Their respective cost curves are given in the first two panels of Fig. 14.9. In the third panel, the *AR* and *MR* curves represent the revenue conditions of the *industry*. The *MC* curve is the summation of *MC* curves of the individual firms. The *MC* and *MR* intersect at point *C* determining the industry output at *OQ*. The market price is determined at *PQ*. The industry output *OQ* is so allocated between firms *A* and *B* that for each of them  $MC = MR$ . The share of each firm in the industry output, *OQ*, can be determined by drawing a line from point *C* and parallel to *X*-axis through  $MC_2$  and  $MC_1$ . The points of intersection  $C_1$  and  $C_2$  determine the level of output for firms *A* and *B*, respectively. Thus, the share of each of the two firms *A* and *B*, is determined at  $Oq_1$  and  $Oq_2$ , respectively, where  $Oq_1 + Oq_2 = OQ$ . Their respective profit can be computed as  $(P_m - \text{firm's } ac) \times \text{firm's output}$ , which is maximum. The total profit of each firm maybe different. But there is no motivation for changing price-quantity combination, since their individual profit is maximum.



**Fig. 14.4** Price and Output Determination under Cartel

Although monopoly solution to joint profit maximisation by cartels look theoretically sound, William Fellner gives the following reasons why joint profits may not be maximised.

**First**, it is difficult to estimate market demand curve accurately since each firm thinks that the demand for its own product is more elastic than the market demand curve because its product is a perfect substitute for the product of other firms.

**Secondly**, similarly an accurate estimation of industry's *MC* curve is highly improbable for lack of adequate and correct cost data. If industry's *MC* is incorrectly estimated, industry output can be only incorrectly determined. Hence joint profit maximisation is doubtful.

**Thirdly**, cartel negotiations take a long time. During the period of negotiation, the composition of the industry and its cost structure may change. This may render the estimates irrelevant, even if they are correct. Besides, if the number of firms increase beyond 20 or so, cartel formation becomes difficult, or even if it is formed, it soon breaks down.

**Fourthly**, there are 'Chiselers' who have a strong temptation to give secret concessions to their customers. This tendency in the members reduces the prospect of joint profit maximisation.

**Fifthly**, if cartel price, like monopoly price, is very high, it may invite government attention and interference. For the fear of government interference, members may not charge the cartel price.

**Sixthly**, another reason for not charging the cartel price is the fear of entry of new firms. The high cartel price which yields monopoly profit may attract new firms to the industry. To prevent the entry of new firms, some firms may decide on their own not to charge the cartel price.

**Finally**, another reason for not charging the cartel price is the desire to build a public image or good reputation. Some firms may, to this end, decide to charge only a fair price and realise only a fair profit.

### *Cartel and Market-Sharing*

The market-sharing cartels are more common because this kind of collusion permits a considerable degree of freedom in respect of style of the product, advertising and other selling activities. There are two main methods of market allocations: (i) non-price competition, and (ii) quota system.

#### *(i) Non-price Competition*

The non-price competition agreements are usually associated with loose cartels. Under this kind of arrangement between the firms, a uniform price is fixed and each firm is allowed to sell as much as it can at the cartel price. The only requirement is that firms are not allowed to reduce the price below the cartel price.

The cartel price is a bargain price. While low-cost firms press for a low price, the high-cost firms press for a higher price. But the cartel price is so fixed by mutual consent that all member firms are able to make some profits. But the firms are allowed to compete with one another in the market on a non-price basis. That is, they are allowed to change the style of their product, innovate new designs, and to promote their sales by advertising.

Whether this arrangement works or breaks down depends on the cost conditions of the individual firms. If some firms expect to increase their profits by violating the price agreements, they will indulge in cheating by charging lower price. This may lead to a price-war and cartel may breakdown.

### (ii) Quota System

The second method of market-sharing is **quota system**. Under this system, cartel fixes a quota of market-share for each firm. There is no uniform principle for fixing quota. In practice, however, the main considerations are (i) bargaining ability of a firm and its relative importance in the industry, (ii) the relative sales of the firms in pre-cartel period, and (iii) production capacity of the firm. The choice of base period depends on the bargaining ability of the firm.

Another popular basis of market-sharing is the geographical division of market. The examples of this kind of market-sharing are mostly found in the case of international markets.

**Unequal Quota for Unequal Firms.** Fixation of quota is a difficult proposition. Nevertheless, some theoretical guidelines for market-sharing have been suggested by the economists: (i) unequal quota for unequal firms, i.e., firms with different cost curves, and (ii) equal quota for equal firms—firms with identical cost and revenue curves.

A reasonable criterion of ideal market-sharing can be to share the total market between the cartel members in such proportions that the industry's marginal cost equals the marginal cost of individual firms. This criterion is illustrated in Fig. 14.4. The profit maximising output of the industry is  $OQ$ . The industry output  $OQ$  is shared between the two firms  $A$  and  $B$ , as  $Oq_1$  and  $Oq_2$ , respectively. Note that  $OQ = Oq_1 + Oq_2$ . At output  $Oq_1$ ,  $mc$  of firm  $A$  equals industry's marginal cost,  $MC$ , and at output  $Oq_2$ ,  $mc$  of firm  $B$  equals  $MC$ . Thus, under quota system, the quota for firms  $A$  and  $B$  may be fixed as  $Oq_1$  and  $Oq_2$ , respectively. Given the quota allocation, the firm may set different prices for their product depending on the position and elasticity of their individual demand curves. This criterion is identical to the one adopted by a multiplant monopolist in the short-run, to allocate the total output between the plants.

**Equal Quota for Equal Firms.** Another reasonable criterion for market-sharing under quota system is equal market-share for equal firms. This criterion is applicable where all have identical cost and revenue curves. This criterion also leads to a monopoly solution. It also resembles Chamberlin's duopoly model.

To illustrate the quota allocation, let us assume that there are only two firms, *A* and *B*. Their *AR*, *MR* and *MC* curves are given as shown in Fig. 14.5(a) and (b). The market revenue and cost curves, which are obtained by adding up individual revenue and cost curves, respectively, are presented in part (c) of the figure. The industry output is determined at  $OQ$ . The quota for each firm, which

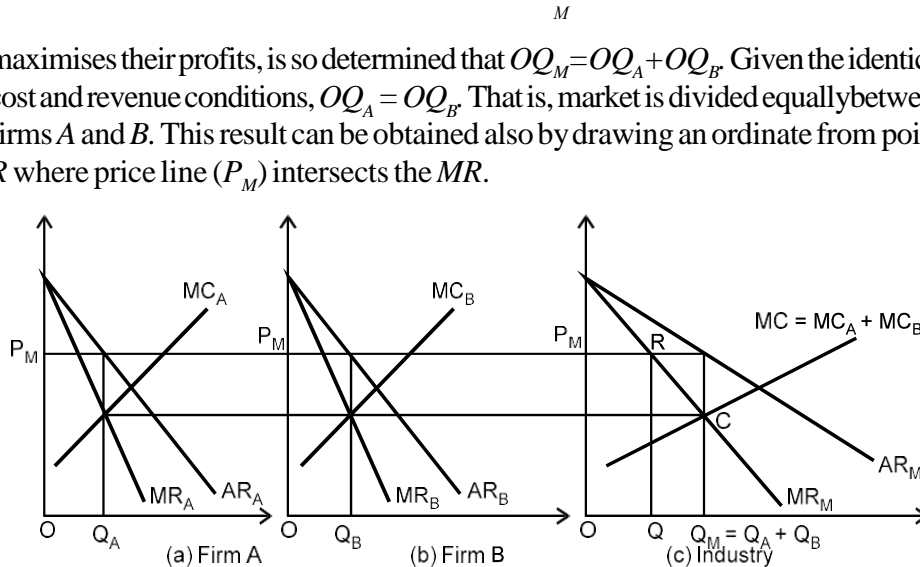


Fig. 14.5 Quota Allocation under Cartel Agreements

It may be mentioned at the end that cartels do not necessarily create the conditions for price stability in an oligopolistic market. Most cartels are loose. Cartel agreements are generally not binding on the members. Cartels do not prevent the possibility of entry of new firms. On the contrary, by ensuring monopoly profits, cartels in fact create conditions which attract new firms to the industry. Besides, chiselers and free-riders create conditions for instability in price and output.

## 2. Price Leadership Models of Oligopoly

Collusion through *price leadership* is another form of collusion between oligopoly firms. Price leadership is an informal position of a firm in an oligopolistic setting to lead other firm in fixing price of their product. This leadership may emerge spontaneously due to technical reasons or out of tacit or explicit agreements between the firms to assign leadership role to one of them.

The *spontaneous* price leadership maybe the result of such technical reasons as size, efficiency, economies of scale or firm's ability to forecast market conditions accurately or a combination of these factors. The most typical case of price leadership is the leading role played by the dominant firm, the largest firm in the industry. The dominant firm takes lead in price changes and the smaller ones follow. Sometimes price leadership is *barometric*. In the barometric price leadership, one of the firms, not necessary the dominant one, takes lead in announcing change in price, particularly when such a change is due but is not affected due to uncertainty in the market.



The price leadership is possible under both *product homogeneity* and *product differentiation* or *heterogeneity*. There may be however price differentials commensurating with product differentiation. Price differentials may also exist on account of cost differentials.

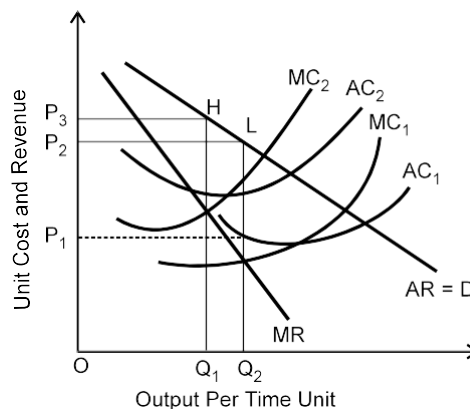
Another important aspect of price leadership is that it often serves as a means to price discipline and price stabilisation. Achievement of this objective establishes an effective price leadership. Such price leadership can however exist only when (i) number of firms is small; (ii) entry to the industry is restricted; (iii) products are, by and large, homogeneous; (iv) demand for industry is inelastic or has a very low elasticity; and (v) firms have almost similar cost curves.

The three common types of price leaderships are:

- (i) Price leadership by a low-cost firm;
- (ii) Price leadership by a dominant firm;
- (iii) Barometric price-leadership.

*(i) Price leadership by a Low-cost Firm*

How price and output decisions are taken under price leadership of a low-cost firm is illustrated in Fig. 14.6. Suppose all the firms face identical revenue curves as shown by  $AR = D$  and  $MR$ . But the largest firm or the low-cost firm, has its cost curves as shown by  $AC_1$  and  $MC_1$  whereas all other rival firms, smaller in size have their cost curves as shown by  $AC_2$  and  $MC_2$ . The largest firm has the economies of scale and its cost of production is lower than that of other firms. Given the cost and revenue conditions, the low-cost firm would find it most profitable to fix its price at  $OP_2 (= LQ_2)$  and sell quantity  $OQ_2$ . Since at this level of output its  $MC = MR$ , its profit will be maximum. On the other hand, the high-cost firms would be in a position to maximise their profit at price  $OP_3$  and quantity  $OQ_1$ .



**Fig. 14.6** Price Leadership

However, if low-cost firms charge profit maximising price  $OP_3$ , they would lose their customers to the low-cost firm charging a lower price  $OP_2$ . The high-cost firms are therefore forced to accept the price  $OP_2$  and recognise the price leadership of the low-cost firm. Note that the low-cost firm can eliminate other firms and become a monopolist, by cutting its price down to  $OP_1$ . At price  $OP_1$ , the low-

cost firm can sell the same quantity  $OQ_2$  and make, of course, only normal profit as its  $AC = \text{price } OP_1$ . But, it may not do so for the fear of anti-monopoly laws.

(ii) *Price leadership by the Dominant Firm*

Price leadership by the dominant firm is more common than by a low-cost firm. In the analysis of price leadership by a dominant firm, it is assumed that there exists a large-size firm in the industry, which supplies a large proportion of the total market. The dominance of the large firm is indicated by the fact that it could possibly eliminate all its rival firms by price-cutting. But then the large firm gains the status of a monopoly which may create legal problems. The dominant firm therefore compromises with the existence of rival firms in the market. It uses its dominance to set its price so as to maximise its price. The smaller firms have no alternative but to accept the price set by the dominant firm. The smaller firms recognise their position and behave just like a firm in a perfectly competitive market. That is, smaller firms assume that their demand curve is a straight horizontal line.

The price leadership and market sharing between the dominant firm and the other firms as a group is illustrated in Fig. 14.12. Suppose that the market demand curve is given by  $DD_m$  in part (a) of the figure. The problem confronting the dominant firm is to determine its price and output that will maximise its profits, leaving the rest of the market to be jointly supplied by the small firms. Now the dominant firm has to find its own demand curve. Given the market demand curve ( $DD_m$ ) and joint supply curve of small firms ( $SS_s$ ), the dominant firm finds its demand curve by deducting from the market demand the quantity supplied jointly by the small firms below the equilibrium price. The part of the market demand not supplied by the small firms will be its own share. Thus, the market share of the dominant firm equals the market demand less the share of small firms.

For example, suppose equilibrium price is set at  $OP_3$ , the total supply by the smaller firms is  $P_3E$  which equals the market demand. Therefore, at price  $OP_3$ , the market left for the dominant firm is zero. When price is  $OP_2$ , market demand is  $\bar{P}B$  out of which  $\bar{P}A$  supplied by smaller firms. The market unsupplied by the smaller firms is  $AB$ . Thus, at price  $OP_2$ , the demand for dominant firm's product equals

$$\bar{P}B - \bar{P}A = AB$$

Similarly, when price is reduced to  $OP_1$ , the demand for dominant firm's product is  $CF$ . Following this process, the market-share of the dominant firm at other prices can be easily obtained.

The information so derived and plotted graphically gives  $P_3D_L$  as the demand curve for the dominant form [Fig. 14.7(b)]. Since the relation between  $AR$  and  $MR$  is known, the  $MR$  curve for the dominant firm can be derived as  $MR_L$  [Fig. 14.7(b)]. If the  $MC$  curve of the dominant firm is assumed to be given as  $MC_L$ , its profit maximising output will be  $OQ_L$  and price  $O\bar{P}Q_L$ .

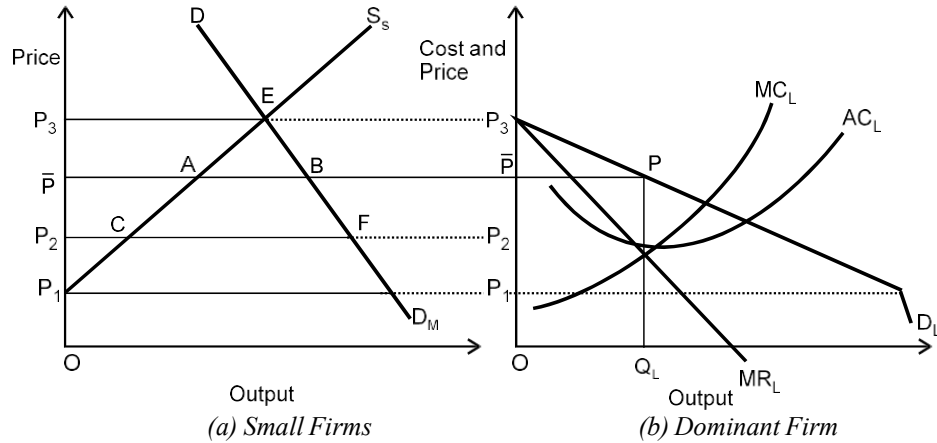


Fig. 14.7 Price Determination by the Dominant Firm

Once the dominant firm sets its price at  $O\bar{P}$ , the market demand curve for the small firms is the horizontal straight line  $\bar{P}B$  because they can sell, at this price, as much as they can produce. But, in order to maximise their joint profits, small firms will produce only  $\bar{P}A$ . Recall that given the price, the line  $AR = MR$  is the same as their  $AR = MR$  line and their supply curve  $P_1S_s$ , intersects  $AR = MR$  at point A. For small firms, therefore, profit-maximising output is  $\bar{P}A$ .

Finally, the dominant firm sets its price at  $O\bar{P}$  which is accepted by the small firms. Thus, the dominant firm plays the role of a price leader. If it wants to eliminate the small firm it may set its price at  $OP_1$  (though at a loss in the short run) at which small firms would not be able to survive. But, for the legal reason mentioned above, the dominant oligopoly firm would not do so. It would prefer, and be content, with its position of a price leader.

### Critical Appraisal of Price Leadership Model

The price leadership model, as presented above, yields a stable solution to the problem of oligopoly pricing and output determination, only if small firms faithfully follow the leader. That is, small firms produce a right quantity and charge the price set by the dominant firm. Besides, the model requires that the dominant firm should be both a large and low-cost firm. For, if a firm does not enjoy the advantages of being large enough and, consequent upon it, the advantages of its low cost, it cannot act as a price leader.

In practice, however, one finds many cases of price leadership by a firm which is neither a large nor a low-cost firm. But such cases are found mostly under

recessionary conditions when a relatively smaller firm reduces its price to survive in the market.

Furthermore, if a leading firm loses its cost advantages, it loses also its leadership. Such cases are frequent in the real business world. Leadership also changes following the innovations of products and techniques of production by the smaller firms.

Besides, where there are many large firms of equal size and have some cost advantage, price leadership of any firm or group of firms becomes less probable, particularly when number of small firms is smaller than that of large firms. Under such conditions, barometric leadership emerges.

Lastly, it is assumed that entry of new firms is prevented either by low-cost or by initial high cost. In practice, however, many firms having the capacity to diversify their products enter the industry with relatively initial low-cost.

For these reasons, leadership model is not a realistic one as it is based on unrealistic assumptions. For the same reasons, the solution given by leadership model may not be stable.

### *(iii) Barometric Price Leadership*

Another form of price leadership is *barometric* price leadership. In this form of price leadership, a firm initiates well publicised changes in price that are generally followed by the rival firms in the industry. The price leader may not necessarily be the largest firm of the industry. The barometric firm is however supposed to have a better knowledge of prevailing market conditions and has an ability to predict the market conditions more precisely than any of its competitors. This qualification of the barometric firm should have been established in the past. Price decisions by a firm having the qualifications of price leadership is regarded as a barometer which reflects the changes in business conditions and environment of the industry. The price changes announced by the barometric firm serves as a barometer of changes in demand and supply conditions in the market.

The barometric leadership evolves for various reasons of which the major ones are following.

**First**, the rivalry between the larger firms may lead to cut-throat competition to the disadvantage of all the firms. On the other hand, rivalry between the larger firms may make them unacceptable as a leader. So a firm which has better predictive ability emerges as price leader.

**Secondly**, most firms in the industry may have neither the capacity nor the desire to make continuous calculations of cost, demand and supply conditions. Therefore, they find it advantageous to accept the price changes made by a firm which has a proven ability to make reasonably good forecasts.

**Thirdly**, Kaplan *et. al.*, state that barometric price leadership often develops as a reaction to a long economic warfare in which all the firms are losers.

## Concluding Remarks on Oligopoly Models

Most oligopoly models concentrate on price competition. In reality, however, as it is obvious from the above discussion that oligopolists may be reluctant to wage price-war and encroach upon each other's market-share. It means that there is an absence of price-competition in the oligopolistic market structure. The absence of price-competition should not mean the absence of competition among oligopoly firms. In fact, the competition among oligopoly firms takes the form of *non-price competition*. The forms of non-price competition are diverse. Yet, there are two most important methods of non-price competition.

**First**, non-price competition involves product differentiation which is intended to attract new customers by creating preference for the new design and variety of product.

**Secondly**, perhaps the most important technique of non-price competition is advertisement. The primary objective of advertising is to make the demand curve for the product shift upward. The sellers try to encroach on the markets of other sellers through advertising. Advertising is also necessary to retain the market-share if there is tough competition between the firms.

---

## 14.3 DUOPOLY

---

When there are only two sellers of a product, there exists duopoly, a special case of oligopoly. Duopoly is a special case in the sense that it is the limiting case of oligopoly as there must be at least two sellers to make the market oligopolistic in nature. In this section, we explain famous 'classical models' of duopoly.

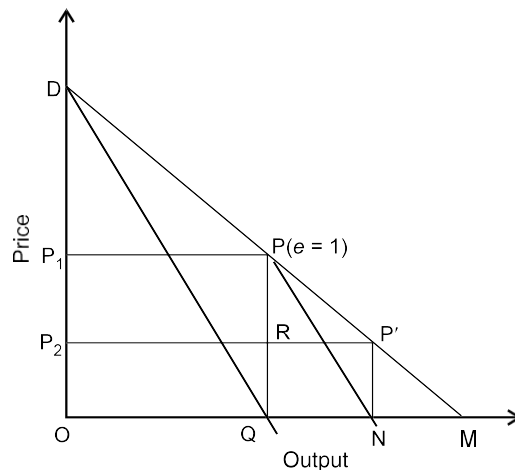
### *Cournot's Duopoly Model*

Augustin Cournot, a French economist, was the first to develop a formal duopoly model in 1838. To illustrate his model, Cournot assumed:

- (a) two firms, *A* and *B*, each owning an artesian mineral water wells;
- (b) both operate their wells at zero marginal cost;
- (c) both face a downward sloping straight line demand curve;
- (d) each seller acts on the assumption that his competitor will not react to his decision to change his output and price. This is Cournot's behavioural assumption.

On the basis of this model, Cournot has concluded that each seller ultimately supplies one-third of the market and both the sellers charge the same price. And, one-third of the market remains unsupplied.

Cournot's duopoly model is presented in Fig. 14.8. To begin the analysis, suppose that *A* is the only seller of mineral water in the market. In order to maximize his profits or revenue, he sells quantity  $OQ$  at which his  $MC = O = MR$ , at price  $OP_1$ . His total profit is  $OP_1PQ$ .



**Fig. 14.8** Price and Output Determination under Duopoly: Cournot's Model

Now let *B* enter the market. The part of market open to him equals  $QM$  which is half of the total market. Note that  $QM$  is the part of the market left unsupplied by *A*. It means that *B* can sell his product in the remaining half of the market,  $PM$  being the relevant part of demand curve for him. *B* assumes that *A* will not change his price and output because he is making the maximum profit. That is, *B* assumes that *A* will continue to sell  $OQ$  at price  $OP_1$ . Thus, the market available to him is  $QM$  and the relevant demand curve is  $PM$ . When he draws his  $MR$  curve,  $PN$ , it bisects  $QM$  at point  $N$  where  $QN = NM$ . In order to maximise his revenue, *B* sells  $QN$  at price  $OP_2 = P_2N$ . His total revenue is maximum at  $QRP_2N$ . Note that *B* supplies only  $QN = 1/4 = 1/2 \times 1/2$  of the market.

With the entry of *B*, price falls to  $OP_2$ . Therefore, *A*'s expected profit falls to  $OP_2RQ$ . Faced with this situation, *A* attempts to adjust his price and output to the changed conditions. He assumes that *B* will not change his output  $QN$  and price  $OP_2$  as he (*B*) is making the maximum profit. Accordingly, *A* assumes that *B* will continue to supply  $1/4$  of the market and, therefore, he has  $3/4 = 1 - 1/4$  of the market available to him. To maximise his profit, *A* will supply  $1/2 (3/4) = 3/8$  of the market. Note that *A*'s market share has fallen from  $1/2$  to  $3/8$ .

Now it is *B*'s turn to react. Following Cournot's assumption, *B* assumes that *A* will continue to supply only  $3/8$  of the market and the market open to him equals  $1 - (3/8) = 5/8$ . To maximise his profit under the new conditions, *B* will supply  $1/2 (5/8) = 5/16$  of the market. It is now for *A* to reappraise the situation and adjust his price and output accordingly.

This process of action and reaction continues in successive periods. In the process, *A* continues to lose his market share and *B* continues to gain. Eventually, a situation is reached when their market share equals at 1/3 each. Any further attempt to adjust output produces the same result. The firms, therefore, reach their equilibrium position with each supplying 1/3 of the market and 1/3 of the market remaining unsupplied.

**Table 14.1** Market Sharing in Cournot's Model

Period	Firm A	Firm B
I	$\frac{1}{2}(1) = \frac{1}{2}$	$\frac{1}{2} \left( \frac{1}{2} \right) = \frac{1}{4}$
II	$\frac{1}{2} \left( 1 - \frac{1}{4} \right) = \frac{3}{8}$	$\frac{1}{2} \left( 1 - \frac{3}{8} \right) = \frac{5}{16}$
III	$\frac{1}{2} \left( 1 - \frac{5}{16} \right) = \frac{11}{32}$	$\frac{1}{2} \left( 1 - \frac{11}{32} \right) = \frac{21}{64}$
IV	$\frac{1}{2} \left( 1 - \frac{21}{64} \right) = \frac{43}{128}$	$\frac{1}{2} \left( 1 - \frac{43}{128} \right) = \frac{85}{256}$
...	.....	.....
...	.....	.....
...	.....	.....
N	$\frac{1}{2} \left( 1 - \frac{1}{3} \right) = \frac{1}{3}$	$\frac{1}{2} \left( 1 - \frac{1}{3} \right) = \frac{1}{3}$

Cournot's equilibrium solution is stable. For, given the action and reaction, it is not possible for any of the two sellers to increase their market share. Cournot's model of duopoly can be extended to the case of general oligopoly. For example, suppose there are three sellers, the industry and firms will be in equilibrium when each firm supplies 1/4 of the market. The three sellers together supply 3/4 = 3 (1/4) of the market, 1/4 of the market remaining unsupplied. The formula for determining the share of each seller in an oligopolistic market is  $Q \div (n + 1)$ , where  $Q$  = market size, and  $n$  = number of sellers.

**Algebra of Cournot's Model**

Cournot's duopoly model may also be presented algebraically. Let us suppose that market demand function is given by a linear function given as

$$Q = 90 - P \quad \dots(14.1)$$

We have noted above that, under zero cost condition, profit is maximum where  $MC = MR = 0$  and profit-maximising output equals  $Q/2$ .

Thus, when firm  $A$  is a monopolist in the market, his profit-maximising output ( $Q_A$ ), according to the profit-maximising rule under zero cost condition, is given by

$$Q_A = 1/2 (90 - P) \quad \dots(14.2)$$

When another firm,  $B$ , enters the market, its profit-maximising output equals

$$Q_B = 1/2[1/2(90 - P)] \quad \dots(14.3)$$

Thus, the respective share of firms,  $A$  and  $B$  is fixed at  $Q_A$  and  $Q_B$ . The division of market output may be expressed as

$$Q = Q_A + Q_B = 90 - P \quad \dots(14.4)$$

The demand function for the firm  $A$  may now be expressed as

$$Q_A = (90 - Q_B) - P \quad \dots(14.5)$$

and for the firm  $B$  as

$$Q_B = (90 - Q_A) - P \quad \dots(14.6)$$

Given the demand function (14.5), the market open to firm  $A$  (at  $P = 0$ ) is  $90 - Q_B$ . The profit-maximising output for  $A$  can be written as

$$Q_A = \frac{90 - Q_B}{2} \quad \dots(14.7)$$

and for  $B$ , as

$$Q_B = \frac{90 - Q_A}{2} \quad \dots(14.8)$$

The Eqs. (14.7) and (14.8) represent the reaction functions of firms  $A$  and  $B$ , respectively. For example, consider Eq. (14.7). The profit-maximising output of firm  $A$  depends on the value of  $Q_B$ , i.e., the output which firm  $B$  is assumed to produce. If firm  $B$  chooses to produce 30 units, (i.e.,  $Q_B = 30$ ), then  $A$ 's output = 30 [=  $(90 - 30)/2$ ]. If firm  $B$  chooses to produce 60 units,  $A$ 's output = 15 [=  $(90 - 60)/2$ ]. Thus, Eq. (14.7) is the reaction function of firm  $A$ . It can be similarly shown that Eq. (14.8) is the reaction function of firm  $B$ .

### Criticism

Although Cournot's model yields a stable equilibrium, it has been criticised on the following grounds:

**First**, Cournot's behavioural assumption [assumption (d) above] is naive to the extent that it implies that firms continue to make wrong calculations about the competitor's behaviour. That is, each seller continues to assume that his rival will not change his output even though he repeatedly observes that his rival firm does change its output.

**Second**, his assumption of zero cost of production is unrealistic though dropping this assumption does not alter his model.



### Bertrand's Duopoly Model

Bertrand, a French mathematician, criticised Cournot's model and developed his own model of duopoly in 1883. Bertrand's model differs from Cournot's model in respect of its behavioural assumption. While under Cournot's model, each seller assumes his rival's output to remain constant, under Bertrand's model each seller determines his price on the assumption that his rival's price, rather than his output, remains constant.

Bertrand's model concentrates on price-competition. His analytical tools are reaction functions of the duopolists. Reaction functions of the duopolists are derived on the basis of iso-profit curves. An iso-profit curve, for a given level of profit, is drawn on the basis of various combinations of prices charged by rival firms. Assuming two firms A and B, the two axis of the plane on which iso-profit curves are drawn measure one each the prices of the two firms. Iso-profit curves of the two firms are convex to their respective price axis, as shown in Figs. 14.9. and 14.10. Iso-profit curves of firm A are convex to its price-axis  $P_A$  (Fig. 14.9) and those of firm B are convex to  $P_B$  (Fig. 14.10).

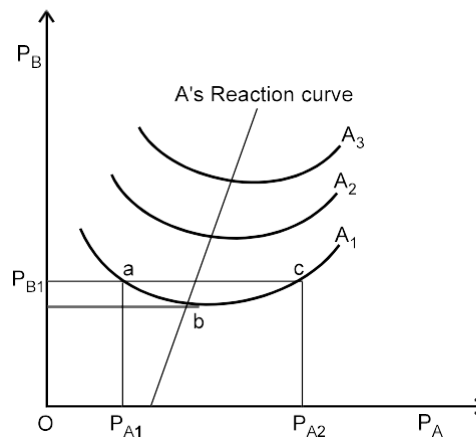


Fig. 14.9 A's Reaction Curve

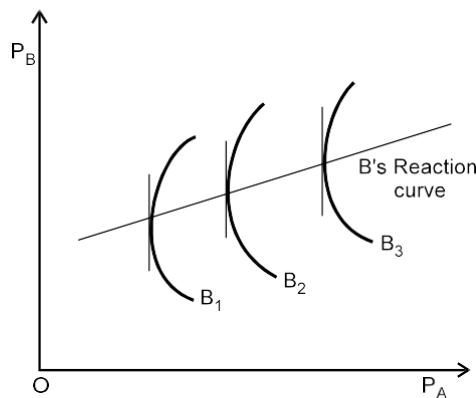


Fig. 14.10 B's Reaction Curve

To explain the implication of an iso-profit curve, consider curve  $A$  in Fig. 14.9. It shows that  $A$  can earn a given profit from the various combinations of its own and its rival's price. For example, price combinations at points  $a$ ,  $b$  and  $c$  on iso-profit curve  $A_1$ , yield the same level of profit. If firm  $B$  fixes its price  $P_{B1}$ , firm  $A$

has two alternative prices,  $P_1$  and  $P_2$ , to make the same level of profits. When  $B$

$A_1$   $A_2$

reduces its price,  $A$  may either raise its price or reduce it.  $A$  will reduce its price when he is at point  $c$  and raise its price when he is at point  $a$ . But there is a limit to which this price adjustment is possible. This point is given by point  $b$ . So there is a unique price for  $A$  to maximise its profits. This unique price lies at the lowest point of the iso-profit curve. The same analysis applies to all other iso-profit curves. If we join the lowest points of the iso-profit curves  $A_1$ ,  $A_2$  and  $A_3$ , we get  $A$ 's reaction curve. Note that  $A$ 's reaction curve has a rightward slant. This is so because, iso-profit curve tend to shift rightward when  $A$  gains market from its rival  $B$ .

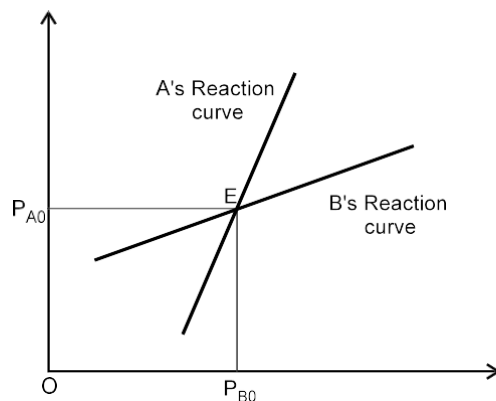


Fig. 14.11 Duopoly Equilibrium: Bertrand's Model

Following the same process,  $B$ 's reaction curve may be drawn as shown in Fig. 14.10. The equilibrium of duopolists suggested by Bertrand's model may be obtained by putting together the reaction curves of the firms  $A$  and  $B$  as shown in Fig. 14.11. The reaction curves of  $A$  and  $B$  intersect at point  $E$  where their expectations materialise. Point  $E$  is therefore equilibrium point. This equilibrium is stable. For, if anyone of the firms deviates from the equilibrium point, it will generate a series of actions and reactions between the firms which will lead them back to point  $E$ .

### Criticism

Bertrand's model has however been criticised on the same grounds as Cournot's model. Bertrand's implicit behavioural assumption that firms never learn from their past experience is naive. Furthermore, if cost is assumed to be zero, price will fluctuate between zero and the upper limit of the price, instead of stabilising at a point.

### Edgeworth's Duopoly Model

Edgeworth developed his model of duopoly in 1897. Edgeworth's model follows Bertrand's assumption that each seller assumes his rival's price, instead of his output, to remain constant. His model is illustrated in Fig. 14.12.

Let us suppose that there are two sellers,  $A$  and  $B$ , in the market. The entire market  $MM$  in Fig. 14.12 is equally divided between the two sellers who face identical demand curves.  $A$  has his demand curve as  $D_A$  and  $B$  as  $D_B$ . Let us also assume that seller  $A$  has a maximum capacity of output  $OM$  and  $B$  has a maximum output capacity of  $OM$ . The ordinate  $OD$  measures the price.

To begin the analysis of Edgeworth's model, let us suppose that  $A$  is the only seller in the market. Following the profit-maximising rule of a monopolyseller, he sells  $OQ$  and charges a price,  $OP_2$ . His monopoly profit, under zero cost, equals  $OP_2EQ$ . Now,  $B$  enters the market and assumes that  $A$  will not change his price since he is making maximum profit. With this assumption,  $B$  sets his price slightly below  $A$ 's price ( $OP_2$ ) and is able to sell his total output and also to capture a substantial position of  $A$ 's market.

Seller  $A$  now realises the reduction in his sale. In order to regain his market,  $A$  sets his price slightly below  $B$ 's price. This leads to price-war between the sellers. The price-war takes the form of price-cutting which continues until price reaches  $OP_1$ . At this price both  $A$  and  $B$  are able to sell their entire output  $A$  sells  $OM$  and  $B$  sells  $OM$ . The price  $OP_1$  could therefore be expected to be stable. But, according to Edgeworth, price  $OP_1$  should not be stable.

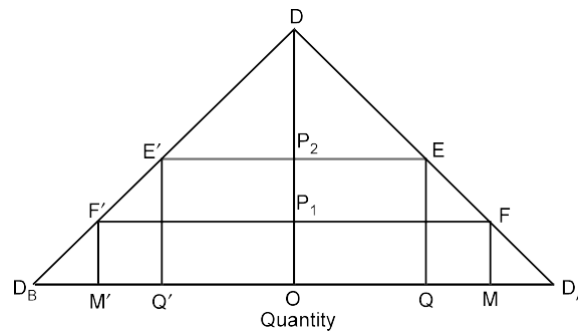


Fig. 14.12 Edgeworth's Model of Duopoly

The reason is that, once price  $OP_1$  is set in the market, the sellers observe an interesting fact. That is, each seller realises that his rival is selling his entire output and he will therefore not change his price, and each seller thinks that he can raise his price to  $OP_2$  and can make pure profit. This realisation forms the basis of their action and reaction. For examples, let seller  $A$  take the initiative and raise his price to  $OP_2$ . Assuming  $A$  to retain his price  $OP_2$ ,  $B$  finds that if he raises his price to a level slightly below  $OP_2$ , he can sell his entire output at a higher price and make greater profits. Therefore,  $B$  raises his price according to his plan.

Now it is  $A$ 's turn to appraise the situation and react.  $A$  finds that his price is higher than  $B$ 's price. His total sale falls. Therefore, assuming  $B$  to retain his price,  $A$  reduces his price slightly below  $B$ 's price. Thus, the price-war between  $A$  and  $B$  begins once again. This process continues indefinitely and price keeps moving up and down between  $OP_1$  and  $OP_2$ . Obviously, **according to Edgeworth's model of duopoly, equilibrium is unstable and indeterminate** since price and output

are never determined. In the words of Edgeworth, there will be an indeterminate tract through which the index of value will oscillate, or, rather will vibrate irregularly for an indefinite length of time.”

Edgeworth’s model, like Cournot’s and Bertrand’s model is based on a naïve assumption, i.e., each seller continues to assume that his rival will never change his price or output even though they are proved repeatedly wrong. But, Hotelling remarked that Edgeworth’s model is definitely an improvement upon Cournot’s model in that it assumes price, rather than output, to be the relevant decision variable for the sellers.

### *Stackelberg’s Leadership Model*

Stackelberg, a German economist, developed, his leadership model of duopoly in 1930. His model is an extension of Cournot’s model. Stackelberg assumes that one of the duopolists (say  $A$ ) is sophisticated enough to play the role of a leader and the other (say  $B$ ) acts as a follower. The leading duopolist  $A$  recognises that his rival firm  $B$  has a definite reaction function which  $A$  uses into his own profit function and maximises his profits.

Suppose market demand function is given as in (14.1), i.e.,  $Q = 90 - P$  and  $B$ ’s reaction function is given as in Eq. (14.9), i.e.,

$$Q_B = \frac{90 - Q_A}{2} \quad \dots(14.9)$$

Now, let  $A$  incorporate  $B$ ’s reaction function into the market function and formulate his own demand function as

$$Q_A = 90 - Q_B - P \quad \dots(14.10)$$

Since  $Q_B = (90 - Q_A)/2$ , Eq. (14.10) may be written as

$$Q_A = 90 - \frac{90 - Q_A}{2} - P$$

$$\text{or } Q_A = 45 + \frac{Q_A}{2} - P$$

$$\text{or } 2Q_A = 90 + Q_A - 2P \quad \dots(14.11)$$

$$Q_A = 90 - 2P$$

Thus, by knowing  $B$ ’s reaction function,  $A$  is able to determine his own demand function. Following the profit-maximisation rule,  $A$  will fix his output at 45 units ( $= 90/2$ ), i.e., half of the total demand at zero price.

Now, if seller  $A$  produces 45 units and seller  $B$  sticks to his own reaction function, he will produce

$$Q_B = \frac{90 - 45}{2} = 22.5 \text{ units} \quad \dots(14.12)$$

Thus, the industry output will be

$$45 + 22.5 = 67.5.$$

The problem with Stackelberg's model is that it does not decide as to which of the firms will act as leader (or follower). If each firm assumes itself to be the leader and the other to be the follower then Stackelberg's model will be indeterminate with unstable equilibrium.

---

## **14.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS**

---

1. One implication of Sweezy's Model is that since elasticity of the demand curve below point P is assumed to be less than unity and MR beyond a point is negative, the conditions of short-run equilibrium are not precise. That is, profit maximisation rule,  $MC=MR$ , cannot be applied to the short-run conditions.
2. A cartel is a formal organisation of the oligopoly firms in an industry. Cartels are the perfect form of collusion.
3. When there are only two sellers of a product, there exists duopoly, a special case of oligopoly.
4. Augustin Cournot, a French economist, was the first to develop a formal duopoly model in 1838.

---

## **14.5 SUMMARY**

---

- Oligopoly is a form of market structure in which a few sellers sell differentiated or homogeneous products.
- The products traded by the oligopolists may be differentiated or homogeneous. Accordingly, the market may be characterised by heterogeneous oligopoly or homogeneous (or pure) oligopoly.
- The nature and degree of competition among the oligopolists makes them interdependent in respect of decision-making.
- An oligopolistic market structure is also characterised, in the long run, by strong barriers to entry of new firms to the industry. If entry is free, new firms attracted by the super-normal profits, if it exists, enter the industry and the market eventually becomes competitive.
- The classical models of duopoly assumed independent action by the rival firms in their attempt to maximise their profits. Chamberlin rejected the

assumption of independent action by the competing firms. He developed his own model of oligopoly assuming interdependence between the competitors.

- In the non-collusive models, oligopoly firms are assumed to act independently. In the collusive models, however, firms are assumed to act in unison, i.e., in collusion with one another.
- A cartel is a formal organisation of the oligopoly firms in an industry. Cartels are the perfect form of collusion. A general purpose of cartels is to centralise certain managerial decisions and functions of individual firm in the industry with a view to promoting common benefits. Cartels may be in the form of open collusion or secret collusion.
- The central management board is provided with cost figures of individual firms. Besides, it is supposed to obtain the necessary data required to formulate the market demand (AR) curve.
- Collusion through price leadership is another form of collusion between oligopoly firms. Price leadership is an informal position of a firm in an oligopolistic setting to lead other firm in fixing price of their product.
- The price leadership is possible under both product homogeneity and product differentiation or heterogeneity.
- The price leadership model, as presented above, yields a stable solution to the problem of oligopoly pricing and output determination, only if small firms faithfully follow the leader.
- Another form of price leadership is barometric price leadership. In this form of price leadership, a firm initiates well publicised changes in price that are generally followed by the rival firms in the industry.
- When there are only two sellers of a product, there exists duopoly, a special case of oligopoly. Duopoly is a special case in the sense that it is the limiting case of oligopoly as there must be at least two sellers to make the market oligopolistic in nature. In this section, we explain famous 'classical models' of duopoly.
- Bertrand, a French mathematician, criticised Cournot's model and developed his own model of duopoly in 1883. Bertrand's model differs from Cournot's model in respect of its behavioural assumption.
- Edgeworth developed his model of duopoly in 1897. Edgeworth's model follows Bertrand's assumption that each seller assumes his rival's price, instead of his output, to remain constant.

---

## 14.6 KEY WORDS

---

- **Assumption:** It refers to something that is accepted as true or as certain to happen, without proof.

- **Criticism:** It refers to the analysis and judgement of the merits and faults of a literary or artistic work.
- **Profit:** It is a financial gain, especially the difference between the amount earned and the amount spent in buying, operating, or producing something.

---

## 14.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

---

### *Short-Answer Questions*

1. State the characteristics of oligopoly.
2. Write a short note on price determination under non-collusive oligopoly.
3. What are the implications of Sweezy's model?

### *Long-Answer Questions*

1. Give a detailed overview of the price determination under collusive models of oligopoly.
2. Describe the price leadership models of oligopoly.
3. Explain Cournot's Duopoly Model in detail. Why was it criticized?
4. How is Bertrand's Duopoly Model different from Edgeworth's Duopoly Model?

---

## 14.8 FURTHER READINGS

---

Dwivedi, D. N. 2002. *Managerial Economics*, 6th Edition. New Delhi: Vikas Publishing House.

Keat Paul G. and K.Y. Philip. 2003. *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th Edition. Singapore: Pearson Education Inc.

Peterson, H. C. and W. C. Lewis. 1999. *Managerial Economics*, 4th Edition.

Singapore: Pearson Education, Inc.

Dwivedi, D. N. 2011. *Principles of Economics*, Second Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Tewari, D. D. 2003. *Principles of Microeconomics*. New Delhi: New Age International (P) Ltd., Published

