INSTITUTE OF DISTANCE EDUCATION IDE
Rajiv Gandhi University

BAECO 101
ECONOMIC THEORY-I

BA (ECONOMICS)
1ST SEMESTER

Rajiv Gandhi University
www.ide.rgu.ac.in

ECONOMIC THEORY - I

BA [Economics]

First Semester

BAECO-101

RAJIV GANDHI UNIVERSITY

Arunachal Pradesh, INDIA - 791 112

BOARD OF STUDIES	
1. **Prof. K C Kapoor** Department of Education Rajiv Gandhi University	**Chairman**
2. **Prof. J C Soni** Department of Education Rajiv Gandhi University	**Member**
3. **Dr. P K Acharya** Department of Education Rajiv Gndhi University	**Member**
4. **Ms. Moyir Riba** Institute of Distance Education Rajiv Gandhi University	**Member**
5. **Dr. Ashan Riddi** Director, IDE	**Member Secretary**

Authors

D.N Dwived
M.C Vaish

Vikas revised edition, 2021

About the University

Rajiv Gandhi University (formerly Arunachal University) is a premier institution for higher education in the state of Arunachal Pradesh and has completed twenty-five years of its existence. Late Smt. Indira Gandhi, the then Prime Minister of India, laid the foundation stone of the university on 4th February, 1984 at Rono Hills, where the present campus is located.

Ever since its inception, the university has been trying to achieve excellence and fulfill the objectives as envisaged in the University Act. The university received academic recognition under Section 2(f) from the University Grants Commission on 28th March, 1985 and started functioning from 1st April, 1985. It got financial recognition under section 12-B of the UGC on 25th March, 1994. Since then Rajiv Gandhi University, (then Arunachal University) has carved a niche for itself in the educational scenario of the country following its selection as a University with potential for excellence by a high-level expert committee of the University Grants Commission from among universities in India.

The University was converted into a Central University with effect from 9th April, 2007 as per notification of the Ministry of Human Resource Development, Government of India.

The University is located atop Rono Hills on a picturesque tableland of 302 acres overlooking the river Dikrong. It is 6.5 km from the National Highway 52-A and 25 km from Itanagar, the State capital. The campus is linked with the National Highway by the Dikrong bridge.

The teaching and research programmes of the University are designed with a view to play a positive role in the socio-economic and cultural development of the State. The University offers Undergraduate, Post-graduate, M.Phil and Ph.D. programmes. The Department of Education also offers the B.Ed. programme.

There are fifteen colleges affiliated to the University. The University has been extending educational facilities to students from the neighbouring states, particularly Assam. The strength of students in different departments of the University and in affiliated colleges has been steadily increasing.

The faculty members have been actively engaged in research activities with financial support from UGC and other funding agencies. Since inception, a number of proposals on research projects have been sanctioned by various funding agencies to the University. Various departments have organized numerous seminars, workshops and conferences. Many faculty members have participated in national and international conferences and seminars held within the country and abroad. Eminent scholars and distinguished personalities have visited the University and delivered lectures on various disciplines.

The academic year 2000-2001 was a year of consolidation for the University. The switch over from the annual to the semester system took off smoothly and the performance of the students registered a marked improvement. Various syllabi designed by Boards of Post-graduate Studies (BPGS) have been implemented. VSAT facility installed by the ERNET India, New Delhi under the UGC-Infonet program, provides Internet access.

In spite of infrastructural constraints, the University has been maintaining its academic excellence. The University has strictly adhered to the academic calendar, conducted the examinations and declared the results on time. The students from the University have found placements not only in State and Central Government Services, but also in various institutions, industries and organizations. Many students have emerged successful in the National Eligibility Test (NET).

Since inception, the University has made significant progress in teaching, research, innovations in curriculum development and developing infrastructure.

SYLLABI-BOOK MAPPING TABLE

Economic Theory - I

Syllabi	Mapping in Book
Unit I: Basic Economic Issues Resource Scarcity, Unlimited Wants, Choice, Opportunity Cost, Economic Problems of Developing Countries, Low Income, Resource Constraints, Low Level of Technology, Low Organizational Ability, Low Degree of Inventiveness and Innovativeness, Low Level of Human and Physical Capital, Acute Poverty and Inequality in the Distribution of Income and Opportunities,	**Unit I:** Basic Economic Issues
Unit II: Economics Economics, Micro and Macro - their Differences and Subject Matters.	**Unit II: Economics**
Unit III: Demand Analysis Basis of demand: utility and income; diminishing marginal utility, income of the consumer and her budget line, constrained utility maximization; demand curve and factors shifting it: income, prices of related goods, etc.; Elasticity of demand: price and income elasticity.	**Unit III: Demand Analysis**
Unit IV: Supply Analysis Production Function, Returns to a Factor and Returns to Scale, Marginal and Average Product of Inputs, Short-run Total, Marginal and Average Cost Curves and their Relationship, Cost Minimization, Total, Average and Marginal Revenue, Profit Maximizing Output, Supply Curve, Shifts in Supply Curve, Elasticity of Supply.	**Unit IV: Supply Analysis**

CONTENTS

UNIT IV SUPPLY ANALYSIS

INTRODUCTION

Economic theory or analysis furnishes the economists with a set of tools which they use to analyse the nature of the observed economic phenomena in the real world. Economic theory may, therefore, be appropriately defined as a 'box of tools' with which the economists construct economic models in order to study the economic phenomena which frequently occur in the real world. Although the analytical tools in the kit-bag of an economist are inadequate to enable him to handle each and every individual economic occurrence in so far as it reveals its own peculiarities, these nevertheless enable him to analyse certain common features of individual economic occurrences. Like other sciences, economic theory also provides us with the general propositions which are employed in the analysis of economic phenomena within certain limits. The limitations of these theoretical economic propositions in analysing the individual economic phenomenon emanate from the assumptions which form the basis of these propositions. Since the assumptions forming the bedrock of economic theory are very seldom realistic, economic theory resembling actual reality is a rare occurrence. To the extent that all economic theories are based on certain assumptions, these theories abstract from reality. The more general or universal is the economic theory the greater is its abstraction from reality.

Formally tracing its origin to Adam Smith's mounmental work entitled *An Inquiry into the Nature and Causes of the Wealth of Nations*, first published in 1776, economic theory can today take pride in calling itself more than two centuries and three decades old. Its development during this long period, however, has failed in following any set pattern, being seldom in the same direction. Economic theory, both past and contemporary, is the product of numerous influences and factors affecting one another. The philosophical thought of a particular period or of a particular writer has also influenced the kind of theory which has developed over this period. Its development has also been influenced by the political biases of writers over the long period of its history. For example, the classical economic theory was influenced, in no small measure, by the political biases of the classical economists. Similarly, the Marxian economic theory was couched in Karl Marx's political philosophy.

It is doubtful to say if David Ricardo would have developed his theory of international trade without a strong animus against the landed class. The theory, however, survives the removal of his prejudices. The development of economic theory has taken place over several periods with each period marked by certain special features not found in the other periods. Consequently, economic theory does not belong to any single individual, country or age. Obviously, its outlook and ownership is essentially cosmopolitan.

There is no unanimity among economists about the nature and purpose of economic theory. Should economic theory accurately describe its assumptions? Or, should it predict actual future events? Or, should it predict consequences of certain causes in an 'ideal' world? The principal function of an economic theory is to explain the nature of economic activity and to predict as to what will happen in the economy at a given time in future.

A perfect theory, besides being realistic in its presentation, should be competent to predict the consequences of certain given events. For instance, assuming that the producer's objective function is only to maximize their profits, given the data about the supply and demand functions and input prices, it should be logically possible to deduce the total amount of the commodity output which producers will produce and the total amount of net profit earned by them. If the government now imposes an *ad valorem* commodity tax on the producers, we can find out its impact on the output, price and profit. By its very nature, however, economic theory cannot always be descriptively realistic. The purpose of economic theory is to develop hypotheses which are abstract from the essential features of the complex real world. Economic theory should formulate questions pertaining to an economic phenomenon. It should also indicate the mode of answering these questions. This book *Economic Theory*, will deal with the various aspects of economic theory.

The book, *Economic Theory*, is written in a self-instructional format and is divided into six units. Each unit begins with an *Introduction* to the topic followed by an outline of the *Unit objectives*. The content is then presented in a simple and easy-to-understand manner, and is interspersed with *Check Your Progress* questions to test the reader's understanding of the topic. A list of *Questions and Exercises* is also provided at the end of each unit, and includes short-answer as well as long-answer questions. The *Summary* and *Key Terms* section are useful tools for students and are meant for effective recapitulation of the text.

NOTES

UNIT 1 BASIC ECONOMIC ISSUES

Structure

1.0 INTRODUCTION

Any activity that produces goods and services is productive activity and any activity that creates goods and services of value is called *economic activity.* The basic objective behind all economic activities is to make income, the source of livelihood. An *important feature* of economic activities is that they are *interrelated* and *interdependent* in the sense that producers produce what consumers want to consume and consumers can consume only what producers produce and they produce only as much as consumers are willing to consume. Similarly, sellers can sell only what buyers are willing to buy and buyers can buy only what is offered for sale; and so on. This interrelatedness and interdependence of economic activities are carried out in a self-operated system.

An economy is a social organism in which people act, interact, cooperate and compete in the process of production and consumption to make their living. An economy is constituted of interrelated and interdependent economic activities of the *economic players.* Economic players include individuals, households, firms, farms, factories, financial institutions and government. All kinds of economic activities are carried out within the framework of an economic system. A free economic system is established and governed by two economic forces—demand for and supply of goods and services. Demand and supply forces create a market system—called market mechanism. The interaction between the market forces of demand and supply makes the economic system of the country. A clear understanding of the economic system and its working is a necessary condition for making appropriate business decisions.

NOTES

This unit will introduce you to the basic problems an economy is always faced with. The economic problem is one of the basic economic theoretical principles being employed in the operation of any economy. The economic problem model asserts that there is resource scarcity, i.e., available resources are not sufficient to satisfy our all wants and needs. Three questions arise from this: first, what to produce; second, how the factors of production, namely capital and labour, are to be allocated to produce it; and third, for whom those goods or service should produce (a problem of allocation of resources). Economics revolves around methods and possibilities of solving this fundamental economic problem.

Developing countries have a unique set of economic problems and challenges to economic development. Economic problems of these countries include rural poverty, low income, inequality in the distribution of income and opportunities, resource constraints and low levels of technology, innovativeness, and human and physical capitals. You will learn about all these problems in this unit.

1.1 UNIT OBJECTIVES

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

- Discuss the major microeconomic and macroeconomic problems faced by an economy
- Assess the production possibilities and opportunity cost of an economy
- Describe the various economic problems of developing countries
- Discuss the major differences between macro and micro economics

1.2 ECONOMY AND ITS BASIC PROBLEMS: RESOURCE SCARCITY AND UNLIMITED WANTS

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.

1.2.1 Why Does the Problem of Making Choice Arise?

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. Unlimited human wants, desires and aspirations

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

NOTES

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 consumers 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 scarcity

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 thing3 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)

(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** (discussed ahead separately in the unit). 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. Gain maximizing attitude

NOTES

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.

1.2.2 Micro and Macroeconomic Problems

The basic problems of an economy lie in the background of all economic decisions, and also form the basis of economic studies and generalization. The major economic problems faced by an economy—whether capitalist, socialist or mixed—may be classified into two broad groups:

(i) Microeconomic problems which are related to the working of the economic system

(ii) 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 subsection.

Microeconomic problems

The basic microeconomic problems are:

- What to produce and how much to produce?
- How to produce?
- 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.

(i) What to Produce?: Problem of Choice between Commodities

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

NOTES

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 the output is maximum and the output-mix is optimum. The objective is to satisfy the maximum needs of the 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 makes 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?: Problem of Choice of Technique

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 the case of some commodities, however, choices are limited. For example, production of woollen carpets and other items of handicrafts is by nature labour-intensive, while production of cars, TV sets, computers, aircraft, etc., is capital-intensive. In the 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 the solution to the problem of choice of technology through the 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 maximizing firms find out an input combination which minimizes their cost of production. This becomes inevitable for the firms because their resources are limited and, with given resources, they intend to maximize their profits.

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

In a modern economy, all the goods and services are produced by business firms. The total output generated by 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 the 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 arise: why do some people have a command over a larger proportion of resources than the others? Why do those who have more, get more and more? Why do those who have less, get less and less? In other words, why do the rich get richer and the 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?

The 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 as 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 the welfare of the society? How can production and distribution be made more efficient or welfare oriented? Economists' attempts to answer these questions has led to the 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 up the subject matter of Microeconomic Theory or 'Price Theory'. Apart from microeconomic problems, there are certain macroeconomic problems of prime importance confronted by an economy. These problems may be specified as follows:

1. **How to increase the production capacity of the economy:** This is essentially the problem related to the economic growth of the country. The need for increasing the production capacity of the economy arises for at least two reasons. First, most economies of the world have realized by experience that their population has grown at a rate much higher than their productive resources. This leads to poverty, especially in the less-developed countries. Poverty, in itself, is a cause of a number of socio-economic problems. Besides, it has frequently jeopardized the sovereignty and integrity of nations. Colonization of poor nations by the richer and powerful imperialist nations during the pre-twentieth century period is evidence of this fact. Therefore, growth of the economy and sparing resources for defence has become a necessity. Second, 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 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 do some economies grow faster than the others? This has led to the Theories of Economic Growth.

NOTES

2. **How to stabilize 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 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, 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 problems of this category are the problems of unemployment and inflation. While widespread unemployment is the biggest problem confronting developing economies, inflation is a global problem. Another set of macro problems is associated with international trade. 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 an economy passes through different phases of economic growth.

1.2.3 Production Possibilities and Opportunity Cost

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 with 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 the 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 sub-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.

Apart from showing the possible alternative combinations of two goods, production possibilities frontier (PPF) also indicates the opportunity cost of one commodity in terms of the other product. Conceptually, opportunity cost is the benefit 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 *PPF*.As can be seen in Fig. 1.1, the movement along the production possibilities frontier, *AF*, shows

a 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 7000 and tons to 6000 tons and increase in the production of clothing from 40 million metres to 55 million metres. It implies that 1000 tons of food can be produced only by sacrificing 55 million metres of clothing. It means that opportunity cost of 1000 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.

NOTES

Increasing Opportunity Cost and Concavity of PPF

The production possibilities frontier reveals another important fact that opportunity cost changes along the *PPF*. In Fig. 1.1, 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 1000 tons for 15 million metres of clothing. It means that the opportunity cost of 15 million metres of clothing is 1000 tons of food. Amovement from point *B* to *C* shows that the opportunity cost of only 9 million metres of clothing, a much lower quantity, is the same 1000 tons of food. It means that opportunity cost of clothing increases as we move downwards along the *PPF*.

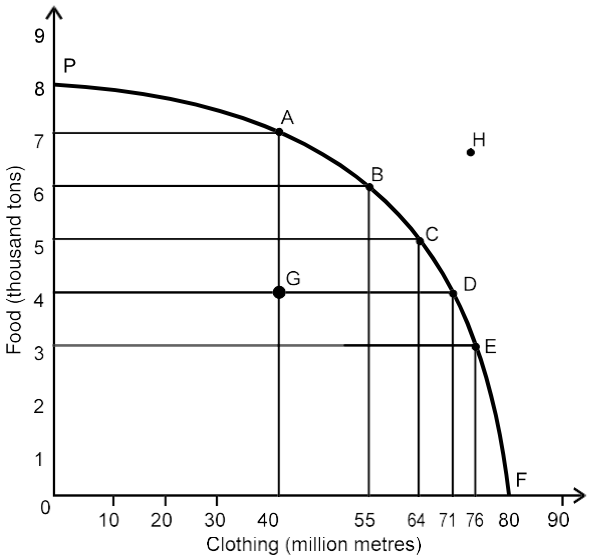


Fig. 1.1 *The Production Possibilities Frontier*

Why is PPF Concave? It can be seen from Fig. 1.1 that *PPF* takes the form of a concave curve. *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* curve 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.

Check Your Progress

1. How can the major economic problems faced by an economy be classified?
2. Why does the problem 'what to produce' arise?
3. Why cannot societies have all that they want?

1.3 ECONOMIC PROBLEMS OF DEVELOPING COUNTRIES

NOTES

Developing areas, including developing countries and regions, have a unique set of economic problems and challenges to economic development. Developing countries, taken as whole, refer to countries characterized by an underdeveloped industrial base, low per capita income, and widespread poverty. Some of the major economic problems of developing countries are described in this section.

1.3.1 Rural Poverty in Developing Countries

The causes of rural poverty are complex and multidimensional. They involve, among other things, culture, climate, gender, markets, and public policy. Likewise, the rural poor are quite diverse both in the problems they face and the possible solutions to these problems.

Broad economic stability, competitive markets, and public investment in physical and social infrastructure are widely recognized as important requirements for achieving sustained economic growth and a reduction in rural poverty. In addition, because the rural poor's links to the economy vary considerably, public policy should focus on issues such as their access to land and credit, education and health care, support services, and entitlements to food through well-designed public works programmes and other transfer mechanisms.

About one-fifth of the world's population is afflicted by poverty—these people live on less than $1 a day. Poverty is not only a state of existence but also a process with many dimensions and complexities. Poverty can be persistent (chronic) or transient, but transient poverty, if acute, can trap succeeding generations. The poor adopt all kinds of strategies to mitigate and cope with their poverty.

To understand poverty, it is essential to examine the economic and social context, including institutions of the state, markets, communities, and households. Poverty differences cut across gender, ethnicity, age, location (rural versus urban), and income source. In households, children and women often suffer more than men. In the community, minority ethnic or religious groups suffer more than majority groups, and the rural poor more than the urban poor; among the rural poor, landless wage workers suffer more than small landowners or tenants. These differences among the poor reflect highly complex interactions of cultures, markets, and public policies.

Rural poverty accounts for nearly 63 per cent of poverty worldwide, reaching 90 per cent in some countries like Bangladesh and between 65 and 90 per cent in sub-Saharan Africa. (Exceptions to this pattern are several Latin American countries in which poverty is concentrated in urban areas.) In almost all countries, the conditions—in terms of personal consumption and access to education, health care, potable water and sanitation, housing, transport, and communications—faced by the rural poor are far worse than those faced by the urban poor. Persistently high levels of rural poverty, with or without overall economic growth, have contributed to rapid population growth and migration to urban areas. In fact, much urban poverty is created by the rural poor's efforts to get out of poverty by moving to cities. Distorted government policies, such as penalizing the agriculture sector and neglecting rural (social and physical) infrastructure, have been major contributors to both rural and urban poverty.

1.3.2 Low Income

NOTES

The links between poverty, economic growth, and income distribution have been studied quite extensively in recent literature on economic development. Absolute poverty can be alleviated if at least two conditions are met:

- Economic growth must occur—or mean income must rise—on a sustained basis
- Economic growth must be neutral with respect to income distribution or reduce income inequality

Generally, poverty cannot be reduced if economic growth does not occur. In fact, the persistent poverty of a substantial portion of the population can dampen the prospects for economic growth. Also, the initial distribution of income (and wealth) can greatly affect the prospects for growth and alleviation of mass poverty. Substantial evidence suggests that a highly unequal distribution of income is not conducive to either economic growth or poverty reduction. Experience has shown that if countries put in place incentive structures and complementary investments to ensure that better health and education lead to higher incomes, the poor will benefit doubly through increased current consumption and higher future incomes.

The pattern and stability of economic growth also matter. On the one hand, traditional capital-intensive, import-substituting, and urban-biased growth—induced by government policies on pricing, trade, and public expenditure—has generally not helped alleviate poverty. On the other hand, agricultural growth—where there is a low concentration of land ownership and labour-intensive technologies are used—has almost always helped reduce poverty. Finally, sharp drops in economic growth—resulting from shocks and economic adjustments—may increase the incidence of poverty. Even when growth resumes, the incidence of poverty may not improve if inequality has been worsened by the crisis.

Low-Income Developing Countries (LIDCs)

The Low-Income Developing Countries (LIDC) group includes all countries that: (a) fall below a modest per capita income threshold (US$2,500 in 2011, based on Gross National Income) and (b) are not conventionally viewed as emerging market economies (EMs). There are 60 countries in this group, accounting for about one-fifth of the world's population; sub-Saharan Africa (SSA) accounts for some 57 per cent of the LIDC population, with a further 28 per cent living in Asia. While sharing characteristics common to all countries at low levels of economic development, the LIDC group is strikingly diverse, with countries ranging in size from oil-rich Nigeria (174 million) to fisheries dependent Kiribati (0.1 million), and in 2013 per capita GDP terms from Mongolia (US$3,770) to Malawi (US$270). The 10 largest economies in the group account for two-thirds of total group output.

1.3.3 Inequality in the Distribution of Income and Opportunities

Inequality in society is not a new phenomenon. And yet it can be fatal. If left unchecked, it can undermine the very foundations of development and social and domestic peace.

Over the last decades, the world has witnessed impressive average gains against multiple indicators of material prosperity. For instance, gross domestic product (GDP) per capita in low- and middle-income countries has more than doubled in real terms since 1990. In the same period, life expectancy in developing countries has risen from 63.2 years to 68.6 years. However, this is only part of the picture. Although the world is

NOTES

globally richer than ever before, more than 1.2 billion people still live in extreme poverty. The richest 1 per cent of the world population owns about 40 per cent of the world's assets, while the bottom half owns no more than 1 per cent. Despite overall declines in maternal mortality, women in rural areas are still up to three times more likely to die while giving birth than women living in urban centres. Social protection has been extended, yet persons with disabilities are up to five times more likely than average to incur catastrophic health expenditures. Women are participating more in the work force, but continue to be disproportionately represented in vulnerable employment. Humanity remains deeply divided.

Nor are recent trends very encouraging. Over the last two decades, income inequality has been growing on average within and across countries. As a result, a significant majority of the world's population lives in societies that are more unequal today than 20 years ago. Remarkably, in many parts of the world, income gaps have deepened—and, with them, the gulf in quality of life between the rich and the poor—despite the immense wealth created through impressive growth performances. In fact, the sharpest increases in income inequality have occurred in those developing countries that were especially successful in pursuing vigorous growth and managed, as a result, to graduate into higher income brackets. Economic progress in these countries has not alleviated disparities, but rather exacerbated them.

The world is more unequal today than at any point since Second World War. However, there are clear signs that this situation cannot be sustained for much longer. Inequality has been jeopardizing economic growth and poverty reduction. It has been stalling progress in education, health and nutrition for large swathes of the population, thus undermining the very human capabilities necessary for achieving a good life. It has been limiting opportunities and access to economic, social and political resources. Furthermore, inequality has been driving conflict and destabilizing society. When incomes and opportunities rise for only a few, when inequalities persist over time and space and across generations, then those at the margins, who remain so consistently excluded from the gains of development, will at some point contest the 'progress' that has bypassed them. Growing deprivations in the midst of plenty and extreme differences between households are almost certain to unravel the fabric that keeps society together. This is especially problematic when we consider that, often, it is precisely those at the margins who tend to pay the biggest price for social unrest. But perhaps most important, extreme inequality contradicts the most fundamental principles of social justice, starting from the notion, enshrined in the Universal Declaration of Human Rights that 'all human beings are born free and equal in dignity and rights'.

There is, however, some good news. There is nothing inevitable about high inequality. The widening of gaps in income, wealth or other dimensions of well-being is not an unavoidable price to pay for development. In fact, many countries over the last years have managed to significantly reduce income and non-income inequality through a combination of progressive economic and social policies, often accompanied by the greater participation and empowerment of those who had been left behind by the development process. Much can be learned from those experiences and applied to other contexts in which inequality continues to be a concern.

The drivers of excessive inequality are well known. Specific aspects of globalization, such as inadequately regulated financial integration and trade liberalization processes, whose benefits have been distributed very unequally across and within countries, have played a significant role in determining the upward trend observed over the last decades.

But domestic policy choices, such as interventions that weakened labour market institutions or resulted in a downsizing of public investments in critical sectors like health, education and social protection, have also played an important role. Often, various economic, social and cultural barriers hindering the political participation of various segments of the population have compounded these processes. In addition, discriminatory attitudes and policies that are marginalizing people on the basis of gender or other cultural constructs such as ethnicity or religious affiliation drive many intergroup inequalities.

The complexity and multi-dimensionality of the drivers of inequality call for a complex and multi-dimensional response. In fact, only a genuinely holistic approach can fully address the multiple factors that cause inequality and create the conditions for a truly inclusive society. Such an approach must shape growth so that market outcomes do not push households further apart, but deliver shared prosperity. But it must also address social and fiscal policy in ways that will allow governments to intervene to re-balance market outcomes through redistribution, when needed, and ensure universal access to critical services. It must strengthen democratic institutions so that there are mechanisms for broad-based participation in political and public life. And it must reverse discriminatory practices so that nobody is excluded because of who he or she is.

The world today is at a critical juncture. The financial and economic crises of recent years have pushed the international community to reconsider long-held views on economic priorities and social cohesion is much more widely recognized as a major factor contributing to resilience and sustainability. The debate on the future of development and international cooperation has started. In this context, inequality has emerged as a major issue of concern—not only among development specialists, but also well beyond. Furthermore, a host of civil society movements have explicitly and forcefully voiced this concern.

Millions of voices are asking the world's decision makers to confront rising inequalities. It is imperative that this demand be met if the ideals of a prosperous, peaceful and sustainable society are to be realized.

1.3.4 Resource Constraints

Basic food insecurity still affects 1 billion people, as many as in 1970. However, the proportion of people who are undernourished declined from about 20 per cent in 1990-1992 to 15 per cent in 2008-2010. Progress has been uneven across regions and the 2007-2008 food and financial crisis posed additional challenges. Under current conditions, the target of halving the proportion of people suffering from hunger by 2015 will not be met in sub-Saharan Africa and South Asia.

Because of low quality and low diversity of available food, the challenge of malnutrition is broader than the issue of hunger or undernourishment. Individuals may take in enough calories for daily subsistence, but still suffer from 'hidden hunger' with low levels of micronutrients owing to the lack of diversification of diets. This is a problem in both developing and developed countries, affecting 30 per cent of the world's population. The excess of calories is another rising major global public-health concern, as overweight and obesity result in more than 2.8 million deaths among adults every year.

Estimates indicate that food production will have to increase 70 per cent globally to feed an additional 2.3 billion people by 2050. Food demand is anticipated to continue to shift towards more resource-intensive agricultural products, such as livestock and dairy products, thereby exerting additional pressure on land, water and biodiversity resources.

NOTES

On the supply side, meeting an increasing food demand is a major concern, given the rise of resource constraints. Current agricultural practices are a leading source of greenhouse gas emissions, while also leading to other problems, such as loss of soil fertility and water pollution from run-off. Increased temperatures and more volatile weather patterns caused by climate change may already be affecting crop yields, affecting incomes and agricultural production.

Increased land use for biofuels will increase constraints on the supply side and may lead to higher food prices, further affecting the most economically disadvantaged. Similarly, current urbanization trends accelerate the diversion of land use from agricultural production.

Human Resource Constraint

The serious shortage of health workers across the world has been identified as one of the most critical constraints to the achievement of health and development goals. The crisis is impairing provision of essential, life-saving interventions such as childhood immunization, safe pregnancy and delivery services for mothers and access to prevention and treatment for HIV/AIDS, malaria and tuberculosis. Health workers are also critical to our preparedness for and response to the global security threats posed by emerging and epidemic-prone diseases such as SARS and avian flu and haemorrhagic fevers as well as the consequences of climate change. Without urgent action, the shortage will worsen, health systems will be weakened even further and health goals will not be achieved.

In its 2006 World Health Report, the World Health Organization estimated that over 4 million more health workers are needed to bridge the gap—with 1.5 million needed for Africa alone. Across the world, 57 countries have been identified as having 'critical shortages'—36 of these are in Africa.

The workforce crisis is made worse by imbalances within countries. There is a general lack of adequate staffing in rural areas compared to cities. To add further pressures, priority disease programmes are competing for scarce staff, to the detriment of integrated health system development. In developed countries, a rise in chronic health problems among ageing populations and ageing of their own workforces has led to an ever-growing demand for health workers. The pull of higher salaries in industrialized countries and the push of poor working conditions at home drive thousands of health workers to jobs abroad each year. Yet developing countries face an escalating double burden of both infectious and non-communicable diseases and are in need of massive scale up of training and retention interventions.

Unfavourable working conditions, widespread shortages and large scale migration of health workers are the challenges we face today. With new killer diseases and issues like climate change threatening global security, aging populations and changing work patterns, there is an ever-growing demand for health workers worldwide.

1.3.5 Low Level of Technology

Technological progress is at the heart of human progress and development. As the 1998 World Development Report on the knowledge economy (World Bank 1998) emphasized, the understanding of how things are created and the communication of that knowledge are critical drivers of economic progress. Central to understanding the role of technology is the recognition that technology and technological progress are relevant to a wide

NOTES

range of economic activities, not just manufacturing and computers. For example, some estimates suggest that technological progress has boosted productivity in agriculture four times as quickly as in manufacturing (Martin and Mitra 2001). Indeed, seemingly low-tech products such as corn or flowers can be the result of relatively high-tech production processes, while in some countries the production of ostensibly high-tech products such as computers is an outcome of relatively low-tech assembly activities. Finally, in many cases technology is embodied in production and management systems rather than in physical goods or software algorithms. A computer loaded with the latest software that sits unused on a desk for most of the day is a very different manifestation of technology than the same computer that is running a production process or managing an accounts payable system.

This defines technology and technological progress in this wider sense, although data limitations may give some of the measures developed the flavour of a more narrow, physical, and manufacturing-oriented definition.

Technology is both a critical determinant and an outcome of rising incomes

Traditionally, economists view the process by which goods and services are produced as one that combines capital, labour, and other factors of production (land and natural resources) using a particular technology. The relative efficiency with which a given economy produces goods and services given a certain quantity of labour and capital is called total factor productivity (TFP). TFP is commonly interpreted as a measure of the technology of production and its rate of growth as a measure of technical progress.

International comparisons of TFP suggest that enormous gaps exist between high income and low- and middle-income countries in the efficiency with which they produce goods and services. In 2005, the average level of TFP in low-income countries was only slightly more than 5 per cent of U.S. levels. The technology lower-middle income countries employed was roughly twice as efficient and that of upper-middle-income countries was approximately four times as efficient. While these gaps have been narrowing for low-income and lower-middle-income countries, upper-middle-income countries have only managed to maintain their relative position in relation to high-income countries. At the regional level, these gaps have widened or remained stagnant in three of six developing regions, with TFP growing faster in high income countries than in Latin America and the Caribbean, the Middle East, and Sub-Saharan Africa.

The relationships between income growth, technological progress, capital accumulation, and welfare are, of course, much more complex than can be summarized in a simple measure of TFP, partly because each factor of production and the technology with which factors are combined are dependent on one another. Capital goods often embody significant technological progress and there is no simple way to distinguish between the contribution that each makes to growth. Similarly, technology in the form of knowledge of business processes and of science and general experience is embodied in labour. Moreover, the contribution of technology to welfare is only imperfectly measured by its impact on GDP.

Improving the Flow of Technology in India

India is in a unique position to mount a strong initiative for affordable innovations for technologies for social and public good by taking advantages of: (a) low expertise costs, (b) vast talent base and, (c) the residual idealism in the society. However, engagement of multiple stakeholders and creating Public-Private-Partnership for promoting people-

NOTES

centric research is a challenge to address national goals with specific targets in a time bound manner. While technologies for public, strategic, and social goods would require collaborative excellence, competitive excellence models for private good would come from industrial sector, as is the case in most developed countries.

Food security of India is closely related to development of technologies for increasing the agriculture outputs through the process of innovations for land saving and water use. The question is how to develop and deploy new agro-biotechnology tools and precision agriculture for increasing the output of agricultural sector in the country by synergizing the strengths of institutions both under public and private sector and adopting a new approach for agriculture research and extension.

To achieve optimal health for its people, India has unique challenges due to its large population, demographic transition and vulnerability to all epidemics. Biomedical devices and instrumentation forms an area of serious gap in the country. Breakthrough innovations, with appropriate stress on translational research for affordable health care, are the need of the hour and would call for new models and mechanisms for evaluating technologies for improving healthcare at individual and public health level, fostering academia-industry linkage; and linking technology developers with industry for translation of lead products and processes.

Water challenge is a major national issue in the country both in terms of quality and availability. Sustainability of research led solutions depends on the interface of technology with policy and societal behaviour. Water related technologies form an ideal theme for building state-centre partnerships. The challenge, therefore, is to convert research outputs from the laboratories into revenue models based solutions in a coordinated manner among the relevant departments in both states and centre for innovative deployment under real field conditions.

India is critically dependent upon import of energy supply sources. Energy security demands integrated approaches and planning. Decoupling energy demand from GDP growth is also essential for complying with responsibilities towards National Action Plan on Climate Change. Therefore, the challenge is to increase the share of clean energy options in the total energy basket of India.

Ministry of Micro Small and Medium Enterprises(MSME) sector in India, which is a strong pillar of economic growth is characterized by low technology levels with some exceptions. This acts as a major handicap in the growth of MSME sector in the emerging global market and is therefore, seen as the next frontier for infusion of technology, by increasing penetration in the MSMEs. It is thus a challenge as to how the MSMEs embrace the new technologies to leap forward and contribute significantly in the inclusive growth process.

1.3.6 Low Degree of Innovativeness

The promotion of innovation, in particular technological innovation, in developing countries is becoming a fashionable subject. The growing interest in the subject stems from a recognition that it is necessary to go back to basics after experiencing the limits of traditional economic policies encapsulated in the 'Washington consensus' approach. This set of privatization, liberalization, and deregulation policies have clearly demonstrated their limits for promoting sustainable growth in the developing world. Similarly, policies focusing on modernization, in the sense of building infrastructure and institutions with a more interventionist government, have not yielded the expected fruits. Thus, there has

been a tendency to look into the black box of the engine of economic development—technology—its creation and diffusion.

Policies supporting technology development are known as 'innovation policies'. Although governments have a long practice of promoting innovation by various measures of both a direct and indirect nature, the explicit formulation of innovation policies began about 40 years ago in the 1960s. Since then such policies have been expanded and improved, while new analytical concepts, such as the concept of 'national innovation system', have been elaborated.

It should be clear that the concept of 'innovation' encompasses not only 'technological innovation', i.e. the diffusion of new products and services of a technological nature into the economy, but equally it includes non-technological forms of innovation, such as 'organization' innovations. The latter include the introduction of new management or marketing techniques, the adoption of new supply or logistic arrangements, and improved approaches to internal and external communications and positioning.

While there is considerable experience accumulated in the field of innovation policy in developed/OECD countries, much of this is not directly applicable to developing countries because of the nature of the challenges the latter are facing. In fact, developing countries face genuine obstacles to innovation and this is precisely why they remain underdeveloped. These obstacles derive from inappropriate business and governance climates and insufficient education. At the same time, there is no choice: innovation policies should cope with these difficult situations. Thus, there is a need to think about innovative approaches adapted to the needs and possibilities of developing countries.

The situation is, however, rendered more complicated because the 'developing world' presents very diverse situations in terms of levels of development and culture. Consequently, innovation policy schemes have to be tailored to countries' specific characteristics in line with the recognized fact that 'one size does not fit all', and the recognized need for working much more on national peculiarities in all walks of development economics and policies.

Innovation Climates in Developing Countries

Major weaknesses in the overall environment: Innovation climates in developing countries are first hampered by weaknesses of other key elements of knowledge-based economies as defined in the World Bank Institutes four pillar framework, namely levels of educational attainment, the business environment and the information infrastructure.

Educational levels are low in developing countries, and, this is a significant barrier to the development and diffusion of innovation in these countries. In fact, one can establish a clear relation between educational needs and the different phases of industrialization. In the pre-industrial phase, educational needs demand only basic literacy. In the industrial phase, more professional and medium-level skills are required. In the post-industrial phase, there is a need for a significant share of a population with tertiary education, with the rest of the population having at least functional literacy.

The influence of the quality of the business environment, linked to governance conditions, on innovation performances is also clearly demonstrated. However, there is a need to approach with some caution the appreciation of business environment. The quality should be seen from the perspective of countries themselves with their own values and cultural specificities. A lack of financial transparency is not necessarily a problem in a number of cultures. On the other hand, a bureaucratic climate which forces

NOTES

an entrepreneur to obtain a hundred authorizations to establish his enterprise is a problem, whatever the culture in question. More generally, when judging the quality of a business environment it is of crucial importance to go beyond the formal appearance of laws and to examine how laws are applied in practice in taking due account of the more or less informal relations regulating transactions among economic agents.

Finally, there is the issue of lack of infrastructure. Of primary importance is, of course, the telephone infrastructure. The telephone is the most important tool for (potential) entrepreneurs. Mobile phone technology has transformed the conditions of telecommunications in developing countries. Yet, the tele-density remains weak in a number of developing countries, inferior to what may be considered the minimal threshold for take-off (around 30 per cent). Progress made with mobile phone technology can lead to rapid improvements in connectivity, however it does not solve the necessity for greater internet penetration—something which remains quite low in most developing countries. Infrastructural needs for innovation in developing countries are, however, not limited to telecommunications. Road and other transport infrastructure are of primary importance, as well as sanitation, water, and other systems.

Innovation Systems

As a consequence of this overall problematic environment, innovation systems in developing countries are poorly constructed and are very fragmented. On the enterprise side, generally a large number of micro-enterprises operate in the informal economy, and a more or less important number of foreign-based firms, which tend, however, to be disconnected from the rest of the economy.

On the knowledge side, there is generally a limited research community, operating usually in an ivory tower, and a university system poorly connected to local realities, particularly to labour market needs and opportunities. Particularly problematic are the lack of technological support services and infrastructure (metrology, quality control, standards, etc.).

Public sector institutions tend to be numerous, including those supporting the promotion of enterprise development, export and foreign investment. In this often overcrowded support system, it is not easy to establish new, efficient organizations for the promotion of innovation. Where this is possible, the organizations are rarely appropriate, lacking the flexibility and drive crucial for entrepreneurship.

These overall conditions keep innovation systems into a low equilibrium trap. They are characterized by low levels of R&D in the business sector, with the bulk of national R&D effort borne by the government, and with questionable relevance for the economy.

Due to a desire not to upset the status quo and the preference of key actors to continue benefiting from vested interests and protected situations rather than taking the risk of unchartered waters, reform is usually difficult.

1.3.7 Low Level of Human and Physical Capital

Classical economists consistently identified three sources and components of national wealth: land, labour, and capital. By contrast, Western economists of the 20th century preferred to focus on capital, understood to be human-made physical capital only—the stock of structures and equipment used for production. Thus, expenses aimed at adding to this stock were the only expenses categorized as investment. Most other expenses,

such as those for education or for environmental protection, were considered to constitute consumption and treated as deductions from potential capital accumulation.

A better understanding of the need for sustainable development first led to attempts to 'green' national accounts—that is, to account for changes in natural capital in calculations of gross domestic product and gross national product—then to the development of statistical methods to account for changes in a country's human capital. Although valuation methods for natural and human capital are still imperfect, they allow experts to explore some critical development issues. These include the changing composition of a country's national wealth and operational indicators of sustainable—or unsustainable—development.

Composition of National Wealth

According to a number of recent World Bank studies, physical capital (produced assets) is not the main—much less the only—component of a country's wealth. Most important for all countries are human resources, which consist of 'raw labour', determined mainly by the number of people in a country's labour, and human capital. Natural capital is another important component of every nation's wealth.

A country's level of development determines the roles played by the different components of its national wealth. The dominance of human capital is particularly marked in the most developed countries, where natural capital is calculated to account for just 2–5 per cent of aggregate wealth. By contrast, in West Africa—one of the world's poorest regions—natural capital still prevails over physical capital, and the share of human resources is among the lowest in the world despite a large population. Comparing West Africa with Western Europe is particularly indicative because in absolute terms the two regions have roughly the same per capita value for natural capital. Thus, the striking difference in the composition of their national wealth can be entirely attributed to the fact that the average West European has 13–14 times as much human and physical capital at his or her disposal.

NOTES

2.0 MICRO AND MACRO ECONOMICS

Like most other sciences, economics is also divided into several branches and sub-branches. The two major branches of economic theory are the *microeconomic* theory and *macroeconomic* theory.

2.1 Difference Between Macro and Micro Economics

Microeconomic theory or *microeconomics,* whose literal translation is 'economics in the small,' studies the economic actions of individuals, firms and groups of individuals and firms in the economy. For example, the determination of equilibrium output and price for a single firm lies in the domain of microeconomic theory. Macroeconomic theory or *macroeconomics* is concerned with the study of economy-wide aggregates, such as the analysis of the total output and employment, total consumption, total investment, total saving and national product. Thus, while the former analysis presents a microscopic view of the economy, the latter furnishes us with its macroscopic view. Microeconomic theories are concerned with the partial equilibrium analysis of the firm's price-output determination under different market situations and the allocation of given economic resources between their rival uses. Macroeconomic theories, on the other hand, are interested in the analysis of the levels of national product, total saving and investment,

Check Your Progress

4. How is urban poverty created?
5. What is total factor productivity?
6. What are innovation policies?
7. Name the sources and components of national wealth as identified by Classical economists.

NOTES

total employment of economy's resources and total money supply. Macroeconomic analysis is the analysis of the *economy-wide* or *aggregate* variables. In short, it is the study of the economy's aggregate output, investment, saving, money supply, employment, general price level and such other macroeconomic variables.

Macroeconomics answers such important and broad questions as: What determines the levels of aggregate employment and output in the economy? How is the general price level in the economy determined? What is the relative importance of the various factors which influence the general price level? What determines the level of aggregate economic activity in the economy and its expansion or contraction over time? Why are cyclical fluctuations caused and how do these affect the overall performance of the economy? On the other hand, the concern of microeconomics is to answer such questions as: How, *ceteris paribus,* is the purchasing behaviour of a consumer influenced by any given change in the price of a commodity? How does a firm, under given market conditions, determine the output of any given good or service which it will produce and the price at which it will sell it? How will a firm in equilibrium combine the different inputs in order to produce any given good or service?

Microeconomic theory employs the technique of partial equilibrium analysis to study the price-output determination of a single commodity or service in any given market situation on the assumption of *ceteris paribus.* It studies the determination of relative prices of particular products and factors and changes in these prices. Macroeconomic theory, on the other hand, employs the technique of general equilibrium in order to study the determination of the general price level, money supply, total employment and output levels and fluctuations in these aggregate magnitudes. General equilibrium analysis stresses interdependence between the different markets and sectors in the economy. Consequently, it studies interdependence between prices and outputs of the entire range of goods and services produced in the economy. In other words, while microeconomics studies happenings in a particular market or sector in splendid isolation, macroeconomics never ignores the fact of close interdependence between economy's different sectors with everything depending on everything else in the economy. According to the general equilibrium approach employed in macroeconomics, a change in any one market or sector has its ramifications on the other markets or sectors of the economy. In short, while macroeconomic theory simplifies by aggregation, microeconomic theory simplifies by assuming 'other things remaining the same'.

In whatever words it is defined, macroeconomics is concerned with the study of the functioning of the *whole economy*, including how the total output of goods and services and the total employment of resources are determined in the economy and what causes fluctuations in their size. It analyses why at one time 15 per cent of the total labour force in the economy is unemployed while at another time only 5 per cent of the total labour force cannot find employment in the economy. It studies the phenomenon of inflation and deflation and seeks to explain why the growth rate of the economy at one time is 8 per cent while at some other time it is only 2 per cent or even less. The Great Depression of the 30s and the war andpost-war hyperinflations are among those economic phenomena which deeply concern the macroeconomist and to prevent which he strives hard to find effective solutions reflected in the monetary and fiscal policies of the government.

According to Gardner Ackley, 'macroeconomics deals with economic affairs "in the large". It concerns the overall dimensions of economic life. It looks at the total size and shape and functioning of the "elephant" of economic experience, rather than the working or articulation or dimensions of the individual parts. To alter the metaphor, it

NOTES

studies the character of the forest, independently of the trees which compose it. More specifically, 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 the national income, with the general price level.' Emphasizing that the subject matter of macroeconomics is the study of economy-wide aggregates, Edward Shapiro has stated that a major task of macroeconomics is the explanation of what determines the economy's aggregate output of goods and services. It deals with the functioning of the economy as a whole including how the economy's total output of goods and services, the price level of goods and services and the total employment of resources are determined and what causes these-magnitudes to fluctuate. According to R. G. D. Allen, 'the term "macro-economies", introduced by Ragnar Frisch in 1933, applies to the study of relations between broad economic aggregates, as opposed to the decision-taking processes of individuals and firms which is the subject matter of micro-economics.'

Microeconomics abstracts from the study of these aggregative macroeconomic variables. Its unit of study is the part and not the whole. Consequently, a micro economist picks up the problem of determination of the profit-maximizing output of a firm for his study. He is interested in finding out what particular output, out of the many possible ones, a firm must produce in order to maximize its total profit function or what particular factor-combination, out of the many possible ones, a firm should choose in order to produce a given quantity of output so as to minimize its total cost function. Microeconomic theory helps him in finding out the equilibrium (most cosy position) of the firm at that level which corresponds to the point of tangency between the firm's isoquant and isocost line.

Similarly, microeconomics is concerned with the study of the manner in which an individual consumer allocates his given income among the many goods and services available to him so as to maximize his total satisfaction or utility. Assuming the economy's total output, total employment and total spending as given, it analyses how the total output and employment are distributed between the different individual firms and industries in the economy. According to Gardner Ackley, 'microeconomics deals with the *division* of total output among industries, products, and firms, and the *allocation* of resources among competing uses. It considers problems of income *distribution.* Its interest is in *relative* prices of particular goods and services.'

Most, though not all, of the contents of the traditional economic theory, until the last 70 years, have consisted of microeconomic theory. Price and value theory, the theory of the household, firm and industry, a major part of protection and welfare theory all belong to the microeconomic theory. However, monetary theory and business cycle theory, which have a long history, are clearly macroeconomic analysis. The classical economic theory was almost wholly macroeconomics while the neoclassical theory was entirely microeconomics. Macroeconomics staged a grand comeback with John Maynard Keynes in the later part of thirties and for over a decade virtually replaced microeconomics.

While microeconomics assumes the aggregate output for the economy as a whole as given, for macroeconomics it is an important variable whose size and changes in that size it aims to explain. On the other hand, while macroeconomics treats the distribution of total output, employment and spending among the various individual goods and services produced by the particular firms and industries as given, these are regarded as variable by microeconomics. Similarly, with regard to prices, while microeconomics regards the relative prices of various different goods and services variable treating the general price level as given, macroeconomics stresses the variability of the general price level treating the relative prices as given. In the language of a metaphor, while macroeconomics is concerned with the study of an elephant as a whole, microeconomics studies the working

NOTES

of the particular parts of it. Macroeconomics studies the forest independently of the trees which compose it while microeconomics looksat the dimensions and characteristics of the individual trees which taken together constitute the forest. To alter the metaphor, while macroeconomics presents a bird's-eye view of the economy, microeconomics presents only a worm's-eye view confined to some specific part of the economy.

2.2 Microeconomics and Macroeconomics are Interdependent

Despite important differences between the microeconomic and macroeconomic theories, there is considerable overlapping between these two. Consequently, it is difficult to draw any precise line of demarcation between these two analyses. The two economic analyses are not mutually exclusive. In practice, the economy is not analysed separately in two watertight compartments. Consequently, there is only one 'economics'. Obviously, this fact should make us aware that macroeconomics has a foundation in microeconomics and *vice versa.* In any meaningful analysis of the macroeconomic variables and their relationships, the role of changes in those microeconomic variables which influence the macroeconomic variables and *vice versa* has to be recognized. For instance, if workers do not move away from the declining industry located in one area to the growing industry which is located in a different region, the total output and employment would be smaller than if the workers were mobile. In any meaningful analysis of the economic processes determining the nation's economic welfare, both the microeconomic and macroeconomic aspects of the nation's economic welfare must be considered.

From macroeconomic consideration alone, the national material welfare will be higher, if the economy attains fuller utilization of the total economic resources taking the allocation of these resources as a given. From the microeconomic point of view, the material welfare of the community will be higher, if the economy is closer to the level of optimum allocation of its given total resources, given the intensity of utilization of these resources. It is, therefore, obvious that microeconomic and macroeconomic analyses are complementary and the maximum national economic welfare will only be achieved when all the available economic resources are both fully utilized and optimally allocated between their different competing uses.

Economics is not a schizophrenic subject; it is a unified and cohesive discipline. As the two branches of economics, macroeconomics and microeconomics differ only in the degree of aggregation involved. For example, while the economy's total output and employment involve aggregation of the output and employment generated in the various sectors of the economy, the total production of the sugar industry is an aggregate of the output of different sugar factories composing the sugar industry in the economy. Similarly, the total consumer demand for sugar is an aggregate of the demands of many household units. A well-meaning general theory of the economy will, therefore, have to be a combination of both the microeconomic and macroeconomic theories.

Emphasizing the fact of interdependence and the difficulty of drawing any sharp line of demarcation between macroeconomic and microeconomic theories, Gardner Ackley has correctly stated:

> Actually, the line between macroeconomic and microeconomic theory cannot be precisely drawn. A truly 'general' theory of the economy would clearly embrace both: it would explain individual behaviour, individual outputs, incomes and prices; and the sums or averages of the individual results would constitute the aggregates with which macroeconomics is concerned. Such a general theory exists; but its very generality leaves it with little substantive content.

NOTES

Although microeconomic and macroeconomic analyses are so closely interrelated that one draws from the other, yet the two analyses differ from one another. Consequently, a microeconomic proposition cannot be extended to macroeconomic situations. For example, an individual can become richer by spending less and saving more out of his given income. A nation cannot, however, become richer unless it produces more. An attempt on the part of all individuals to save more out of their given incomes will not lead to an increase in total national savings because to the extent one individual spends less, the incomes of the rest of the people in the economy are reduced. Consequently, their savings are reduced. In fact, efforts to save more out of the given national income on the part of all individuals in the community may actually end up in reduced total savings. This is the so-called famous 'paradox of thrift' in macroeconomic theory.

Similarly, an individual can withdraw from his bank account his entire deposit money without the bank failing. If, however, all the depositors were to withdraw their bank deposits simultaneously, the bank will certainly fail. Again, a person becomes richer when he wins prize money in a lottery but the nation does not become richer because to the extent the winner of the lottery has gained, the other lottery ticket-holders have lost. Similarly, when one finds a ten-rupee bank note on the road one becomes richer to the extent of his find but the community's income remains unchanged because someone's gain is someone else's loss.

Furthermore, while in a fully employed economy it is possible for a single firm to increase its total output by weaning away inputs from other competing uses by offering higher factor rewards, it is not possible for the economy to increase the total output by such resort on the part of some firms because the increase in the output of some firms is cancelled out by the decrease in the output of others. Moreover, it is possible that in certain situations the fall in the output of other firms may more than neutralize the increase in a particular firm's output resulting in a net fall in the total output.

It is, therefore, obvious that microeconomic statements cannot always be valid for the macroeconomic decisions. Nor are the macroeconomic statements always reliable to draw correct microeconomic conclusions. For example, a substantial increase in the total agricultural output as a result of a bumper crop harvest causes an increase in the national product causing a substantial increase in the community's economic welfare. It does not, however, always follow from this that the economic condition of the agriculturists has also improved although macroeconomic theory lends strong support to this belief. In fact, since the elasticity of demand for most agricultural products is less than unity (at any rate it is so in the short period), the larger output of bumper harvest will have to be sold at more than proportionately reduced price yielding lower total revenue (income) to the farmers. Consequently, the economic condition of the agriculturists would deteriorate rather than improve while the nation as a whole would enjoy larger material well-being resulting from the bumper harvest. This is known as the familiar 'fallacy of composition'.

Again, it would be wrong to say that with the higher gross national product every one in the economy necessarily becomes richer. Even a higher per capita income is not an infallible indicator of the better economic condition of each and every individual living in the country. Ahigher national product may co-exist with greater mass poverty if in the process of producing the larger national product, the distribution of the national product becomes more skewed. For example, in the oil-producing Gulf countries very high national product and mass poverty coexist as the distribution of national product in these countries is highly skewed.

Check Your Progress

8. What does microeconomic theory study?
9. Fill in the blanks with appropriate terms.
   (i) Economics is not a schizophrenic subject; it is a _____ and _____ discipline.
   (ii) It is obvious that microeconomic statements cannot always be valid for _____ the .

2.3 SUMMARY

NOTES

In this unit, you have learnt that,

- The basic problems of an economy lie in the background of all economic decisions, and also form the basis of economic studies and generalization.
- 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.
- 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.
- In a modern economy, all the goods and services are produced by business firms. The total output generated by business firms is known as 'society's total product' or 'national output'. The total output ultimately flows to the households.
- The need for increasing the production capacity of the economy arises for at least two reasons. First, most economies of the world have realized by experience that their population has grown at a rate much higher than their productive resources. This leads to poverty, especially in the less-developed countries.
- Second, 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 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.
- 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 controlled economies, free enterprise economies have experienced it more frequently and more severely.
- 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 with technology becoming more and more efficient and productive.
- Apart from showing the possible alternative combinations of two goods, production possibilities frontier (PPF) also indicates the opportunity cost of one commodity in terms of the other product.
- Developing areas, including developing countries and regions, have a unique set of economic problems and challenges to economic development. Developing countries, taken as whole, refer to countries characterized by an underdeveloped industrial base, low per capita income, and widespread poverty.
- Broad economic stability, competitive markets, and public investment in physical and social infrastructure are widely recognized as important requirements for achieving sustained economic growth and a reduction in rural poverty.
- To understand poverty, it is essential to examine the economic and social context, including institutions of the state, markets, communities, and households. Poverty

differences cut across gender, ethnicity, age, location (rural versus urban), and income source.

- Substantial evidence suggests that a highly unequal distribution of income is not conducive to either economic growth or poverty reduction. Experience has shown that if countries put in place incentive structures and complementary investments to ensure that better health and education lead to higher incomes, the poor will benefit doubly through increased current consumption and higher future incomes.
- Inequality in society is not a new phenomenon. And yet it can be fatal. If left unchecked, it can undermine the very foundations of development and social and domestic peace.
- The world is more unequal today than at any point since Second World War. However, there are clear signs that this situation cannot be sustained for much longer. Inequality has been jeopardizing economic growth and poverty reduction.
- The world today is at a critical juncture. The financial and economic crises of recent years have pushed the international community to reconsider long-held views on economic priorities and social cohesion is much more widely recognized as a major factor contributing to resilience and sustainability.
- On the supply side, meeting an increasing food demand is a major concern, given the rise of resource constraints. Current agricultural practices are a leading source of greenhouse gas emissions, while also leading to other problems, such as loss of soil fertility and water pollution from run-off.
- Technological progress is at the heart of human progress and development. As the 1998 World Development Report on the knowledge economy (World Bank 1998) emphasized, the understanding of how things are created and the communication of that knowledge are critical drivers of economic progress.
- Policies supporting technology development are known as 'innovation policies'. Although governments have a long practice of promoting innovation by various measures of both a direct and indirect nature, the explicit formulation of innovation policies began about 40 years ago in the 1960s.
- As a consequence of this overall problematic environment, innovation systems in developing countries are poorly constructed and are very fragmented.
- Classical economists consistently identified three sources and components of national wealth: land, labour, and capital.
- A country's level of development determines the roles played by the different components of its national wealth. The dominance of human capital is particularly marked in the most developed countries, where natural capital is calculated to account for just 2–5 per cent of aggregate wealth.
- Like most other sciences, economics is also divided into several branches and sub-branches. The two major branches of economic theory are the microeconomic theory and macroeconomic theory.
- Microeconomic theory or microeconomics, whose literal translation is 'economics in the small,' studies the economic actions of individuals, firms and groups of individuals and firms in the economy.
- Macroeconomic theory or macroeconomics is concerned with the study of economy-wide aggregates, such as the analysis of the total output and employment, total consumption, total investment, total saving and national product.

NOTES

NOTES

- While microeconomics assumes the aggregate output for the economy as a whole as given, for macroeconomics it is an important variable whose size and changes in that size it aims to explain. On the other hand, while macroeconomics treats the distribution of total output, employment and spending among the various individual goods and services produced by the particular firms and industries as given, these are regarded as variable by microeconomics.
- Despite important differences between the microeconomic and macroeconomic theories, there is considerable overlapping between these two. Consequently, it is difficult to draw any precise line of demarcation between these two analyses.
- Macroeconomics has a foundation in microeconomics and *vice versa.*

2.4 KEY TERMS

- **Society's total product/National output:** The total output generated by business firms is known as 'society's total product' or 'national output'.
- **Total factor productivity:** The relative efficiency with which a given economy produces goods and services given a certain quantity of labour and capital is called total factor productivity (TFP).

2.5 ANSWERS TO 'CHECK YOUR PROGRESS'

1. The major economic problems faced by an economy—whether capitalist, socialist or mixed—may be classified into two broad groups:
   - Microeconomic problems which are related to the working of the economic system.
   - Macroeconomic problems related to the growth, employment, stability, external balance, and macroeconomic policies for the management of the economy as a whole.
2. 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.
3. 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 with technology becoming more and more efficient and productive.
4. Urban poverty is created by the rural poor's efforts to get out of poverty by moving to cities.
5. The relative efficiency with which a given economy produces goods and services given a certain quantity of labour and capital is called total factor productivity (TFP).
6. Policies supporting technology development are known as 'innovation policies'.
7. Classical economists consistently identified three sources and components of national wealth: land, labour, and capital.

8. Microeconomic theory or *microeconomics*, whose literal translation is 'economics in the small,' studies the economic actions of individuals, firms and groups of individuals and firms in the economy.
9. (i) unified; cohesive

   (ii) macroeconomic decisions

2.6 QUESTIONS AND EXERCISES

Short-Answer Questions

1. What are the basic problems of an economy? How can they be classified?
2. State the basic microeconomic problems faced by an economy.
3. How can the production capacity of an economy be increased?
4. What is meant by opportunity cost? How can it be increased?
5. What are the main contributors to urban and rural poverty?
6. Who are the Low-Income Developing Countries? How can absolute poverty be alleviated?
7. Write a note on inequality in the distribution of income and opportunities.
8. 'Technological progress is at the heart of human progress and development.' Describe.
9. What is the problem of low degree of innovativeness in the developing countries?
10. What is the 'paradox of thrift' in macroeconomic theory?

Long-Answer Questions

1. Discuss the major microeconomic problems faced by an economy.
2. Describe the major macroeconomic problems of an economy.
3. Assess the production possibilities of an economy.
4. Assess the problem of poverty, low income and inequality in the distribution of income and opportunities in the developing countries.
5. Evaluate the problem of resource constraints and low level of technology in the developing countries.
6. Critically analyse the problem of low degree of innovativeness and low level of human and physical capital in the developing countries.
7. Discuss the major differences between macro and micro economics.
8. Assess the statement, 'Microeconomics and macroeconomics are interdependent'.

2.7 FURTHER READING

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.

NOTES

Keating, B. and J. H. Wilson. 2003. *Managerial Economics: An Economic Foundation for Business Decisions,* 2nd Edition. New Delhi: Biztantra.

Mansfield, E., W. B. Allen, N. A. Doherty and K. Weigelt. 2002. *Managerial Economics: Theory, Applications and Cases,* 5th Edition. NY: W. Orton & Co.

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

Salvantore, Dominick. 2001. *Managerial Economics in a Global Economy*, 4th Edition. Australia: Thomson-South Western.

Thomas, Christopher R. and Maurice S. Charles. 2005. *Managerial Economics: Concepts and Applications,* 8th Edition. New Delhi: Tata McGraw-Hill.

Adhikary, M. 2000. *Business Economics.* New Delhi: Excel Books.

Baumol, W. J. 1996. *Economic Theory and Operations Analysis,* 3rd edition. New Jersey: Prentice-Hall.

Chopra, O. P. 1985. *Managerial Economics.* New Delhi: Tata McGraw-Hill.

Kautsoyiannis, A. 1991. *Modern Microeconomics*. New York: Macmillan.

UNIT III DEMAND ANALYSIS

Structure

NOTES

3.0 INTRODUCTION

Consumer demand is the basis of all productive activities. Just as 'necessity is the mother of invention', *demand is the mother of production*. Increasing demand for a product offers high business prospects for it in future and decreasing demand for a product diminishes its business prospect. For example, increasing demand for computers, cars and mobile phones in India has enlarged the business prospect for both domestic and foreign companies selling these goods. On the other hand, declining demand for black and white TV sets and manual typewriters is forcing their companies to switch over to modern substitutes or else go out of business. It is, therefore, essential for business managers to have a clear understanding of the following aspects of demand for their products:

- What is the basis of demand for a commodity?
- What are the determinants of demand?
- How do the buyers decide the quantity of a product to be purchased?
- How do the buyers respond to change in product prices, their incomes and prices of the related goods?
- How can the total or market demand for a product be assessed and forecasted?

These questions are answered by the **Theory of Demand**. The analysis of total demand for a firm's product plays a crucial role in business decision-making. The market demand or the size of the market at a point in time at different prices gives the overall scope of business; it gives prospects for expanding business; and it plays a crucial role in planning for future production, inventories of raw materials, advertisement, and setting up sales outlets. Therefore, the information regarding the magnitude of the current and future demand for the product is indispensable. Theory of demand provides an insight into these problems. From the analysis of market demand, business executives can know:

NOTES

- The factors that determine the size of demand
- Elasticities of demand, i.e., how responsive or sensitive is the demand to the changes in its determinants
- Possibility of sales promotion through manipulation of prices
- Responsiveness of demand to advertisement expenditure
- Optimum levels of sales, inventories and advertisement cost

In this unit, we discuss the basis of demand, diminishing marginal utility, income of the consumer and the budget line, constrained utility maximization, demand curve and factors shifting it and elasticities of demand and their measurement.

3.1 UNIT OBJECTIVES

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

- Discuss utility as a basis of consumer demand
- Explain the law of diminishing marginal utility and its necessary conditions
- Assess the income of the consumer and the budget line
- Analyse the concept of consumer equilibrium and constrained utility maximization
- Evaluate the demand curve and the factors shifting the demand curve
- Discuss price elasticity of demand
- Explain the income elasticity of demand and the use of income-elasticity in business decisions

3.2 BASIS OF DEMAND: UTILITY AND INCOME

The consumers demand a commodity because they derive or expect to derive *utility* from that commodity. The expected utility from a commodity is the basis of demand for it. Though 'utility' is a term of common usage, it has a specific meaning and use in the analysis of consumer demand. We will, therefore, describe in this section the meaning of utility, the related concepts and the law associated with utility.

Meaning of Utility

The concept of utility can be looked upon from two angles—from the commodity angle and from the consumer's angle. From the commodity angle, *utility is the want-satisfying property of a commodity*. From the consumer's angle, *utility is the psychological feeling of satisfaction, pleasure, happiness or well-being which a consumer derives from the consumption, possession or the use of a commodity*.

There is a subtle difference between the two concepts which must be borne in mind. The concept of a want-satisfying property of a commodity is 'absolute' in the sense that this property is ingrained in the commodity irrespective of whether one needs it or not. For example, a pen has its own utility irrespective of whether a person is literate or illiterate. Another important attribute of the 'absolute' concept of utility is that it is *ethically neutral* because a commodity may satisfy a frivolous or socially immoral need, e.g., alcohol, drugs or a profession like prostitution.

On the other hand, from a consumer's point of view, utility is a post-consumption phenomenon as one derives satisfaction from a commodity only when one consumes or

uses it. Utility in the sense of satisfaction is a 'subjective' or a 'relative' concept. In the subjective sense, utility is a matter of one's own feeling of satisfaction. In the relative sense: (*i*) a commodity need not be useful for all, for example, cigarettes do not have any utility for non-smokers, and meat has no utility for strict vegetarians; (*ii*) utility of a commodity varies from person to person and from time to time; and (*iii*) a commodity need not have the same utility for the same consumer at different points of times, at different levels of consumption and at different moods of a consumer. In consumer analysis, only the 'subjective' concept of utility is used.

behaviour.

3.2.1 Measurability of Utility

Utility is a psychological phenomenon. It is a feeling of satisfaction, pleasure or happiness. Is utility measurable quantitatively? Measurability of utility has, however, been a contentious issue. The classical economists, viz., Jeremy Bentham, Leon Walrus, Carl Menger and the neo-classical economist, 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. The modern economists, most notably J. R. Hicks and R. G. D. Allen, however, hold the view that utility is not quantitatively measurable—it is not measurable in absolute terms. Utility can be expressed only ordinally comparatively 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. This is known as the *ordinal* concept of utility. Let us now look into the origin of the two concepts of utility and their use in the analysis of demand.

(i) Cardinal measurement of utility: Some early psychological experiments on an individual's responses to various stimuli led classical and neo-classical economists to believe that utility is measurable and cardinally quantifiable. 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. The neo-classical economists, especially Marshall, devised a method of measuring utility. According to Marshall, utility of a commodity for a person equals the amount of money he is willing to pay for a unit of the commodity. In other words, price one is prepared to pay for a unit of a commodity equals the utility he expects to derive from the commodity. They formulated 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 economic analysis, they assumed (*i*) that one 'util' equals one unit of money, and **(ii)** that utility of money remains constant.

It has, however, been realised 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., the utility. Numerous factors affect the state of consumer's mood, which are impossible to determine and quantify. *Utility is therefore immeasurable in cardinal terms*.

(ii) Ordinal measurement of utility: The modern economists have discarded the concept of *cardinal utility* and have instead employed the concept of *ordinal utility* for analysing 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 numerical terms, but it is always possible for them to tell introspectively whether a commodity is more or less or equally useful as compared to another. For example, a consumer may not be able to tell that a bottle of Pepsi gives 5 utils and a glass of fruit juice gives 10 utils. But he or she can always tell whether a glass of fruit juice gives more or less utility than a bottle of Pepsi. This assumption forms the basis of the ordinal theory of consumer

NOTES

NOTES

To sum up, the neo-classical economists maintained that cardinal measurement of utility is practically possible and is meaningful in consumer analysis. The modern economists, on the other hand, maintain that utility being a psychological phenomenon is inherently immeasurable quantitatively. They also maintain that the concept of ordinal utility is a feasible concept and it meets the conceptual requirement of analysing the consumer behaviour. However, both the concepts of utility are used in analysing consumer behaviour.

3.2.2 Two Approaches to Consumer Demand Analysis

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

- **Cardinal Utility Approach**, attributed to Alfred Marshall and his followers, is also called the Neo-classical Approach or Marshallian approach.
- **Ordinal Utility Approach**, pioneered by J. R. Hicks, a Nobel laureate and R. G. D. Allen, is also called Hicks-Allen approach or 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 managerial decisions 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 neo-classical demand theory based on cardinal utility has retained its appeal and applicability to the analysis of market behaviour. Besides, the study of neo-classical demand theory serves as a foundation for understanding the advanced theories of consumer behaviour. The study of neo-classical theory of demand is of particular importance and contributes a great deal in managerial decisions.

3.3 DIMINISHING MARGINAL UTILITY

The law of diminishing marginal utility is one of the fundamental laws of economics. This law states that *as the quantity consumed of a commodity increases, the utility derived from each successive unit decreases, consumption of all other commodities remaining the same*. In simple words, when a person consumes more and more units of a commodity per unit of time, e.g., *rasgullas*, keeping the consumption of all other commodities constant, the utility which he derives from the successive *rasgullas* he consumes goes on diminishing. This law applies to all kinds of consumer goods—durable and non-durable sooner or later.

Check Your Progress

1. Give an important attribute of the 'absolute' concept of utility.
2. What gave rise to the concept of cardinal utility?
3. What are the two approaches to the analysis of consumer behaviour?

To explain the law of diminishing marginal utility, let us suppose that a consumer consumes 6 units of a commodity *X* and his total and marginal utility derived from various units of *X* are as given in Table 2.1.

Table 2.1 *Total and Marginal Utility Schedules of X*

No. of units consumed	*Total utility*	*Marginal utility*
1	30	30
2	50	20
3	60	10
4	65	5
5	60	– 5
6	45	– 15

NOTES

As shown in Table 2.1, with the increase in the number of units consumed per unit of time, the *TU* increases but at a diminishing rate. The diminishing rate of increase in the total utility gives the measure of marginal utility. The diminishing *MU* is shown in the last column of the table. Fig. 2.1 illustrates graphically the law of diminishing *MU*. The rate of increase in *TU* as the result of increase in the number of units consumed is shown by the *MU* curve in Fig. 2.1. The downward sloping *MU* curve shows that marginal utility goes on decreasing as consumption increases. At 4 units consumed, the *TU* reaches its maximum level, i.e., 65 utils. Beyond this, *MU* becomes negative and *TU* begins to decline. The downward sloping *MU* curve illustrates the law of diminishing marginal utility.

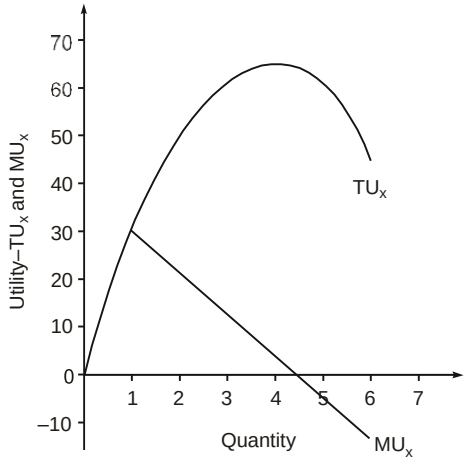


Fig. 2.1 *Diminishing Marginal Utility*

Why the *MU* Decreases: The utility gained from a unit of a commodity depends on the intensity of the desire for it. When a person consumes successive units of a commodity, his need is satisfied by degrees in the process of consumption and the intensity of his need goes on decreasing. Therefore, the utility obtained from each successive unit goes on decreasing.

Necessary Conditions: The law of diminishing marginal utility holds only under certain conditions. These conditions are referred to as the *assumptions* of the law. The assumptions of the law of diminishing marginal utility are listed below.

First, the unit of the consumer good must be a standard one, 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, the consumer's taste or preference must remain the same during the period of consumption.

Third, there must be continuity in consumption. Where a break in continuity is necessary, the time interval between the consumption of two units must be appropriately short.

Fourth, the mental condition of the consumer must remain normal during the period of consumption. Otherwise, the law of diminishing *MU* may not apply.

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, melodious songs, etc. the marginal utility may initially increase, but eventually it does decrease. As a matter of fact, the law of marginal utility generally operates universally.

3.4 INCOME OF THE CONSUMER AND THE BUDGET LINE

NOTES

Given the indifference curves and indifference map, the consumer is free to choose an *IC* curve and opt for any point on the chosen *IC*. Given the option, the consumer would like to choose the highest *IC.* But, he cannot because he faces a budgetary limitation. Recall that the consumer has a limited income and goods he consumes have a price. Limited income and prices impose constraints on consumer's choice, called *budgetary constraints*. Given the budgetary constraint, the consumer cannot opt for the highest *IC*. Let now see the implications of budgetary constraints on consumer's choices.

Given the indifference map, a utility maximizing consumer would like to reach the highest possible indifference curve on his indifference map. But, as noted above, the consumer is assumed to have a limited income. The limitedness of income acts as a *constraint on how high a consumer can ride on his indifference map*. This is known as *budgetary constraint*. In a two-commodity model, assuming a limited money income (*M*), the *budgetary constraint*, may be expressed through a *budget equation* as shown in Eq. 2.1.

$$M = P_x \cdot Q_x + P_y \cdot Q_y \qquad ...(2.1)$$

where P_x and P_y are prices of *X* and *Y*, respectively; Q_x and Q_y are their respective quantities; and *M* is the consumer's money income.

The budget equation states that the total expenditure of the consumer on goods *X* and *Y* cannot exceed his total income, *M*. The total quantity of *X* and *Y* that can be bought with given *M*, P_x and P_y can be easily obtained from the budget equation, as shown below.

$$Q_x = \frac{M}{P_x} - \frac{P_y}{P_x} Q_y \qquad ...(2.2)$$

and

$$Q_y = \frac{M}{P_y} - \frac{P_x}{P_y} Q_x \qquad ...(2.3)$$

These equations are also called budget equations. Given the budget equations, if *M*, P_x and P_y are known, the values of Q_x and Q_y and different combinations thereof can be easily calculated by assigning a numerical value to Q_y or to Q_x. When the values of Q_x and Q_y are plotted on the *X* and *Y* axes, and joined by a line, it produces a line which is called the *budget line* or *price line*, as shown in Fig. 2.2.

There is a simple method of drawing the Budget Line. Given the Eq. (2.2), find Q_x at $Q_y = 0$. Q_x equals M/P_x. Mark M/P_x as a point on *X*-axis. Similarly, given Eq. (2.3), find Q_y at $Q_x = 0$. $Q_y = M/P_y$. Mark M/P_y point at *Y*-axis. Both M/P_x and M/P_y points are shown in Fig. 2.2. By joining those points by a line, we get the budget line. *The budget line shows the quantity-combinations available to the consumer given his income and the prices of X and Y.*

NOTES

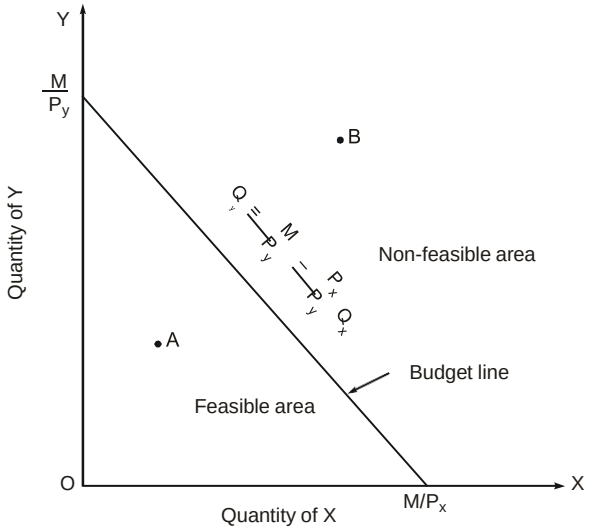


Fig. 2.2 *Budget Line and Budget Space*

The budget line divides the *commodity space* into two parts: (*i*) feasibility area, and (*ii*) non-feasibility area. The area under the budget line (including the budget line) is *feasibility* area (Fig. 2.2). For, any combination of goods *X* and *Y* represented by a point within this area (e.g., point *A*) or on the boundary line (i.e., on the budget line) is a feasible combination, given *M*, P_x and P_y. The area beyond the budget line is *non-feasible* area because any point falling in this area, e.g., point *B*, is unattainable (given *M*, P_x and P_y).

Shifts in the Budget Line: The budget line shifts upward or downward or swivels up and down due to change in the consumer's income and prices of the commodities. If the consumer's income increases, prices remaining the same, the budget line shifts upwards remaining parallel to the original budget line. Suppose the original budget line is given by line *AB* in Fig. 2.3. If money income (*M*) increases (prices remaining the same), the budget line *AB* will make a parallel shift to *CD*. And, if *M* decreases by the same amount, the budget line will shift downward to its original position *AB*. Income remaining the same, if prices change, the budget line changes its position. For example, if *M* and P_y remain constant and P_x decreases to a half then the budget line will be *AF*. Similarly, *M* and P_x remaining constant, if P_y increases, the budget line shifts to *EB*.

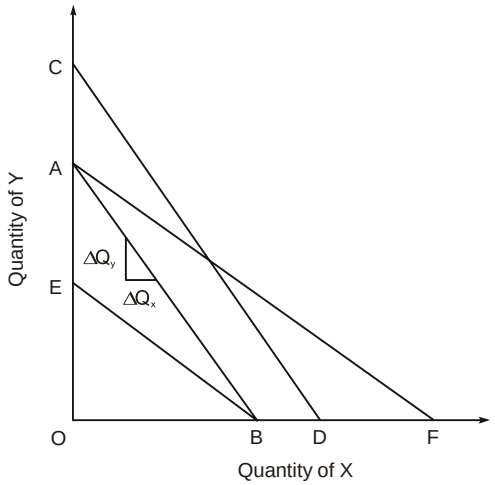


Fig. 2.3 *Shift in the Budge Space*

NOTES

Slope of the Budget Line: **Another important aspect of the budget line that matters indetermining a consumer's equilibrium is its *slope*. The slope of the budget line (*AB*) inFig. 2.3, is given as:**

$$\frac{\Delta Q_y}{\Delta Q_x} = \frac{OA}{OB}$$

Since $OA = M/P_y$ (when $X = 0$) and $OB = M/P_x$ (when $Y = 0$), the *slope of the budget line AB* in Fig. 2.3 may be rewritten as:

$$\frac{OA}{OB} = \frac{M/P_y}{M/P_x} = \frac{P_x}{P_y}$$

Thus, the slope of the budget line is the same as the *price ratio* of the two goods.

3.4.1 Consumer Equilibrium

In this section onwards, we take up the main theme of the theory of consumer behaviour as developed under the ordinal utility approach. The main issue is how a consumer attains his equilibrium. As noted earlier, a consumer attains his equilibrium when he maximizes his total utility, given his income and market prices of the goods and services that he consumes. The ordinal utility approach specifies two conditions for the consumer's equilibrium:

- Necessary or the first order condition
- Supplementary or the second order condition

In a two-commodity model, the necessary or the *first order* condition under *ordinal utility approach* is the same as equilibrium condition under *cardinal utility approach*. It is given as:

$$\frac{MU_x}{MU_y} = \frac{P_x}{P_y}$$

Since, by implication, $MU_x/MU_y = MRS_{x,y}$, the necessary condition of equilibrium under ordinal utility approach can be written as:

$$MRS_{x,y} = \frac{MU_x}{MU_y} = \frac{P_x}{P_y}$$

This is a necessary but not a *sufficient condition* of consumer's equilibrium. The *second order* or *supplementary condition* requires that the necessary condition be fulfilled at the *highest possible indifference curve*.

Consumer's equilibrium is illustrated in Fig. 2.4. The indifference curves IC_1, IC_2 and IC_3 present a hypothetical indifference map of the consumer. The line *AB* is the hypothetical budget line. Both the budget line *AB* and the indifference curve IC_2 pass through point *E*. Therefore, the slopes of the indifference curve IC_2 and the budget line (*AB*) are equal. Thus, both the *necessary* and *supplementary* conditions are fulfilled at point *E*. Therefore, consumer is in equilibrium at point *E*. This point can be proved as follows.

We know that between any two points on an indifferent curve, MU_y of $\Delta Y = MU_x$ of ΔX and, therefore, the slope of an indifference curve is given by:

$$\frac{\Delta Y}{\Delta X} = \frac{MU_y}{MU_x} = MRS_{x,y}$$

We know also that the slope of the budget line is given by:

$$\frac{OA}{OB} = \frac{P_y}{P_x}$$

NOTES

As shown in Fig. 2.4, at point *E*, $MRS_{y,x} = P_y / P_x$. Therefore, the consumer is in equilibrium at point *E*. The tangency of IC_2 with the budget line *AB*, indicates that IC_2 is the highest possible indifference curve which the consumer can reach, given his budgetary constraint and the prices. At equilibrium point *E*, the consumer consumes OQ_x of *X* and OQ_y of *Y*, which yield him the maximum satisfaction.

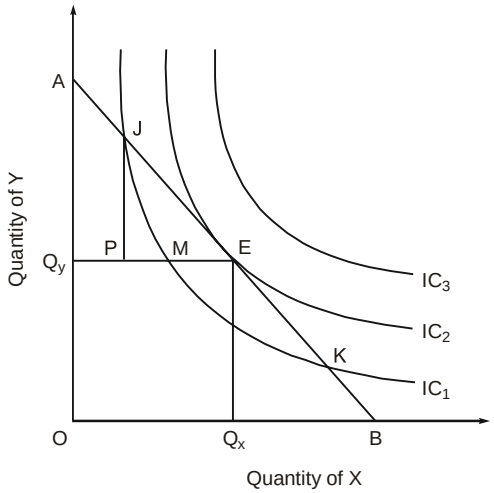


Fig. 2.4 *Equilibrium of the Consumer*

Although the necessary condition is also satisfied on two other points, *J* and *K* (i.e., the points of intersection between the budget line *AB* and a lower indifference curve IC_1), these points do not satisfy the *second order* condition. Indifference curve IC_1 is not the highest possible curve on which the necessary condition is fulfilled. Since indifference curve IC_1 lies below the curve IC_2, at any point on IC_1, the level of satisfaction is lower than the level of satisfaction indicated by IC_2. So long as the utility maximizing consumer has an opportunity to reach the curve IC_2, he would not like to settle on a lower indifference curve.

From the information contained in Fig. 2.4, it can be proved that the level of satisfaction at point *E* is greater than that on any other point on IC_1. Suppose the consumer is at point *J*. If he moves to point *M*, he would be equally well-off because points *J* and *M* are on the same indifference curve. If he moves from point *J* to *M*, he will have to sacrifice *JP* of *Y* and take *PM* of *X*. But in the market, he can exchange *JP* of *Y* for *PE* of *X*. That is, he gets extra *ME* (= *PE* – *PM*) of *X*. Since *ME* of *X* gives him extra utility, the consumer moves to point *E*. Since point *E* falls on a higher *IC*, it represents a utility higher than the point *M*. Therefore, point *E* is preferable to point *M*. The consumer will, therefore, have a tendency to move to point *E* on a higher IC_2 from any other point on the curve IC_1, all other things (taste, preference and prices of goods) remaining the same.

Another fact which is obvious from Fig. 2.4 is that, due to budget constraint, the consumer cannot move to an indifference curve placed above and to the right of IC_2. For example, his income would be insufficient to buy any combination of two goods at the curve IC_3. Note that the indifference curve IC_3 falls in the infeasibility area.

NOTES

3.4.2 Constrained Utility Maximization

The central theme of the consumption theory—be it based on ordinal utility or cardinal utility approach—is the *utility maximizing behaviour* of the consumer. The fundamental postulate of the consumption theory is that all the consumers—individuals and households—aim at *utility maximization* and all their decisions and actions as consumers are directed towards utility maximization. The specific questions that the consumption theory seeks to answer are:

- How does a consumer decide the optimum quantity of a commodity that he or she chooses to consume, i.e., how does a consumer attain his/her equilibrium in respectto each commodity?
- How does he or she allocate his/her disposable income between various commodities of consumption so that his/her total utility is maximized?

The theory of consumer behaviour seeks to answer these questions on the basis of the postulates that consumers seek to maximize their total utility or satisfaction.

Constrained utility maximization is a process wherein under certain constraints the highest possible level of utility is obtained through the consumption of goods and services. This happens when the highest overall level of utility cannot be obtained. The concept of constrained utility maximization is an alteration of the more general utility maximization assumption. The general utility maximization is based on the notion that the consumers might be regulated or restricted from achieving the absolute maximum level of utility. The major restriction would be the amount of income available in comparison to the price paid. Constrained utility maximization generally does reach the peak of the total utility curve.

While the idea of utility maximization as an unrestricted quest of utility is very essential in the consumer demand theory and the study of economics, our everyday life is directed by the idea of constrained utility maximization.

3.5 DEMAND CURVE AND FACTORS SHIFTING IT

When the demand curve changes its position (retaining its slope though not necessarily), the change is known as a *shift in the demand curve*. For example, suppose that the original demand curve for commodity X is given as D_1 in Fig. 2.5. As shown in the figure, at price OP_2, the consumer would buy OQ_1 units of X, other factors remaining constant. But, if any of the other factors (e.g., consumer's income or price of the substitutes) changes, it will change the consumer's ability and willingness to buy commodity X. For example, if consumer's disposable income increases due to decrease in income tax, he would be able to buy OQ_2 units of X instead of OQ_1. This is true for the whole range of prices of X; consumers would be able to buy more at all other prices. This will cause an *upward* shift in demand curve from D_1 to D_2. Similarly, decrease in disposable income of the consumer due to, say, rise in taxes may cause a *downward shift* in the demand curve from D_2 to D_1.

Check Your Progress

4. List any two assumptions of the law of diminishing marginal utility.
5. When does a consumer attain equilibrium?
6. What is the central theme of the consumption theory?

NOTES

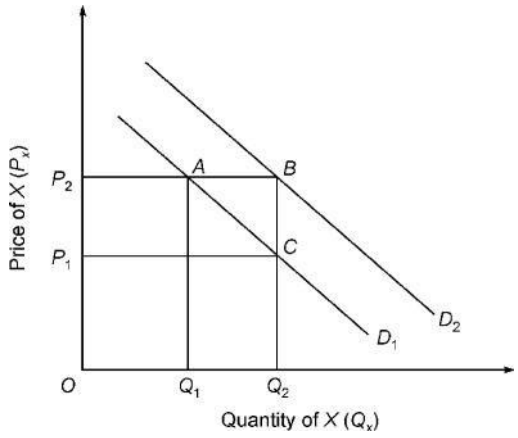


Fig. 2.5 *Shift in Demand Curve*

2.5.1 Factors Behind Shifts in the Demand Curve

Shifts in a price-demand curve may take place owing to the change in one or more non-price determinants of the demand for a commodity. Consider, for example, the increase in demand for commodity *X* by Q_1Q_2 in Fig. 2.5. Given the price OP_2, the demand for *X* might have increased by Q_1Q_2 for any of the following reasons.

- Increase in consumer's income so that he can buy OQ_2 of *X* at price OP_2; this is *income effect*
- Price of the substitute of *X* rises so that the consumers find it gainful to prefer Q_1Q_2 of *X* for its substitute: this is *substitution effect*
- Advertisement by the producer of the commodity *X* changes consumer's taste or preference in favour of commodity *X* so much that the consumer buys more of *X* or he prefers Q_1Q_2 to its substitute, again a *substitution effect*
- Price of a complement of *X* falls so much that the consumer can afford OQ_2 of *X*
- Price remaining the same, demand for *X* might increase also for such reasons as *X* gaining fashion status, improvement in its quality, change in production technology and seasonality of the product

It is important for the business decision-makers to bear in mind the distinction between changes in demand due to: (*i*) shift in price-demand curve; and (*ii*) movement along the demand curve. For instance, in Fig. 2.5, the increase in quantity demanded from OQ_1 to OQ_2 can be explained in two different ways: ***one***, by moving down from point *A* to *C* along the demand curve D_1 which results from a fall in price from P_2 to P_1; and ***two***, through upward shift in demand curve from D_1 to D_2. In the former case, additional demand is obtained at the cost of some revenue. In the latter case, demand increases due to a shift in the demand curve on account of some other factors, such as increase in consumer's income, increase in the price of substitutes, increase in population, etc. This kind of increase in demand results in increase in revenue. However, in case the demand curve is made to shift through advertisement or other sales promotion devices, the additional demand is not free of cost. Moreover, it is the latter kind of increase in demand which is hoped for and attempted by business firms.

Increase and Decrease vs Extension and Contraction of Demand: Economists sometimes make a distinction between: (*a*) increase and decrease in demand, and (*b*) extension and contraction in demand. *Increase and decrease* in demand are associated

NOTES

with non-price-quantity relationships of demand whereas *extension and contraction* of demand are associated with the price-quantity relationship of demand. For example, in Fig. 2.5, movement from point *A* to *B* is an *increase in demand* and movement from *B* to *A* is a *decrease in demand*. On the other hand, movement from *A* to *C* is an *extension of demand* and movement from *C* to *A* is a *contraction of demand*. In other words, movement along the demand curve implies extension or contraction of demand.

This kind of distinction of terminology between a change in demand caused by different factors is, however, a matter of convenience. It has no theoretical basis.

3.6 ELASTICITIES OF DEMAND

Importance of the Elasticity Concept

We have earlier discussed the *nature* of *relationship* between demand and its determinants. From managerial point of view, however, the knowledge of nature of relationship alone is not sufficient. What is more important is the extent of relationship or the degree of responsiveness of demand to the changes in its determinants. The degree of responsiveness of demand to the change in its determinants is called *elasticity of demand*.

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 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 depends on:

(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.

In this section, we will discuss various methods of measuring elasticities of demand. The concepts of demand elasticities used in business decisions are: (*i*) Price elasticity, (*ii*) Cross-elasticity; (*iii*) Income elasticity; and (*iv*) Advertisement elasticity, and (*v*) Elasticity of price expectation.

Check Your Progress

7. What is meant by shift in the demand curve?
8. State one reason for the occurrence of shifts in a price-demand curve.

2.6.1 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 as a result of 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, derived from this definition of elasticity, is given as follows:

$$e_p = \frac{\Delta Q}{Q} \div \frac{\Delta P}{P} = \frac{\Delta Q}{Q} \times \frac{P}{\Delta P}$$

$$= \frac{\Delta Q}{\Delta P} \times \frac{P}{Q} \quad \ldots(2.4)$$

where Q = original quantity demanded, P = original price, ΔQ = change in quantity demanded and ΔP = 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 elasticity can be measured between any two points on a demand curve (called *arc elasticity*) or at a point (called *point elasticity*).

APC Elasticity

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. 2.6) is the measure of arc elasticity. The movement from point *J* to *K* on the demand curve (D_x) shows a fall in the price from ` 20 to ` 10 so that $\Delta P = 20 - 10 = 10$. The fall in price causes an increase in demand from 43 units to 75 units so that $\Delta Q = 43 - 75 = -32$. The elasticity between points *J* and *K* (moving from *J* to *K*) can be calculated by substituting these values into the elasticity formula as follows:

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

$$= -\frac{-32}{10} \frac{20}{43} = 1.49 \quad \ldots(2.5)$$

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.

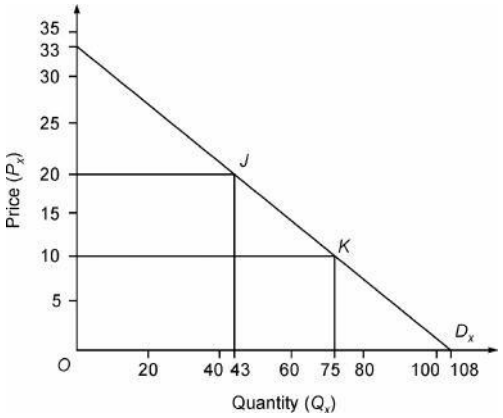


Fig. 2.6 *Linear Demand Curve*

Problem in Using Arc Elasticity The arc elasticity should be measured and used carefully, otherwise it may lead to wrong decisions. Arc elasticity co-efficients 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. (2.5), the elasticity between

NOTES

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 co-

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

$$e_p = - \frac{32}{-10} \cdot \frac{10}{75} = 0.43 \qquad ...(2.6)$$

The measure of elasticity co-efficient in Eq. (2.6) for the reverse movement in price is obviously different from one given by Eq. (2.5). 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.

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, so that

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

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

$$e_p = - \frac{32}{10} \cdot \frac{10}{43} = 0.74 \qquad ...(2.7)$$

This method is however devoid of the logic of calculating percentage change 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

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

or

$$e_p = - \frac{Q_1 - Q_2}{P_1 - P_2} \cdot \frac{(P_1 + P_2)/2}{(Q_1 + Q_2)/2} \qquad \ldots(2.8)$$

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

Substituting the values from our example, we get,

$$e_p = - \frac{43 - 75}{20 - 10} \cdot \frac{(20 + 10)/2}{(43 + 75)/2} = 0.81$$

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. 2.6)—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 *B* on the linear demand curve *MN* in Fig. 2.7. This is in contrast to the arc elasticity between points *P* and *B*. A movement from point *B* towards *P* implies change in price (□*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.

Point elasticity $(e_p) = \frac{P}{Q} \frac{\partial Q}{\partial P}$ ...(2.9)

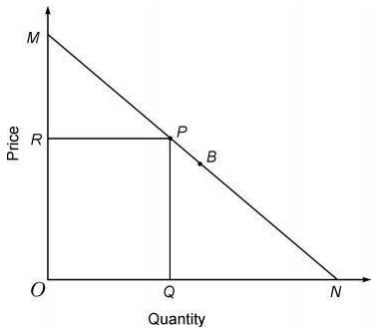


Fig. 2.7 *Point Elasticity*

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 *Q* = *OQ*. By substituting these values in Eq. (2.9), 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. 2.6. At point *J*,

NOTES

NOTES

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

Note that $e_p = 1.51$ 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 we will see below, geometrically, $QN/OQ = PN/PM$. Therefore, elasticity of demand at point P (Fig. 2.7) may be expressed as

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

Proof. The fact that $e_p = PN/PM$ can be proved as follows. Note that in Fig. 8.8, there are three triangles—$\square MON$, $\square MRP$ and $\square PQN$—and $\square MON$, $\square MRP$ and $\square PQN$ are right angles. Therefore, the other corresponding angles of the three triangles will always be equal and hence, $\square MON$, $\square MRP$ and $\square PQN$ 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 $\square PQN$ and $\square MRP$,

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

Since $RP = OQ$, by substituting OQ for RP in the above equation, we get

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

It follows that

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

It means that price elasticity of demand at point P (Fig. 2.7) 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.,

$$e_p = \frac{\text{Lower segment}}{\text{Upper segment}}$$

Point elasticity on a non-linear demand curve. The ratio $\square Q/\square P$ 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. 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. 2.8) 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

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

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

NOTES

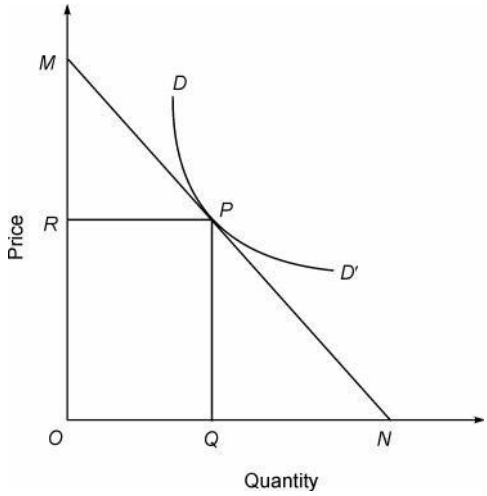


Fig. 2.8 *Non-linear Demand Curve*

To conclude, at midpoint of a linear demand curve, $e_p = 1$. Note that in Fig. 2.9, 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*.

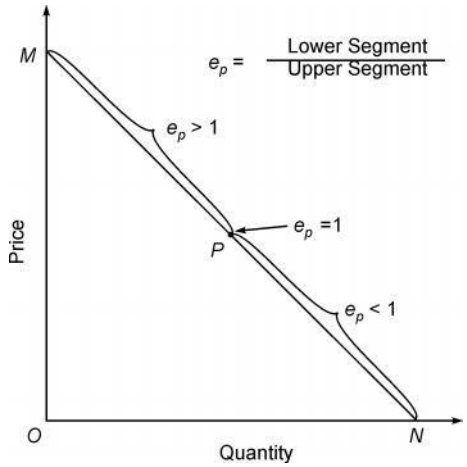


Fig. 2.9 *Point Elasticities of Demand*

Measuring Price Elasticity fro a Demand Function

The price elasticity of demand for a product can be measured directly from the demand function. In this section, we will 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 ΔP and ΔQ. We will, therefore, confine ourselves here to point elasticity of demand with respect to price.

Price Elasticity from a Linear Demand Function

NOTES

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 we want to measure elasticity at $P = 10$. We know that

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

The term $\delta Q/\delta P$ in the elasticity formula is the slope of the demand curve. The slope of the demand curve can be found by differentiating the demand function. That is,

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

Having obtained the slope of the demand curve as $\delta Q / \delta P = -5$, e_p at $P = 10$ can be calculated as follows. Since, $P = 10$, $Q = 100 - 5(10) = 50$. By substituting these values into the elasticity formula, we get,

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

Similarly, at $P = 8$, $Q = 100 - 5(8) = 60$ and

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

And at $P = 15$, $Q = 100 - 5(15) = 25$, and

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

Price Elasticity from a Non-Linear Demand Function

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.,

$$e_p = \frac{\delta Q}{\delta P} \cdot \frac{P}{Q} \qquad \ldots(2.10)$$

What we need to compute the price-elasticity coefficient is to find first the value of the first term, $\delta Q/\delta P$, i.e., the slope of the demand curve. The slope can be obtained by differentiating the demand function, Thus,

$$\frac{\delta Q}{\delta P} = -baP^{-b-1} \qquad \ldots(2.11)$$

By substituting Eq. (2.11) in Eq. (2.10), e_p can be expressed as

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

$$= \frac{-baP^{-b}}{Q} \qquad \ldots(2.12)$$

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

$$e_p = \frac{-baP^{-b}}{aP^{-b}} = -b \qquad \ldots(2.13)$$

Equation (2.13) 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.

NOTES

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 the price will increase the total revenue and if $e_q < 1$, then increasing the price will increase the 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 $TR = Q.P$, we need to know *P* and *Q*. This information can be obtained through the demand function. The demand function is given as:

$$Q = 100 - 5P$$

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

$$P = 20 - 0.2Q \qquad \ldots(2.14)$$

Given the price function, *TR* can be obtained as:

$$TR = P \,.\, Q = (20 - 0.2Q)Q = 20Q - 0.2Q^2$$

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

$$MR = \frac{\partial TR}{\partial Q} = 20 - 0.4Q$$

The *TR*-function is graphed in panel (*a*) and the demand and *MR* functions are presented in panel (*b*) of Fig. 2.10. 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. 2.10 that *TR* increases so long as $e > 1$; *TR* reaches its maximum level where $e = 1$; and it decreases when $e < 1$.

The relationship between price-elasticity and *TR* is summed up in Table 2.2. 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*.

In case of an *inelastic demand* (i.e., $e_p < 1$), quantity demanded increases by less than the proportionate decrease in price and hence the total revenue falls when price falls. The total revenue increases when price increases because quantity demanded decreases by less than the proportionate increase 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. 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, the total revenue increases when price falls and *vice versa.*

NOTES

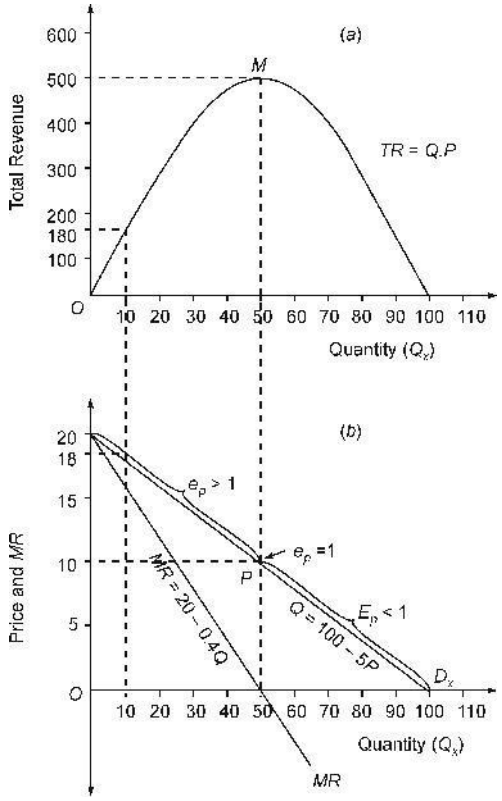


Fig. 2.10 *Price Elasticity and Total Revenue*

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 2.2 *Elasticity, Price-change and Change in TR*

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

*Subject to the size of the market.

Price Elasticity and Marginal Revenue

NOTES

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

Since $TR = P.Q$,

$$MR = \frac{\partial(P.Q)}{\partial Q} = P + Q\frac{\partial P}{\partial Q}$$

$$= P\left(1 + \frac{Q}{P}\cdot\frac{\partial P}{\partial Q}\right) \quad \ldots(2.15)$$

Note that $\frac{Q}{P}\cdot\frac{\partial P}{\partial Q}$ in Eq. (2.15) is the reciprocal of elasticity. That is,

$$\frac{Q}{P}\cdot\frac{\partial P}{\partial Q} = -\frac{1}{e_p}$$

By substituting $-\frac{1}{e}$ for $\frac{Q}{P}\cdot\frac{\partial P}{\partial Q}$ in Eq. (2.15), we get:

$$MR = P\left(1 - \frac{1}{e_p}\right) \quad \ldots(2.16)$$

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*. If $e < 1$, $MR < 0$, *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. (2.16), the formula for price elasticity (e_p) can be expressed in terms of *AR* and *MR*. We know that $P = AR$. So Eq. (2.16) can be written as:

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

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

By rearranging the terms, we get:

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

or

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

NOTES

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

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

Determinants of Price Elasticity of Demand

We have noted above that price-elasticity of 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 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 both these 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*

NOTES

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 TV sets 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 TV set 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, 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, $e < 1$.

2.6.2 Income Elasticity of Demand

Apart from the price of a product and its substitutes, consumer's income is another basic determinant of demand for a product. The relationship between quantity demanded and income is of positive nature, unlike the negative price-demand relationship. The demand for most goods and services increases with increase in consumer's income and *vice versa*. *The responsiveness of demand to the changes in income is known as income-elasticity of demand*.

Income-elasticity of demand for a product, say X, (i.e., e_y) may be defined as:

$$e_y = \frac{\frac{\Delta X_q}{X_q}}{\frac{\Delta Y}{Y}} = \frac{Y}{X_q} \cdot \frac{\Delta X_q}{\Delta Y} \qquad \ldots(2.17)$$

(where X_q = quantity of X demanded; Y = disposable income; ΔX_q = change in quantity of X demanded; and ΔY = change in income).

Obviously, the formula for measuring income-elasticity of demand is the same as that for measuring the price-elasticity. The only change in the formula is that the variable 'income' (Y) is substituted for the variable 'price' (P). Here, income refers to the disposable income, i.e., income net of taxes. All other formulae for measuring price-elasticities may by adopted to measure the income-elasticities, keeping in mind the difference between them and the purpose of measuring income-elasticity.

NOTES

To estimate income-elasticity, suppose, for example, government announces a 10 per cent dearness allowance to its employees. As a result average monthly salary of government employees increases from ` 20,000 to ` 22,000. Following the pay-hike, monthly petrol consumption of government employees increases from 150 litre per month to 165 litre. The income-elasticity of petrol consumption can now be worked out as follows. In this case, ΔY = ` 22,000 – ` 20,000 = ` 2,000, and ΔQ (oil demand) = 165 litre – 150 litre = 15 litre. By substituting those values in Eq. (2.17), we get:

$$e_y = \frac{20,000}{150} \times \frac{15}{2,000} = 1$$

It means that income elasticity of petrol consumption by government employees equals 1. It means that a one per cent increase in income results in a one per cent increase in petrol consumption.

Unlike price-elasticity of demand, which is always negative, income-elasticity of demand is always positive because of a positive relationship between income and quantity demanded of a product. But there is an *exception* to this rule. Income-elasticity of demand for an inferior good is negative, because of the inverse substitution effect. The demand for inferior goods decreases with increase in consumer's income. The reason is that when income increases, consumers switch over to the consumption of superior substitutes, i.e., 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 food grains; non-vegetarians buy more of meat and less of potato, and travellers travel more by plane and less by train.

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 of the three categories, viz., necessities, comforts and luxuries have different elasticities. The general pattern of income-elasticities of different goods for increase in income and their effect on sales are given in Table 2.3.

Table 2.3 *Income-Elasticities*

	Consumer goods	*Co-efficient of income-elasticity*	*Effect on sales with change in income*
1.	Essential goods	Less than one ($e_y < 1$)	Less than proportionate change in sale
2.	Comforts	Almost equal to unity ($e_y \cong 1$)	Almost proportionate change in sale
3.	Luxuries	Greater than unity ($e_y > 1$)	More than proportionate increase in sale

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

Uses of Income-elasticity in Business Decisions

While price and cross elasticities of demand are of greater significance in the pricing of a product aimed at maximizing the total revenue in the short run, income-elasticity of a product is of a greater significance in production planning and management in the long run, particularly during the period of a business cycle. The concept of income-elasticity can be used in estimating future demand provided that 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 a change in personal incomes is expected, other things remaining the same. It also helps in avoiding over-production or under-production.

NOTES

In forecasting demand, however, only the relevant concept of income and data should be used. It is generally believed that the demand for goods and services increases with increase in GNP, depending on the marginal propensity to consume. This may be true in the context of aggregate national demand, but not necessarily for a particular product. It is quite likely that increase in GNP flows to a section of consumers who do not consume the product in which a businessman is interested. For instance, if the major proportion of incremental GNP goes to those who can afford a car, the growth rate in GNP should not be used to calculate income-elasticity of demand for bicycles. Therefore, the income of only a relevant class or income-group should be used. Similarly, where the product is of a regional nature, or if there is a regional division of market between the producers, the income of only the relevant region should be used in forecasting the demand.

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 'normal goods'. On the other hand, goods whose income-elasticities are negative beyond a certain level of income are termed 'inferior goods'.

3.7 SUMMARY

In this unit, you have learnt that,

3.7.1 The consumers demand a commodity because they derive or expect to derive utility from that commodity. The expected utility from a commodity is the basis of demand for it.

3.7.2 The concept of utility can be looked upon from two angles—from the commodity angle and from the consumer's angle. From the commodity angle, utility is the want-satisfying property of a commodity. From the consumer's angle, utility is the psychological feeling of satisfaction, pleasure, happiness or well-being which a consumer derives from the consumption, possession or the use of a commodity.

3.7.3 Some early psychological experiments on an individual's responses to various stimuli led classical and neo-classical economists to believe that utility is measurable and cardinally quantifiable. This belief gave rise to the concept of cardinal utility.

3.7.4 The law of diminishing marginal utility is one of the fundamental laws of economics. This law states that as the quantity

Check Your Progress

9. Define price elasticity of demand.
10. What is point elasticity?
11. How is the closeness of the substitutes related to the elasticity of demand for the commodity?
12. Why is income-elasticity of demand always positive?

consumed of a commodity increases, the utility derived from each successive unit decreases, consumption of all othercommodities remaining the same.

NOTES

- The consumer has a limited income and goods he consumes have a price. Limited income and prices impose constraints on consumer's choice, called budgetary constraints.
- A consumer attains his equilibrium when he maximizes his total utility, given his income and market prices of the goods and services that he consumes. The ordinal utility approach specifies two conditions for the consumer's equilibrium:
  - o Necessary or the first order condition
  - o Supplementary or the second order condition
- The central theme of the consumption theory—be it based on ordinal utility or cardinal utility approach—is the utility maximizing behaviour of the consumer.
- Constrained utility maximization is a process wherein under certain constraints the highest possible level of utility is obtained through the consumption of goods and services. This happens when the highest overall level of utility cannot be obtained.
- When the demand curve changes its position—retaining its slope though not necessarily—the change is known as a shift in the demand curve.
- Shifts in a price-demand curve may take place owing to the change in one or more non-price determinants of the demand for a commodity.
- Increase and decrease in demand are associated with non-price-quantity relationships of demand whereas extension and contraction of demand are associated with the price-quantity relationship of demand.
- The degree of responsiveness of demand to the change in its determinants is called elasticity of demand.
- 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.
- 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 as a result of one per cent change in the price of the commodity.
- The measure of elasticity of demand between any two finite points on a demand curve is known as arc elasticity.
- 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.
- 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.
- 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.
- Apart from the price of a product and its substitutes, consumer's income is another basic determinant of demand for a product.
- The responsiveness of demand to the changes in income is known as income-elasticity of demand.

- Unlike price-elasticity of demand, which is always negative, income-elasticity of demand is always positive because of a positive relationship between income and quantity demanded of a product.

NOTES

3.8 KEY TERMS

- **Utility:** It is the psychological feeling of satisfaction, pleasure, happiness or well-being which a consumer derives from the consumption, possession or the use of a commodity.
- **Shift in the demand curve:** When the demand curve changes its position—retaining its slope though not necessarily—the change is known as a shift in the demand curve.
- **Elasticity of demand:** The degree of responsiveness of demand to the change in its determinants is called elasticity of demand.
- **Price elasticity of demand:** It is generally defined as the responsiveness or sensitiveness of demand for a commodity to the changes in its price.
- **Arc elasticity:** The measure of elasticity of demand between any two finite points on a demand curve is known as arc elasticity.
- **Income-elasticity of demand:** The responsiveness of demand to the changes in income is known as income-elasticity of demand.

3.9 ANSWERS TO 'CHECK YOUR PROGRESS'

1. An important attribute of the 'absolute' concept of utility is that it is ethically neutral because a commodity may satisfy a frivolous or socially immoral need, e.g., alcohol, drugs or a profession like prostitution.
2. Some early psychological experiments on an individual's responses to various stimuli led classical and neo-classical economists to believe that utility is measurable and cardinally quantifiable. This belief gave rise to the concept of cardinal utility.
3. 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 or Marshallian approach.
   (ii) Ordinal Utility Approach, pioneered by J. R. Hicks, a Nobel laureate and R. G. D. Allen, is also called Hicks-Allen approach or the Indifference Curve Analysis.
4. The assumptions of the law of diminishing marginal utility are listed below.
   - First, the unit of the consumer good must be a standard one, 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, the consumer's taste or preference must remain the same during the period of consumption.
5. A consumer attains his equilibrium when he maximizes his total utility, given his income and market prices of the goods and services that he consumes.

NOTES

6. The central theme of the consumption theory—be it based on ordinal utility or cardinal utility approach—is the utility maximizing behaviour of the consumer.
7. When the demand curve changes its position—retaining its slope though not necessarily—the change is known as a shift in the demand curve.
8. Shifts in a price-demand curve may take place owing to the change in one or more non-price determinants of the demand for a commodity.
9. Price elasticity of demand is generally defined as the responsiveness or sensitiveness of demand for a commodity to the changes in its price.
10. 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.
11. 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.
12. Unlike price-elasticity of demand, which is always negative, income-elasticity of demand is always positive because of a positive relationship between income and quantity demanded of a product.

3.10 QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is utility? What is the difference between the two concepts of utility?
2. Is utility a measurable concept? Justify.
3. What does the law of diminishing marginal utility state?
4. What are budgetary constraints? What is consumer equilibrium? How is it attained?
5. Write a note on constrained utility maximization.
6. What are the factors behind shifts in the demand curve?
7. 'The concept of elasticity of demand plays a crucial role in business-decisions regarding manoeuvring of prices with a view to making larger profits.' Describe.
8. What are the problems in using arc elasticity? What are the determinants of price elasticity of demand?
9. Define and distinguish between:
   (a) Arc elasticity and point elasticity
   (b) Price elasticity and cross-elasticity, and
   (c) Income elasticity and price elasticity.

Long-Answer Questions

1. Discuss utility as a basis of consumer demand.
2. Describe the two approaches to consumer demand analysis and the measurability of utility.
3. Explain the law of diminishing marginal utility and its necessary conditions.

4. With regard to budgetary constraints, assess the income of the consumer and the budget line.
5. Critically analyse the concept of consumer equilibrium and constrained utility maximization.
6. Evaluate the demand curve and the factors shifting the demand curve.
7. What is elasticity of demand? Discuss price elasticity of demand in detail.
8. Explain the income elasticity of demand and the use of income-elasticity in business decisions.

NOTES

3.11 FURTHER READING

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.

Keating, B. and J. H. Wilson. 2003. *Managerial Economics: An Economic Foundation for Business Decisions*, 2nd Edition. New Delhi: Biztantra.

Mansfield, E., W. B. Allen, N. A. Doherty and K. Weigelt. 2002. *Managerial Economics: Theory, Applications and Cases*, 5th Edition. NY: W. Orton & Co.

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

Salvantore, Dominick. 2001. *Managerial Economics in a Global Economy*, 4th Edition. Australia: Thomson-South Western.

Thomas, Christopher R. and Maurice S. Charles. 2005. *Managerial Economics: Concepts and Applications*, 8th Edition. New Delhi: Tata McGraw-Hill.

Adhikary, M. 2000. *Business Economics*. New Delhi: Excel Books.

Baumol, W. J. 1996. *Economic Theory and Operations Analysis,* 3rd edition. New Jersey: Prentice-Hall.

Chopra, O. P. 1985. *Managerial Economics.* New Delhi: Tata McGraw-Hill.

Kautsoyiannis, A. 1991. *Modern Microeconomics*. New York: Macmillan.

UNIT IV SUPPLY ANALYSIS

NOTES

Structure

4.0 INTRODUCTION

We have so far been concerned with the demand side of the market. In this unit, we turn to the supply side of the market. Supply of a product depends on the production of the commodity. Production of the commodity depends on the availability of inputs (labour and capital) and technology. Given the technology, output of a product depends on the relationship between the output and input. The input-output relationship is brought out by the theory of production.

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.

NOTES

4.1 UNIT OBJECTIVES

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

- Discuss production function as a tool of analysis
- Explain the returns to a factor and the laws of returns to scale
- Assess the cost-output relations through cost curves
- Discuss the problems of cost minimization in managerial decision-making
- Analyse the importance of profit maximization assumption
- Evaluate the supply curve and the shift in the supply curve
- Assess the determinants of the price elasticity of supply

4.2 PRODUCTION FUNCTION

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?

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. Besides, 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:

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

NOTES

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'. 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.

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:

$$Q = f(K, L) \qquad ...(3.1)$$

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

The production function (3.1) 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.

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:

$$Q = f(L) \qquad ...(3.2)$$

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

$$Q = f(K, L) \qquad ...(3.3)$$

NOTES

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

(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.

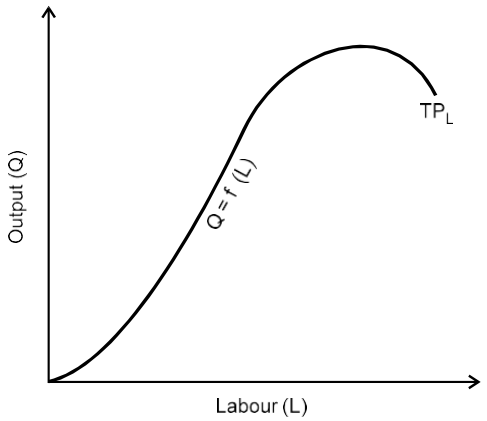


Fig. 3.1 *Single Variable Production Function*

The production function (3.2) may be graphically presented, as shown in Fig. 3.1, 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.

4.2.1 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 the marginal product (*MP*) curve.

From the production function (3.2), 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 (∂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,

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

or for a large change in L, $\frac{\Delta Q}{\Delta L}$ ...(3.4)

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. 3.2.

NOTES

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.

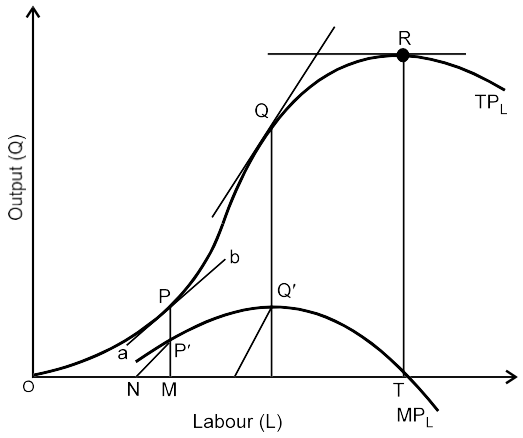


Fig. 3.2 *Derivation of MP_L Curve*

4.2.2 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:

$$AP_L = \frac{Q}{L} = \frac{f(L)}{L} \qquad ...(3.5)$$

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. 3.3.

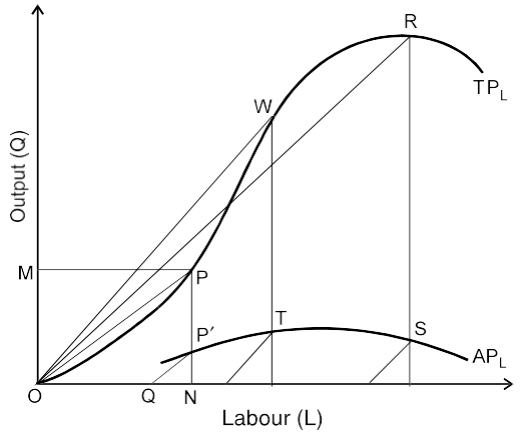


Fig. 3.3 *Derivation of AP_L Curve*

NOTES

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:

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

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.

4.3 RETURNS TO FACTOR AND RETURNSTO SCALE

Returns to a factor relate to the short-period production function when one factor is varied keeping the other factor fixed in order to have more output, the marginal returns or marginal product of the variable factor diminishes.

This relates to the Law of Variable Proportions. On the other hand, returns to scale relate to the long-period production function when a firm changes its scale to production by changing one or more of its factors. This refers to the Law of Returns to Scale.

4.3.1 Short-Run Laws of Production

Production with one Variable Input

Some factors of production have elastic supply even during the short period. Such factors are called **variable factors.** In the short-run, therefore, the firms can employ a large 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 combination leads to *variation in factor proportions.* The laws which bring out the relationship between varying factor proportions and output are therefore known also as the *Laws of Returns to a Variable proportions*. This law is more popularly known as the *Law of Diminishing Returns.* In this section, we explain the laws of returns to variable input.

The Law of Returns to a Variable Input: The Law of Diminishing Returns

The law of diminishing returns states that *when 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 and then at a constant rate but it will eventually increase at diminishing rates.* In other words, when a firm using two inputs—labour and capital—increases the number of labour, capital remaining constant, the marginal productivity of labour may initially increase, but it does decrease eventually. This is called the law of diminishing returns to the variable input.

Check Your Progress

1. In economics, what does the term production mean?
2. State the reason behind the reduction of number of variables used in a production function.
3. Why does the concept of marginal product play an important role in explaining the laws of returns?

NOTES

Assumptions. The law of diminishing returns is based on the following assumptions: (*i*) the state of technology is given, (*ii*) labour is homogeneous, and (*iii*) input prices—wages and interest—are given.

To illustrate the law of diminishing returns, let us assume (*i*) that the coal-mining firm (in our earlier example) has a set of mining machinery as its capital (*K*), fixed in the short run, and (*ii*) that it can employ more of mine-workers to increase its coal production. Thus, the short run production function for the firm will take the following form.

$$Q_c = f(L)$$

Let us assume that the labour-output relationship in coal production based on actual data is given by a hypothetical production function of the following form.

$$Q_c = -L^3 + 15L^2 + 10L \qquad \ldots(3.6)$$

Given the production function (3.6), we may substitute different numerical values for L in the function and work out a series of Q_c, i.e., the quantity of coal (say, thousand tonnes) that can be produced with different number of workers. For example, if $L = 5$, then by substitution,

$$\begin{aligned} Q_c &= -5^3 + 15 \times 5^2 + 10 \times 5 \\ &= -125 + 375 + 50 \\ &= 300 \text{ (thousand tonnes)} \end{aligned}$$

A tabular array of output levels associated with different number of workers from 1 to 12, in our hypothetical coal-production example, is given in Table 3.1 (Cols. 1 and 2).

What we need now is to work out *marginal productivity of labour* (MP_L) to find the trend in the contribution of the marginal labour and *average productivity of labour* (AP_L) to find the average contribution of labour.

Tables 3.1 *Three Stages of Production*

No. of workers (N)	*Total product (TP_L) (000 tonnes)*	*Marginal Product* (MP_L)*	*Average Product (AP_L)*	*Stages of Production*
(1)	*(2)*	*(3)*	*(4)*	*(5)*
1	24	24	24	I
2	72	48	36	Increasing
3	138	66	46	returns
4	216	78	54	
5	300	84	60	
6	384	84	64	
7	462	78	66	II
8	528	66	66	Diminishing
9	576	48	64	returns
10	600	24	60	
11	594	– 6	54	III
12	552	– 42	46	Negative returns

*$MP_L = TP_n - TP_{n-1}$. MP_L calculated by differential method will be different from that given in Col. 3.

NOTES

Marginal Productivity of Labour (MP_L) can be obtained by differentiating the production function (3.6). Thus,

$$MP_L = \frac{\partial Q}{\partial L} = -3L^2 + 30L + 10 \quad \ldots(3.7)$$

By substituting numerical value for labour (L) in Eq. (3.7), MP_L can be obtained at different levels of labour employment. However, this method can be used only where labour is perfectly divisible and δL 0. Since, in our example, each unit of $L = 1$, calculus method cannot be used.

Alternatively, where labour can be increased at least by one unit, i.e., $\Delta L = 1$, MP_L can be obtained as

$$MP_L = TP_L - TP_{L-1}$$

The MP_L worked out by this method is presented in col. 3 of Table 3.1.

Average Productivity of labour (AP_L) can be obtained by dividing the production function by L. Thus,

$$AP_L = \frac{-L^3 + 15L^2 + 10L}{L} = -L^2 + 15L + 10 \quad \ldots(3.8)$$

Now AP_L can be obtained by substituting the numerical value for L in Eq. (3.8). AP_L obtained by this method is given in col. 4 of Table 3.1.

The information contained in Table 3.1 is presented graphically in panels (*a*) and (*b*) of Fig. 3.4. Panel (*a*) of Fig. 3.4 presents the total product curve (TP_L) and panel (*b*) presents marginal product (MP_L) and average product (AP_L) curves. The TP_L schedule demonstrates the law of diminishing returns. As the curve TP_L shows, the total output increases at an increasing rate till the employment of the 5th worker, as indicated by the increasing slope of the TP_L curve. (See also col. 3 of the table.) Beyond the 6th worker, TP_L increases (until the 10th worker) but the rate of increase in TP_L (i.e., marginal addition to TP_L) begins to fall and turns negative 11th worker onwards. This shows the operation of the law of diminishing returns, i.e., as units of labour increase, capital remaining constant, the total output increases initially at an increasing rate but finally at a decreasing rate.

Three Stages in Production

Table 3.1 and Fig. 3.4 present the three usual stages in the application of the laws of diminishing returns. In Stage I, TP_L increases at increasing rate. This is indicated by the rising MP_L till the employment of the 5th worker. Given the production function (Eq. 3.6), the 6th worker produces as much as the 5th worker. The output from the 5th and the 6th workers represents an intermediate stage of constant returns to the variable factor, labour.

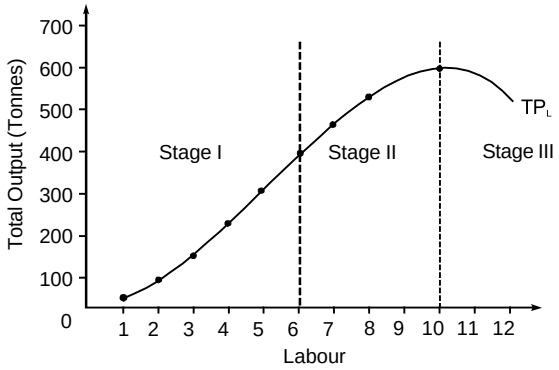
In Stage II, TP_L continues to increase but at diminishing rates, *i.e.*, MP_L begins to decline. This stage in production shows the law of diminishing returns to the variable factor. Total output reaches its maximum level at the employment of the 10th worker. Beyond this level of labour employment, TP_L begins to decline. This marks the beginning of Stage III in production.

To conclude, given the employment of fixed factor (capital), when more and more workers are employed, the return from the additional worker may initially increase but will eventually decrease.

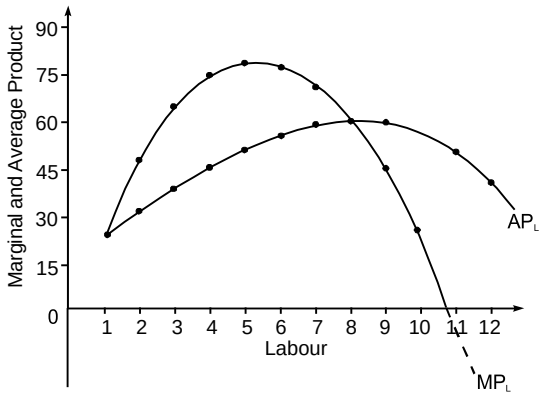
Factors Behind the Laws of Returns

As shown in Fig. 3.4, the marginal productivity of workers (MP_L) increases in Stage I and it decreases in Stage II. Stage I shows the Law of Increasing Returns and Stage II shows the Law of Diminishing Returns.

NOTES



(a)



(b)

Fig. 3.4 *Total, Average and Marginal Products*

The reasons which underly the application of the laws of returns in Stages I and II may be described as follows. One of the important factors causing increasing returns to a variable factor is the ***indivisibility of fixed factor*** (capital). It results in under-utilisation of capital if labour is less than its optimum number. Let us suppose that optimum capital-labour combination is 1:6. If capital is indivisible and less than 6 workers are employed, then capital would remain underutilised. When more and more workers are added, utilization of capital increases and also the productivity of additional worker. Another reason for increase in labour productivity is that employment of additional workers leads 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 amounts to substitution of capital with labour. But, technically, one factor can substitute another only upto a limited extent. Therefore, with increase in labour, capital per unit of labour decreases. This causes decrease in the productivity of the marginal labour. That is, employment of more and more labour against a given capital, causes decrease in MP_L.

NOTES

Empirical Validity of the Law of Diminishing Returns. The law of diminishing returns is *an empirical law,* frequently observed in various production activities. This law, however, may not apply universally to all kinds of productive activities since the law is not as true as the law of gravitation. 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. This law has been found to operate in agricultural production more regularly than in industrial production. The reason is, in agriculture, natural factors play a predominant role whereas man-made factors play the major role in industrial production. Despite these variations and limitations of the law, if increasing units of an input are applied to the fixed factors, the marginal returns to the variable input decrease eventually.

The Law of Diminishing Returns and Business Decisions

We have discussed above the law of diminishing returns in a theoretical framework. Let us now look at the applicability of this law to business decision-making.

The law of diminishing returns as presented graphically has a relevance to the business decisions. The graph can help in identifying the rational and irrational stages of operations. It can also provide answer to such questions as (*i*) how much labour to employ to maximise the output; and (*ii*) what number of workers to apply to a given fixed input so that per unit cost in minimized when output is maximumized. Fig. 3.4 exhibits the three stages of production. In Stage III, has a very high labour-capital ratio. As a result, employment of additional workers proves not only unproductive but also causes a decline in the *TP*. Similarly, in Stage I, capital is presumably underutilized. So a firm operating in Stage I is required to increase labour, and a firm operating in Stage III is required to reduce labour, with a view to maximising its total production. From the firm's point of view, setting an output target in Stages I and III is irrational. The only meaningful and rational stage from the firm's point of view is Stage II in which the firm can find answer to the questions 'how many workers to employ'.

Figure 3.4 shows also that the firm should employ a minimum of 6 workers and a maximum of 10 workers even if labour is available free of cost. This means that the firm has a limited choice ranging from 6 to 10 workers. How many workers to employ against the fixed capital and how much to produce can be answered, only when the price of labour, i.e., wage rate, and that of the product are known. This question is answered below.

Optimum Employment of Labour

It may be recalled from Fig. 3.4 that an output maximizing coal-mining firm would like to employ 10 workers—since at this level of employment, the output is maximum and $MP_L = 0$. The firm can, however, employ 10 workers only if workers are available free of cost. But labour is not available free of cost—the firm is required to pay wages to the workers. Therefore, the question arises 'how many workers should the firm employ—10 or less or more than 10—to maximise its profit. A simple answer to this question is that the number of workers to be employed depends on the output that maximizes firm's profit, given the product price and the wage rate. This point can be proved as follows.

As a rule, total profit is maximum at the level of output at which

$MC = MR$

In our example here, since labour is the only variable input, marginal cost (*MC*) equals marginal wages (*MW*), i.e., $MC = MW$.

NOTES

As regards *MR*, in case of factor employment, the concept of ***Marginal Revenue Productivity*** is used. The marginal revenue productivity of labour is the money value of marginal product of labour (MP_L). In specific terms, marginal revenue productivity of labour (MRP_L) equals marginal physical productivity (MP_L) of labour multiplied by the price (*P*) of the product, i.e.,

$$MRP_L = MP_L \times P$$

For example, suppose that the price (*P*) of coal is given at ` 10 per quintal. Now, $MRP_L$ can be known by multiplying its $MP_L$ (as given in Table 3.1) by ` 10. For example, MRP_L of the 3rd worker (see Table 3.1) equals 66 □ 10 = ` 660 and of the 4th worker, 78 □ 10 = 780. Likewise, if whole column ($MP_L$) is multiplied by ` 10, it gives us a table showing the number of workers and the marginal revenue productivity of workers. Let us suppose that wage rate (per time unit) is given at ` 660. Given the wage rate, the profit maximising firm will employ only 8 workers because at this employment, $MRP_L$ = wage rate = $MRP_L$ of 8th worker = 66 □ 10 = ` 660. If the firm employs the 9th worker, MRP_L = 48 □ 10 = ` 480 < ` 660. Clearly, the firm loses ` 180 on the 9th worker. And, if the firm employs less than 8 workers, it will not maximize its profit.

To generalize, if relevant series of MRP_L is graphed, it gives the MRP_L curve as shown in Fig. 3.5. Similarly, the MRP_L curve for any input may be drawn and compared with *MC* (or *MW*) curve. Labour being the only variable input, in our example, let us suppose that wage rate in the labour market is given at *OW* (Fig. 3.5). When wage rate remains constant, average wage (*AW*) and marginal wage (*MW*) are equal, i.e., *AW* = *MW*, for the whole range of employment in the short run. When *AW* = *MW*, a large number of labour can be employed at wage rate *OW* as shown by a straight horizontal line marked *AW* = *MW*.

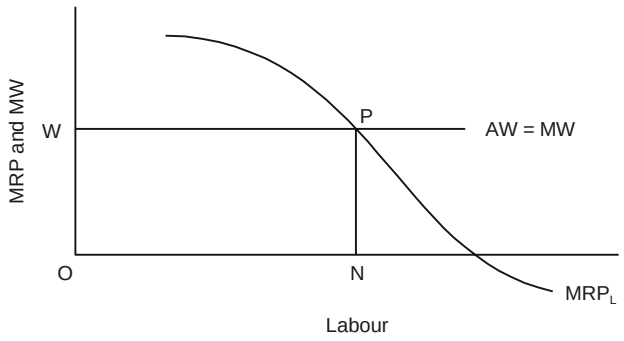


Fig. 3.5 *Determination of Labour Employment in the Short-Run*

With the introduction of *MRP* curve and *AW* = *MW* line (Fig. 3.5), a profit maximising firm can easily find the maximum number of workers which can be optimally employed against a fixed quantity of capital. Once the maximum number of workers is determined, the optimum quantity of the product is automatically determined.

The marginality principle of profit maximization tells that profit is maximum where *MR* = *MC*. This is a necessary condition of profit maximisation. Fig. 3.5 shows that MRP_L = *MW* (= *MC*) are equal at point *P*, the point of intersection between MRP_L and *AW* = *MW*. The number of workers corresponding to this point is *ON*. A profit maximising firm should therefore employ only *ON* workers. Given the number of workers, the total output can be known by multiplying *ON* with average labour productivity (*AP*).

NOTES

4.3.2 Returns to Scale

The laws of returns to scale state the behaviour of output in response to a proportional and simultaneous change in inputs. Increasing inputs proportionately and simultaneously is, in fact, an expansion of the scale of production.

When a firm expands its scale, i.e., it increases both the inputs proportionately, then there are three technical possibilities:

- Total output may increase more than proportionately
- Total output may increase proportionately
- Total output may increase less than proportionately Accordingly, there are three kinds of returns to scale
  (i) Increasing returns to scale
  (ii) Constant returns to scale
  (iii) Diminishing returns to scale

So far as the sequence of the laws of 'returns to scale' is concerned, the law of increasing returns to scale is followed by the law of constant and then by the law of diminishing returns to scale. This is the most common sequence of the laws of returns to scale.

Let us now explain the laws of returns to scale with the help of isoquants for a two-input and single output production system.

Increasing Returns to Scale

When a certain proportionate increase in both the inputs, *K* and *L*, leads to a more than proportionate increase in output, it exhibits *increasing returns to scale.* For example, if both the inputs, *K* and *L*, are successively doubled and the corresponding output is more than doubled, the returns to scale is said to be increasing. The increasing returns to scale is illustrated in Fig. 3.6. The movement from point *a* to *b* on the line *OB* means doubling the inputs. It can be seen in Fig. 3.6 that the combination of inputs *L* and *K*, increases from 1*K* + 1*L* to 2*K* + 2*L*. As a result of doubling the inputs, output is more than doubled: it increases from 10 to 25 units, *i.e.*, a 100 per cent increase in inputs results in 120 per cent increase in output. Similarly, the movement from point *b* to point *c* indicates a 50 per cent increase in inputs as a result of which the output increases from 25 units to 50 units, i.e., by 200 per cent. This kind of relationship between the inputs and output shows *increasing returns to scale*.

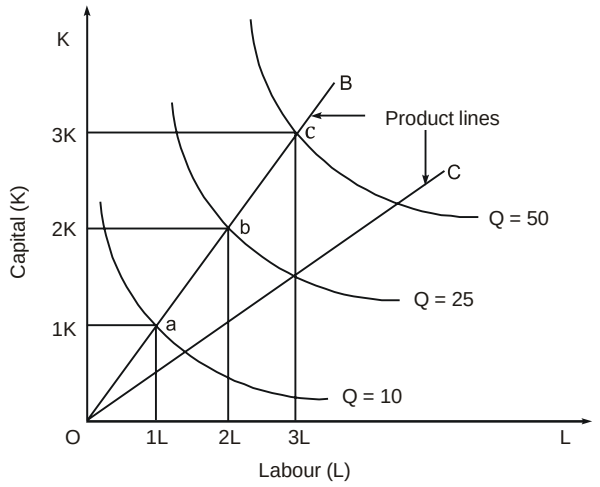


Fig. 3.6 *Increasing Returns to Scale*

Factors Leading to Increasing Returns to Scale: There are at least three plausible reasons for increasing returns to scale, called *economics of scale*.

(i) **Indivisibility of machinery and managerial manpower:** Certain inputs, particularly mechanical equipment and managerial manpower, used in the process of production are available in a given size. Such inputs cannot be divided into parts to suit small scale of production. For example, half a turbine cannot be used to produce electricity and one-third of a composite harvester and earth-movers cannot be used productively. Similarly, half of a production manager cannot be employed, if part-time employment is not acceptable to the manager. Because of indivisibility of machinery and managers, given the state of technology, they have to be employed in a minimum quantity even if scale of production is much less than the capacity output. Therefore, when scale of production is expanded by increasing all the inputs, the productivity of indivisible factors increases exponentially because of technological advantage. This results in increasing returns to scale.

(ii) **Higher degree of specialization:** Another factor causing increasing returns to scale is higher degree of specialization of labour, manager and machinery, which becomes possible with increase in scale of production. The use of specialized labour suitable to job needs and composite machinery increases productivity per unit of inputs. Their cumulative effects contribute to the increasing returns to scale. Besides, employment of specialized managerial personnel, e.g., administrative manager, production managers, sales manager and personnel manager, contributes a great deal in increasing production.

(iii) **Dimensional relations:** Increasing returns to scale is also a matter of dimensional relations. For example, when the length and breadth of a room ($15' \times 10' = 150$ sq. ft.) are doubled, then the size of the room is more than doubled. It increases to $30' \times 20' = 600$ sq. ft. which is more than double the room size. Similarly, when diameter of a pipe is doubled, the flow of water is more than doubled. In accordance with this dimensional relationship, when the labour and capital are doubled, the output is more than doubled and so on.

Constant Returns to Scale

When the increase in output is proportional to the increase in inputs, it exhibits *constant returns to scale.* For example, if both the inputs, *K* and *L*, are doubled subsequently and output is also doubled, subseqently then the returns to scale are said to be constant. Constant returns to scale are illustrated in Fig. 3.7. The lines *OA* and *OB* are 'product lines' indicating two hypothetical techniques of production. The isoquants marked $Q = 10$, $Q = 20$ and $Q = 30$ indicate the three different levels of output. In the figure, the movement from points *a* to *b* indicates doubling both the inputs. When inputs are doubled, output is also doubled, i.e., the output increases from 10 to 20, i.e., a 50 per cent increase in output.

Similarly, movement from point *b* to *c* indicates a 50 per cent increase in labour as well as capital. This increase in inputs results in an increase of output from 20 to 30 units, i.e., a 50 per cent increase in output. In simple words, a 50 per cent increase in inputs leads to a 50 per cent increase in output. This relationship between the proportionate change in inputs and proportional change in output may be summed up as follows:

$$1K + 1L \Rightarrow 10$$
$$2K + 2L \Rightarrow 20$$
$$3K + 3L \Rightarrow 30$$

This relationship between inputs and output exhibits *constant returns* to *scale*.

NOTES

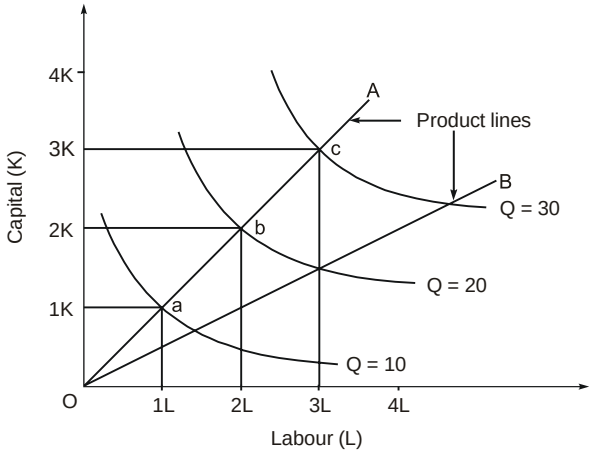


Fig. 3.7 *Constant Returns to Scale*

The constant returns to scale are attributed to the limits of the economies of scale: With expansion in the scale of production, economies arise from such factors as indivisibility of fixed factors, greater possibility of specialization of capital and labour, use of labour-saving techniques of production, etc. But there is a limit to the economies of scale. When economies of scale reach their limits and diseconomies are yet to begin, returns to scale become constant. The constant returns to scale also take place where factors of production are perfectly divisible and where technology is such that capital-labour ratio is fixed. When the factors of production are perfectly divisible, the production function is homogeneous of degree 1 showing constant returns to scale.

Decreasing Returns to Scale

The firms are faced with *decreasing returns to scale* when a certain proportionate increase in inputs, *K* and *L*, leads to a less than proportional increase in output. For example, when inputs are doubled and output is less than doubled, then decreasing returns to scale is in operation. The decreasing returns to scale is illustrated in Fig. 3.8. As the figure shows, when the inputs *K* and *L* are doubled, i.e., where capital-labour combination is increased from 1*K* + 1*L* to 2*K* + 2*L*, the output increases from 10 to 18 units, which is less that the proportionate increase. The movement from point *b* to *c* indicates a 50 per cent increase in the inputs. But, the output increases by only 33.3 per cent. This exhibits *decreasing returns to scale*.

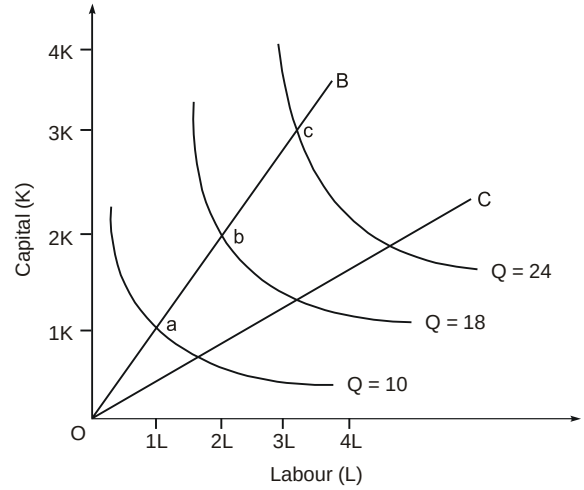


Fig. 3.8 *Decreasing Returns to Scale*

Causes of Diminishing Returns to Scale: The decreasing returns to scale are attributed to the diseconomies of scale. The most important factor causing diminishing returns to scale is 'the diminishing return to management', i.e., managerial diseconomies. As the size of the firms expands, managerial efficiency decreases. Another factor responsible for diminishing returns to scale is the limitedness or exhaustibility of the natural resources. For example, doubling of coal mining plant may not double the coal output because of limitedness of coal deposits or difficult accessibility to coal deposits. Similarly, doubling the fishing fleet may not double the fish output because availability of fish may decrease in the ocean when fishing is carried out on an increased scale.

NOTES

4.4 COST CONCEPTS

Business decisions are generally taken on the basis of money values of the inputs and outputs. *Inputs multiplied by their respective prices and added together give the money value of the inputs*, i.e., *the cost of production.* The cost of production is an important factor in almost all business analysis and business decision-making, specially those pertaining to: (*a*) locating the weak points in production management; (*b*) minimizing the cost; (*c*) finding the optimum level of output; (*d*) determining price and dealers, margin; and (*e*) estimating or projecting the cost of business operation. Also, cost analysis assumes a great significance in all major business decisions because the term 'cost' has different meanings under different settings and is subject to varying interpretations. It is, therefore, essential that only the relevant concept of costs is used in the business decisions.

The cost concepts that are relevant to business operations and decisions can be grouped on the basis of their nature and purpose under two overlapping categories: (*i*) cost concepts used for accounting purposes, and (*ii*) analytical cost concepts used in economic analysis of business activities.

Accounting Cost

1. **Opportunity cost and actual cost:** The opportunity cost is the opportunity lost. An opportunity to make income is lost because of scarcity of resources like land, labour, capital, etc. We know that resources available to any person, firm or society are scarce but have alternative uses with different returns. Income maximizing resource owners put their scarce resources to their most productive use and thus, they forego the income expected from the second best use of the resources. Thus, the *opportunity cost* may be defined as the expected returns from the second best use of the resources that are foregone due to the scarcity of resources. The opportunity cost is also called *alternative cost.* Had the resource available to a person, a firm or a society been unlimited, there would be no opportunity cost.

   To explain and illustrate the concept of opportunity cost, suppose a firm has a sum of ` 100,000 for which it has only two alternative uses. It can buy either a printing machine or a photo copier, both having a productive life of 10 years. From the printing machine, the firm expects an annual income of ` 20,000 and from the photo copier, ` 15,000. A profit maximizing firm would invest its money in the printing machine and forego the expected income from the photo copier. The opportunity cost of the income from printing machine is the expected income from the photo copier, i.e., ` 15,000. In assessing the alternative cost, both explicit and implicit costs are taken into account.

Check Your Progress

4. State the law of diminishing returns.
5. What are the three kinds of returns to scale?
6. State one factor causing increasing returns to scale.

NOTES

Associated with the concept of opportunity cost is the concept of ***economic rent*** or ***economic profit.*** In our example of expected earnings firm printing machine, the *economic rent* of the printing machine is the excess of its earning over the income expected from the photo copier. That is, economic rent of the printing machine equals ` 20,000 – ` 15,000 = ` 5,000. The implication of this concept for a businessman is that investing in the 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, it will be helpful in choosing the best investment avenue.

2. **Business costs and full costs:** *Business costs* include all the expenses that are incurred to carry out a business. The concept of business costs is similar to the actual or real costs. Business costs include all the payments and contractual obligations made by the firm together with the book cost of depreciation on plant and equipment. Business costs are used for calculating business profits and losses and for filing returns for income-tax and also for other legal purposes.

   The concept of ***full cost***, includes business costs, opportunity cost and normal profit. The opportunity cost includes the foregone expected earning from the second best use of the resources, or the market rate of interest on the total money capital and also the value of an entrepreneur's own services that are not charged for in the current business. Normal profit is a necessary minimum earning in addition to the opportunity cost, which a firm must receive to remain in its present occupation.

3. **Actual or explicit costs and implicit or imputed costs:** The ***actual*** or ***explicit costs*** are those which are actually incurred by the firm in payment for labour, material, plant, building, machinery, equipment, travelling and transport, advertisement, etc. The total money expenses, recorded in the books of accounts are, for all practical purposes, the *actual costs*. In our example, the cost of printing machine, i.e., ` 100,000 is the actual cost. Actual cost comes under the accounting cost concept.

   In contrast to explicit costs, there are certain other costs that 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***. Opportunity cost is an important example of implicit cost. For example, suppose an entrepreneur does not utilize his services in his own business and works as a manager in some other firm on a salary basis. If he sets up his own business, he foregoes his salary as manager. This loss of salary is the opportunity cost of income from his own business. This is an implicit cost of his own business. Thus, implicit wages, rent, and implicit interest are the wages, rent and interest that an owner's labour, building and capital respectively, can earn from their second best use.

   Implicit costs are not taken into account while calculating the loss or gains of the business, but they form an important consideration in deciding whether or not to retain a factor in its present use. The explicit and implicit costs together make the ***economic cost.***

4. **Out-of-pocket and book costs:** The items of expenditure that involve cash payments or cash transfers, both recurring and non-recurring, are known as ***out-of-pocket costs***. All the explicit costs (e.g., wages, rent, interest, cost of materials and maintenance, transport expenditure, electricity and telephone expenses, etc.)

fall in this category. On the contrary, there are certain actual business costs that do not involve cash payments, but a provision is therefore made in the books of account and they are taken into account while finalizing the profit and loss accounts. Such expenses are known as ***book costs***. In a way, these are payments made by a firm to itself. Depreciation allowances and unpaid interest on the owner's own funds are the example of *book costs*.

Analytical Cost

1. **Fixed and variable costs:** ***Fixed costs*** are those that are fixed in volume for a certain quantity of output. Fixed cost does not vary with variation in the output between zero and a certain level of output. In other words, costs that do not vary or are fixed for a certain level of output are known as *fixed costs*. The fixed costs include: (*i*) costs of managerial and administrative staff, (*ii*) depreciation of machinery, building and other fixed assets, (*iii*) maintenance of land, etc. The concept of fixed cost is associated with the short-run.

   Variable costs are those which vary with the variation in the total output. Variable costs include cost of raw material, running cost of fixed capital, such as fuel, repairs, routine maintenance expenditure, direct labour charges associated with the level of output, and the costs of all other inputs that vary with output.

2. **Total, average and marginal costs:** ***Total cost* (*TC*)** is the total actual cost incurred on the production of goods and service. It refers to the total outlays of money expenditure, both explicit and implicit, on the resources used to produce a given level of output. It includes both fixed and variable costs. The total cost for a given output is given by the cost function.

   ***Average cost* (*AC*)** is of statistical nature—it is not actual cost. It is obtained by dividing the total cost (*TC*) by the total output (*Q*), i.e.,

$$AC = \frac{TC}{Q}$$

   ***Marginal cost* (*MC*)** is defined as the addition to the total cost on account of producing one additional unit of the product. Or, marginal cost is the cost of the marginal unit produced. Marginal cost is calculated as $TC_n - TC_{n-1}$ where *n* is the number of units produced. Using cost function, *MC* can be defined as:

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

   Total, average and marginal cost concepts are used in the economic analysis of firm's production activities.

3. **Short-run and long-run costs:** Short-run and long-run cost concepts are related to variable and fixed costs, respectively, and often figure in economic analysis interchangeably.

   Short-run costs are those that have a short-run implication in the process of production. Such costs are made once e.g., payment of wages, cost of raw materials, etc. Such costs cannot be used again and again. From analytical point of view, short-run costs are those that vary with the variation in output, the size of the firm remaining the same. Therefore, *short-run* costs are treated as *variable costs*.

NOTES

Long-run costs, on the other hand, are those that have long-run implications in the process of production, i.e., they are used over a long range of output. The costs that are incurred on the fixed factors like plant, building, machinery, etc., are known as long-run costs. It is important to note that the running cost and depreciation of the capital assets are included in the short-run or variable costs.

Furthermore, ***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 variables in the utilization of fixed plant or other facilities whereas long-run costs are associated with the changes in the size and kind of plant.

4. **Incremental costs and sunk costs:** Conceptually, ***incremental costs*** are closely related to the concept of marginal cost but with a relatively wider connotation. While marginal cost refers to the cost of the marginal unit (generally one unit) of output, incremental cost refers to the total additional cost associated with the decisions to expand the output or to add a new variety of product, etc. The concept of incremental cost is based on the fact that in the real world, it is not practicable (for lack of perfect divisibility of inputs) to employ factors for each unit of output separately. Besides, in the long run, when firms expand their production, they hire more of men, materials, machinery and equipment. The expenditures of this nature are incremental costs—not the marginal cost (as defined earlier). Incremental costs arise also owing to the change in product lines, addition or introduction of a new product, replacement of worn out plant and machinery, and replacement of old technique of production with a new one.

   The ***sunk costs*** are those which are made once and for all and cannot be altered, increased or decreased, by varying the rate of output, nor can they be recovered. For example, once it is decided to make incremental investment expenditure and the funds are allocated and spent, all the preceding costs are considered to be the sunk costs. The reason is, such costs are based on the prior commitment and cannot be revised or reversed or recovered when there is a change in market conditions or change in business decisions.

5. **Historical and replacement costs:** *Historical cost* refers to the cost incurred in past on the acquisition of productive assets, e.g. land, building, machinery, etc., whereas *replacement cost* refers to the outlay that has to be made for replacing an old asset. These concepts owe their significance to the unstable nature of price behaviour. Stable prices over time, other things given, keep historical and replacement costs on par with each other. Instability in asset prices makes the two costs differ from each other.

   Historical cost of assets is used for *accounting purposes*, in the assessment of the net worth of the firm. The replacement cost figures in business decisions regarding the renovation of the firm.

6. **Private and social costs:** We have so far discussed the cost concepts that are related to the working of the firm and that are used in the cost-benefit analysis of business decisions. Such costs fall in the category of *private costs.* There are, however, certain other costs that arise due to the functioning of the firm but do not normally figure in the business decisions nor are such costs explicitly borne by the firms. The costs in this category are borne by the society. Thus, the total cost

NOTES

generated by a firm's working may be divided into two categories: (*i*) those paid out or provided for by the firms, and (*ii*) those not paid or borne by the firms including the use of resources freely available plus the disutility created in the process of production. The costs of the former category are known as *private costs* and of the latter category are known as *external* or *social costs*. To mention a few examples of social cost, Mathura Oil Refinery discharging its wastage in the Yamuna River causes water pollution. Mills and factories located in a city cause air pollution, environment pollution and so on. Such costs are termed as *external costs* from the firm's point of view and *social costs* from the society's point of view.

The relevance of the social costs lies in the social cost-benefit analysis of the overall impact of a firm's operation on the society as a whole and in working out the social cost of private gains. A further distinction between private cost and social cost is, therefore, in order.

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 internalized costs that are incorporated in the firm's total cost of production.

Social costs on the other hand, refer to the total cost borne by the society due to production of a commodity. Social costs includes both private cost and the external cost. Social cost includes: (*a*) the cost of resources for which the firm is not required to pay a price, i.e., atmosphere, rivers, lakes, etc., and also for the use of public utility services like roadways, drainage system, etc., and (*b*) the cost in the form of 'disutility' created through air, water, noise and environment pollution, etc. The costs of category (*b*) are generally assumed to equal the total private and public expenditure incurred to safeguard the individual and public interest against the various kinds of health hazards and social tension created by the production system. The private and public expenditure, however, serve only as an indicator of 'public disutility'—they do not give the exact measure of the public disutility or the social costs.

3.4.1 Cost-Output Relations Through Cost Curves

The theory of cost deals with the behaviour of cost in relation to a change in output. In other words, the cost theory deals with cost-output relations. The basic principle of the cost behaviour is that the *total cost increases with increase in output*. This simple statement of an observed fact is of little theoretical and practical importance. What is of importance from a theoretical and managerial point of view is not the absolute increase in the total cost but the direction of change in the average cost (*AC*) and the marginal cost (*MC*). The direction of change in *AC* and *MC*—whether *AC* and *MC* decrease or increase or remain constant—depends on the nature of the cost function. Acost function is a symbolic statement of the technological relationship between the cost and output. The general form of the cost function is written as:

$$TC = f(Q)$$
$$\Delta TC/\Delta Q > 0 \qquad \ldots(3.9)$$

The specific form of the cost function depends on whether the time framework chosen for cost analysis is short-run or long-run. It is important to recall here that some costs remain constant in the short-run while all costs are variable in the long-run. Thus,

depending on whether cost analysis pertains to short-run or to long-run, there are two kinds of cost functions: (*i*) short-run cost functions, and (*ii*) long-run cost functions. Accordingly, the cost output relations are analysed in short-run and long-run framework.

(1) Short-Run Cost-Output Relations

NOTES

In this section, we will analyse the cost-output relations in the short-run. The long-run cost output relations are discussed in the following section.

Before we discuss the cost-output relations, let us first look at the cost concepts and the components used to analyze the short-run cost-output relations.

The basic analytical cost concepts used in the analysis of cost behaviour are Total, Average and Marginal costs. The total cost (*TC*) is defined as the actual cost that must be incurred to produce a given quantity of output. The short-run *TC* is composed of two major elements: (*i*) *total fixed cost* (*TFC*), and (*ii*) *total variable cost* (*TVC*). That is, in the short-run,

$$TC = TFC + TVC \qquad ...(3.10)$$

As mentioned earlier, *TFC* (i.e., the cost of plant, building, etc.) remains fixed in the short run, whereas *TVC* varies with the variation in the output.

For a given quantity of output (*Q*), the average cost, (*AC*), average fixed cost (*AFC*) and average variable cost (*AVC*) can be defined as follows.

$$AC = \frac{TC}{Q} = \frac{TFC + TVC}{Q}$$

$$= \frac{TFC}{Q} + \frac{TVC}{Q} = AFC + AVC$$

Thus, $$AFC = \frac{TFC}{Q} \quad \text{and} \quad AVC = \frac{TVC}{Q}$$

and $$AC = AFC + AVC \qquad ...(3.11)$$

Marginal cost (*MC*) is defined as the change in the total cost divided by the change in the total output, i.e.,

$$MC = \frac{\Delta TC}{\Delta Q} \qquad ...(3.12)$$

In fact, *MC* is the first derivative of cost function, i.e., $\frac{\partial TC}{\partial Q}$

It may be added here that since $\Delta TC = \Delta TFC + \Delta TVC$ and, in the short-run, $\Delta TFC = 0$, therefore, $\Delta TC = \Delta TVC$. Furthermore, under the marginality concept, where $\Delta Q = 1$, $MC = \Delta TVC$. Now we turn to cost function and derivation of various cost curves.

The cost-output relations are determined by the cost function and are exhibited through cost curves. The shape of the cost curves depends on the nature of the cost function. Cost functions are derived from actual cost data of the firms. Given the cost data, cost functions may take a variety of forms, e.g., linear, quadratic or cubic, yielding different kinds of cost curves. The cost curves produced by *linear, quadratic* and *cubic cost functions* are illustrated below.

1. **Linear Cost Function:** A linear cost function takes the following form.

$$TC = a + bQ \qquad \ldots(3.13)$$

(where TC = total cost, Q = quantity produced, $a = TFC$, and b = Change in TVC due to change in Q).

Given the cost function [Eq. (3.13)], AC and MC can be obtained as follows.

$$AC = \frac{TC}{Q} = \frac{a + bQ}{Q} = \frac{a}{Q} + b$$

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.

To illustrate a linear cost function, let us suppose that an actual cost function is given as:

$$TC = 60 + 10Q \qquad \ldots(3.14)$$

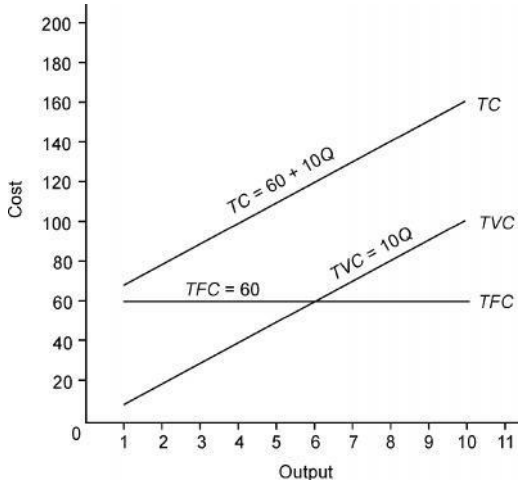


Fig. 3.9 *Linear Cost Functions*

Give the cost function (3.14), one can easily work out *TC*, *TFC*, *TVC*, *MC* and *AC* for different levels of output (*Q*) and can present them in the form of a table as shown in Table 3.2.

Table 3.2 *Tabular Cost Function*

Output Q	*TFC = 60*	*TVC = 10Q*	*TC = 60 + 10Q*	*MC = b = 10*	*AC = 60/Q + 10*
1	60	10	70	–	70.0
2	60	20	80	10	40.0
3	60	30	90	10	30.0
4	60	40	100	10	25.0
5	60	50	110	10	22.0
6	60	60	120	10	20.0
7	60	70	130	10	18.6
8	60	80	140	10	17.5
9	60	90	150	10	16.6
10	60	100	160	10	16.0

NOTES

NOTES

Table 3.2 presents a series of *Q* and corresponding *TFC*, *TVC*, *TC*, *MC* and *AC* for output *Q* from 1 to 10. The figures in Table 3.2, graphed in Fig. 3.9, shows the relationship between total costs and output.

Furthermore, given the cost function (3.14), *AC* can be worked out as follows:

$$AC = \frac{60}{Q} + 10 \qquad \text{and} \qquad MC = 10$$

Fig. 3.9 shows the behaviour of *TC*, *TVC* and *TFC*. The horizontal line shows *TFC* and the line *TVC* = 10*Q* shows the movement in *TVC* with change in *Q*. The total cost function is shown by *TC* = 60 + 10*Q*.

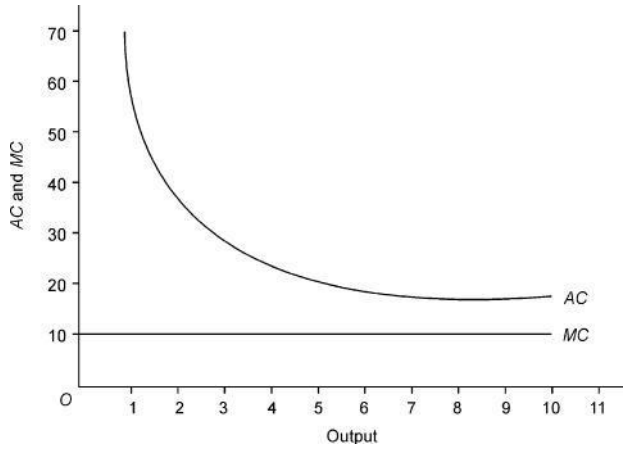


Fig. 3.10 *AC and MC Curves Derived from Linear Cost Function*

More important is the behaviour of *AC* and *MC* curves in Fig. 3.10. Note that, incase of a linear cost function, while *MC* remains constant, *AC* continues to declinewith the increase in output. This is so simply because of the logic of the linear costfunction.

2. **Quadratic Cost Function:** A quadratic cost function is of the form:

$$TC = a + bQ + Q^2 \cdots\cdots\cdots\cdots\cdots\cdots (3.15)$$

where *a* and *b* are constants and *TC* and *Q* are total cost and total output respectively.

Given the cost function (3.15), *AC* and *MC* can be obtained as follows.

$$AC = \frac{TC}{Q} = \frac{a + bQ + Q^2}{Q} = \frac{a}{Q} + b + Q \qquad ...(3.16)$$

$$MC = \frac{\partial TC}{\partial Q} = b + 2Q \qquad ...(3.17)$$

Let the actual (or estimated) cost function be given as:

$$TC = 50 + 5Q + Q^2 \qquad \ldots(3.18)$$

Given the cost function (3.18),

$$AC = \frac{50}{Q} + Q + 5 \quad \text{and} \quad MC = \frac{\partial C}{\partial Q} = 5 + 2Q$$

The cost curves that emerge from the cost function (3.18) are graphed in Fig. 3.11 (*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 trend 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* declines till output 8 and then begins to increase.

NOTES

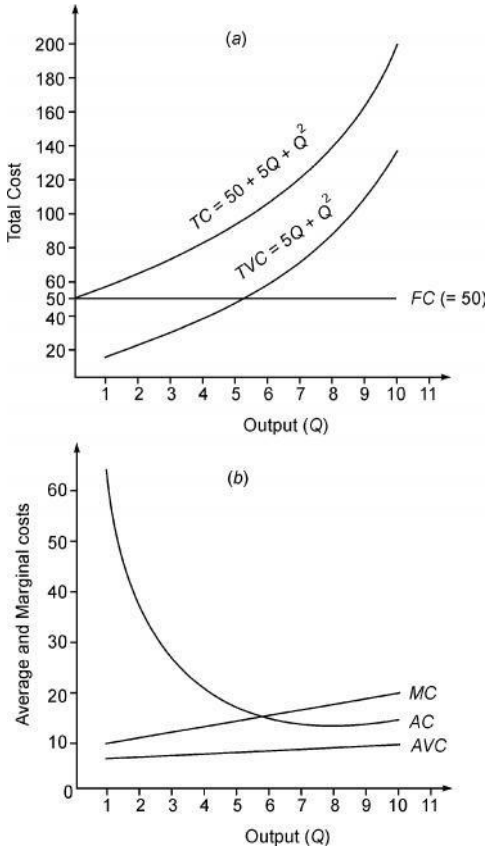


Fig. 3.11 *Cost Curves Derived from a Quadratic Cost Function*

3. **Cubic Cost Function:** A cubic cost function is of the form:

$$TC = a + bQ - cQ^2 + Q^3 \qquad \ldots(3.19)$$

where *a*, *b* and *c* are the parametric constants.

From the cost function (3.19), *AC* and *MC* can be derived as follows.

$$AC = \frac{TC}{Q} = \frac{a + bQ - cQ^2 + Q^3}{Q}$$

$$= \frac{a}{Q} + b - cQ + Q^2$$

and $$MC = \frac{\partial TC}{\partial Q} = b - 2\,cQ + 3Q^2$$

Let us suppose that the cost function is empirically estimated as:

$$TC = 10 + 6Q - 0.9Q^2 + 0.05Q^3 \qquad \ldots(3.20)$$

Note that fixed cost equals 10. *TVC* can be obtained by subtracting 10—the fixed cost—from *TC*-function (3.20).

Thus, $$TVC = 6Q - 0.9Q^2 + 0.05Q^3 \quad \ldots(3.21)$$

NOTES

The *TC* and *TVC*, based on Eqs. (3.20) and (3.21) respectively, have been calculated for $Q = 1$ to 16 and presented in Table 3.3. The *TFC*, *TVC* and *TC* have been graphically presented in Fig. 3.12. As the figure shows, *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 rate and then at an increasing rate with the increase in the output. The rate of increase can be obtained from the slope of *TVC* curve. The pattern of change in the *TVC* stems directly from the law of increasing and diminishing returns to the variable inputs. As output increases, larger quantities of variable inputs are required to produce the same quantity of output due to diminishing returns. This causes a subsequent increase in the variable cost for producing the same output.

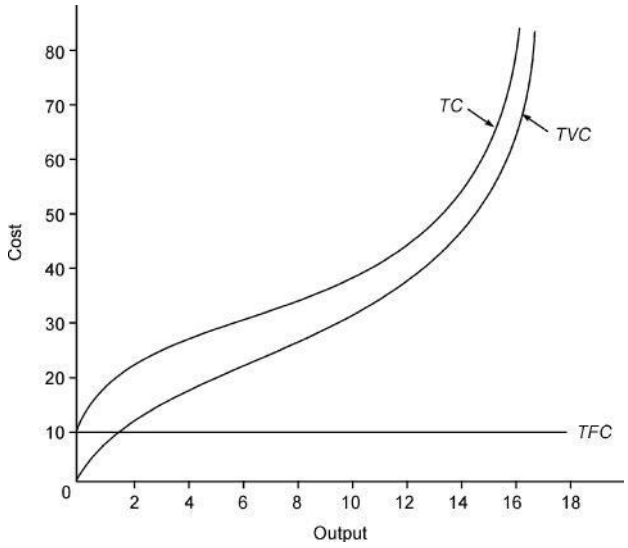


Fig. 3.12 *TC, TFC and TVC Curves*

Table 3.3 *Cost-Output Relations*

Q	FC	TVC	TC	AFC	AVC	AC	MC
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
0	10	0.0	10.00	–	–	–	–
1	10	5.15	15.15	10.00	5.15	15.15	5.15
2	10	8.80	18.80	5.00	4.40	9.40	3.65
3	10	11.25	21.25	3.33	3.75	7.08	2.45
4	10	12.80	22.80	2.50	3.20	5.70	1.55
5	10	13.75	23.75	2.00	2.75	4.75	0.95
6	10	14.40	24.40	1.67	2.40	4.07	0.65
7	10	15.05	25.05	1.43	2.15	3.58	0.65
8	10	16.00	26.00	1.25	2.00	3.25	0.95
9	10	17.55	27.55	1.11	1.95	3.06	1.55
10	10	20.00	30.00	1.00	2.00	3.00	2.45
11	10	23.65	33.65	0.90	2.15	3.05	3.65
12	10	28.80	38.80	0.83	2.40	3.23	5.15
13	10	35.75	45.75	0.77	2.75	3.52	6.95
14	10	44.80	54.80	0.71	3.20	3.91	9.05
15	10	56.25	66.25	0.67	3.75	4.42	11.45
16	10	70.40	80.40	0.62	4.40	5.02	14.15

NOTES

From equations (3.20) and (3.21), we may derive the behavioural equations for *AFC*, *AVC* and *AC*. Let us first consider *AFC*.

Average Fixed Cost (AFC): As already mentioned, the costs that remain fixed for a certain level of output make the total fixed cost in the short-run. The fixed cost is represented by the constant term '*a*' in Eq. (3.19) and *a* = 10 in Eq. (3.20). We know that:

$$AFC = \frac{TFC}{Q} \qquad \ldots(3.22)$$

Substituting 10 for *TFC* in Eq. 3.22, we get:

$$AFC = \frac{10}{Q} \qquad \ldots(3.23)$$

Equation (3.23) expresses the behaviour of *AFC* in relation to change in *Q*. The behaviour of *AFC* for *Q* from 1 to 16 is given in Table 3.3 (Col. 5) and presented graphically by the *AFC* curve in Fig. 3.13. The *AFC* curve is a rectangular hyperbola.

Average Variable Cost (AVC): As defined above, $AVC = \frac{TVC}{Q}$

Given the *TVC* function [Eq. (3.21)], we may express *AVC* as follows.

$$AVC = \frac{6Q - 0.9Q^2 + 0.05Q^3}{Q}$$

$$= 6 - 0.9\,Q + 0.05Q^2 \qquad \ldots(3.24)$$

Having derived the *AVC* function in Eq. (3.24), we may easily obtain the behaviour of *AVC* in response to change in *Q*. The behaviour of *AVC* for output from *Q* = 1 to 16 is given in Table 3.3 (Col. 6), and graphically presented in Fig. 3.13 by the *AVC* curve.

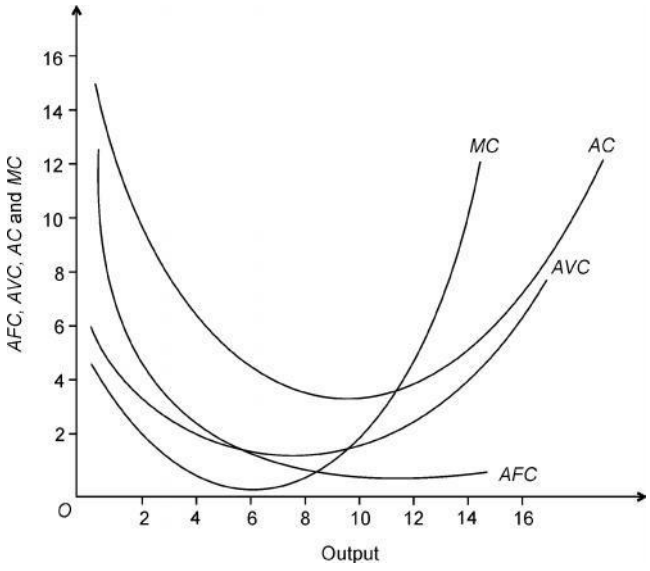


Fig. 3.13 *Short-run Cost Curves*

Critical Value of AVC: *The critical value of Q (in respect of AVC) is one that minimizes AVC*. From Eq. (3.24), we may compute the critical value of *Q* in respect of *AVC*. The *AVC* will be minimum when its (decreasing) rate of change

NOTES

equals zero. This can be accomplished by differentiating Eq. (3.24) and setting it equal to zero. Thus, critical value of Q can be obtained as follows.

$$\text{Critical value of } Q = \frac{\partial AVC}{\partial Q} = -0.9 + 0.10Q = 0$$

$$0.10\,Q = 0.9$$

$$Q = 9$$

In our example, the critical value of $Q = 9$. This can be verified from Table 3.3. The *AVC* is minimum (1.95) at output 9.

Average Cost (AC): The average cost (*AC*) is defined as $AC = \frac{TC}{Q}$.

Substituting Eq. (3.20) for *TC* in the above equation, we get:

$$AC = \frac{10 + 6Q - 0.9Q^2 + 0.05Q^3}{Q}$$

$$= \frac{10}{Q} + 6 - 0.9Q + 0.05Q^2 \qquad ...(3.25)$$

The Eq. (3.25) gives the behaviour of *AC* in response to change in *Q*. The behaviour of *AC* for $Q = 1$ to 16 is given in Table 3.3 (Col. 7) and graphically presented in Fig. 3.13by the *AC* curve. Note that *AC* curve is *U*-shaped.

Minimization of AC: One objective of business firms is to minimize *AC* of their product. The level of output that minimizes *AC* can be obtained by differentiating Eq. (3.25) and setting it equal to zero. Cost-minimizing *Q* can be obtained as follows.

$$\frac{\partial AC}{\partial Q} = \frac{10}{Q^2} - 0.9 + 0.1Q = 0$$

When we simplify this equation by multiplying it by Q^2, it takes the form of a quadratic equation as:

$$10 - 0.9Q^2 + 0.1Q^3 = 0$$

When this equation is multiplied by 10, for simplification, it takes the form,

$$Q^3 - 9Q^2 - 100 = 0 \qquad ...(3.26)$$

By solving equation (3.26), we get $Q = 10$.

Thus, the critical value of output in respect of *AC* is 10. That is, *AC* reaches its minimum at $Q = 10$. This can be verified from Table 3.3.

Marginal Cost (MC): The concept of marginal cost (*MC*) is particularly useful in economic analysis. *MC* is technically the first derivative of the *TC* function. Given the *TC* function in Eq. (3.20), the *MC* function can be obtained as:

$$MC = \frac{\partial TC}{\partial Q} = 6 - 1.8Q + 0.15Q^2 \qquad ...(3.27)$$

Equation (3.27) represents the behaviour of *MC*. The behaviour of *MC* for $Q = 1$ to 16 computed as $MC = TC_n - TC_{n-1}$ is given in Table 3.3 (Col. 8) and graphically presented by the *MC* curve in Fig. 3.13. The critical value of *Q* with respect to *MC* is 6 or 7. This can be seen from Table 3.3.

Cost Curves and the Law of Diminishing Returns

NOTES

We now return to the law of diminishing returns and explain it through the cost curves. Figs. 3.12 and 3.13 present the short-term law of production i.e., the law of diminishing returns. Let us recall the law: it states that when more and more units of a variable input are applied, other inputs held constant, the returns from the marginal units of the variable input may initially increase but they decrease eventually. The same law can also be interpreted in terms of decreasing and increasing costs. The law can then be stated as, if more and more units of a variable input are applied to a given amount of a fixed input, the marginal cost initially decreases, but eventually increases. Both interpretations of the law yield the same information—one in terms of marginal productivity of the variable input, and the other in terms of the marginal cost. The former is expressed through a production function and the latter through a cost function.

Fig. 3.13 presents the short-run laws of return to a variable input 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 *AC = AFC +AVC, AC* is also declining. This shows the operation of the law of increasing returns in the initial stage of production. But beyond a certain level of output (i.e., 9 units in our example), while *AFC* continues to fall, *AVC* starts increasing because of a faster increase in the *TVC*. Consequently, the rate of fall in *AC* decreases. The *AC* reaches its minimum when output increases to 10 units. Beyond this level of output, *AC* starts increasing which shows that the law of diminishing returns comes into operation. The *MC* curve represents the change in both the *TVC* and *TC* curves due to change in output. A downward trend in the *MC* shows increasing marginal productivity of the variable input due mainly to internal economies resulting from increase in production. Similarly, an upward trend in the *MC* shows increase in *TVC*, on the one hand, and decreasing marginal productivity of the variable input, on the other.

Some Important Cost Relationships

Some important relationships between costs used in analysing the short-run cost-behaviour may now be summed up as follows:

(a) Over the range of output *AFC* and *AVC* fall, *AC* also falls.

(b) When *AFC* falls but *AVC* increases, change in *AC* depends on the rate of change in *AFC* and *AVC*.

  (i) If decrease in *AFC* > increase in *AVC*, then *AC* falls

  (ii) If decrease in *AFC* = increase in *AVC*, *AC* remains constant

  (iii) If decrease in *AFC* < increase in *AVC*, then *AC* increases

(c) *AC* and *MC* are related in following ways.

  (i) When *MC* falls, *AC* follows, over a certain range of output. When *MC* is falling, the rate of fall in *MC* is greater than that of *AC*, because while *MC* is attributed to a single marginal unit, *AC* is distributed over the entire output. Therefore, *AC* decreases at a lower rate than *MC*.

  (ii) Similarly, when *MC increases, AC also increases but at a lower rate* for the reason given in (*i*). There is, however, a range of output over which the relationship does not exist. Compare the behaviour of *MC* and *AC* over the range of output from 6 units to 10 units (see Fig. 3.13). Over this range of output, *MC* begins to increase while *AC* continues to decrease. The reason for this can be seen in Table 3.2: when *MC* starts increasing, it increases at

NOTES

a relatively lower rate that is sufficient only to reduce the rate of decrease in *AC*—not sufficient to push the *AC* up.

(iii) *MC curve intersects AC curve at its minimum*. The reason is, while *AC* continues to decrease, *MC* begins to rise. Therefore, they are bound to intersect. Also, when *AC* is at its minimum, it is neither increasing nor decreasing: it is constant. When *AC* is constant, *AC* = *MC*. That is the point of intersection.

Output Optimization in the Short-Run

Let us suppose that a short-run cost function is given as

$$TC = 200 + 5Q + 2Q^2 \qquad \ldots(3.28)$$

As noted earlier, the level of output is optimized at the level of production at which *MC* = *AC*. In other words, at optimum level of output, *AC* = *MC*. Given the cost function in Eq. (3.28),

$$AC = \frac{200 + 5Q + 2Q^2}{Q}$$

$$= \frac{200}{Q} + 5 + 2Q \qquad \ldots(3.29)$$

and

$$MC = \frac{\partial TC}{\partial Q} = 5 + 4Q \qquad \ldots(3.30)$$

By equating *AC* and *MC* equations, i.e., Eqs. (3.29) and (3.30), respectively, and solving them for *Q*, we get the optimum level of output. Thus,

$$\frac{200}{Q} + 5 + 2Q = 5 + 4Q = 2Q$$

$$2Q^2 = 200$$

$$Q^2 = 100$$

$$Q = 10$$

Thus, given the cost function (3.28), the optimum output is 10.

(2) Long-Run Cost-Output Relations

In the context of production theory, long-run is defined as a period in which all the inputs become variable. The variability of inputs is based on the assumption that in the long-run, supply of all the inputs, including those (especially capital) held constant in the short-run, becomes elastic. The firms are, therefore, in a position to expand the scale of their production by hiring a larger quantity of all the inputs. The long-run cost output relations, therefore, imply the relationship between the changing scale of the firm and the total output, whereas in the short-run, this relationship is essentially one between the total output and the variable cost (labour). Specifically, long-run cost-output relations refers to the behaviour of *TC*, *AC* and *MC* in response to simultaneous and proportionate charge in both labour and capital costs.

To understand the long-run-cost-output relations and to derive long-run cost curves, it will be helpful to imagine that a long-run is composed of a series of short-run production decisions. As a corollary of this, long-run cost curve is composed of a series of short-run

cost curves. With this perception of long-run-cost-out relationship, we may now show the derivation of the long-run cost curves and study their relationship with output.

NOTES

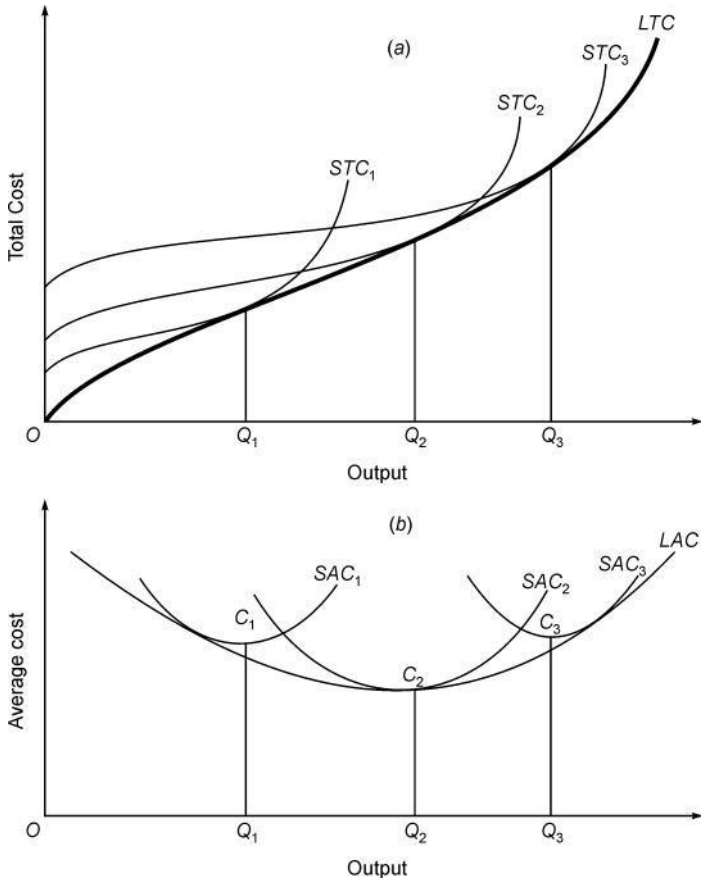


Fig. 3.14 *Long-run Total and Average Cost Curves*

Long-run Total Cost Curve (LTC)

In order to draw the long-run total cost curve, let us begin with a short-run situation. Suppose that a firm having only one plant has its **short-run total** cost curve as given by STC_1, in panel (*a*) of Fig. 3.14. Let us now suppose that the firm decides to add two more plants over time, one after the other. As a result, two more short-run total cost curves are added to STC_1, in the manner shown by STC_2 and STC_3 in Fig. 3.14 (*a*). The *LTC* can now be drawn through the minimum points of STC_1, STC_2 and STC_3 as shown by the *LTC* curve corresponding to each *STC*.

Long-run Average Cost Curve (LAC)

Like *LTC*, long-run average cost curve (*LAC*) is derived by combining the short-run average cost curves (SAC_s). Note that there is one *SAC* associated with each *STC*. The curve SAC_1 in panel (*b*) of Fig. 3.14 corresponds to STC_1 in panel (*a*). Similarly, SAC_2 and SAC_3 in panel (*b*) correspond to STC_2 and STC_3 in panel (*a*), respectively. Thus, given the STC_1, STC_2, STC_3 curves in panel (*a*) of Fig. 3.14, there are three corresponding *SAC* curves as given by SAC_1, SAC_2, and SAC_3 curves in panel (*b*) of Fig. 3.14. Thus, the firm has a series of *SAC* curves, each having a bottom point showing the minimum *SAC*. For instance, C_1Q_1 is minimum *AC* when the firm has only one plant. The *AC* decreases to C_2Q_2 when the second plant is added and then rises to C_3Q_3 after the addition of the third plant. The *LAC* curve can be drawn through the SAC_1, SAC_2 and

NOTES

SAC_3 as shown in Fig. 3.14 (*b*). The *LAC* curve is also known as the **'Envelope Curve'** or **'Planning Curve'** as it serves as a guide to the entrepreneur in his plans to expand production.

The *SAC* curves can be derived from the data given in the *STC* schedule, from *STC* function or straightaway from the *LTC* curve. Similarly, *LAC* curve can be derived from *LTC*-schedule, *LTC* function or from *LTC*-curve.

The relationship between *LTC* and output, and between *LAC* and output can now be easily derived. It is obvious from the *LTC* that the long-run cost-output relationship is similar to the short-run cost-output relation. With the subsequent increases in the output, *LTC* first increases at a decreasing rate, and then at an increasing rate. As a result, *LAC* initially decreases until the optimum utilization of the second plant and then it begins to increase. These cost-output relations follow the 'laws of returns to scale'. When the scale of the firm expands, unit cost of production initially decreases, but ultimately increases as shown in Fig. 3.14 (*b*). The decrease in unit cost is attributed to the internal and external economies and the eventual increase in cost, to the internal and external diseconomies. The economies and diseconomies of scale are discussed in the following section.

Long-run Marginal Cost Curve (LMC)

The *long-run marginal cost curve* (*LMC*) is derived from the short-run marginal cost curves (*SMCs*). The derivation of *LMC* is illustrated in Fig. 3.15 in which *SACs*, SMC_s and *LAC* are the same as in Fig. 3.14 (*b*). To derive the *LMC*, consider the points of tangency between *SACs* and the *LAC*, i.e., points *A*, *B* and *C*. In the long-run production planning, these points determine the output at the different levels of production. Each of these outputs has an *SMC*. For example, if we draw a perpendicular from point *A*, it intersects SMC_1 at point *M* determining *SMC* at MQ_1. The same process can be repeated for points *B* and *C* to find out *SMC* at outputs Q_2 and Q_3. Note that points *B* and *C* determine *SMC* at BQ_2 and CQ_3 respectively. A curve drawn through points *M*, *B* and *N*, as shown by the *LMC*, represents the behaviour of the marginal cost in the long-run. This curve is known as the long-run marginal cost curve, *LMC*. It shows the trends in the marginal cost in response to the changes in the scale of production.

Some important inferences may be drawn from Fig. 3.14. 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*. Another important point to notice is that *LMC* intersects *LAC* when the latter is at its minimum, i.e., 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:

$$SAC_2 = SMC_2 = LAC = LMC$$

Optimum Plant Size and Long-Run Cost Curves

The short-run cost curves are helpful in showing how a firm can decide on the *optimum utilization of the plant*—the fixed factor, or how it can determine the output level that minimizes cost. Long-run cost curves, on the other hand, can be used to show how a firm can decide on the *optimum size of the firm*.

NOTES

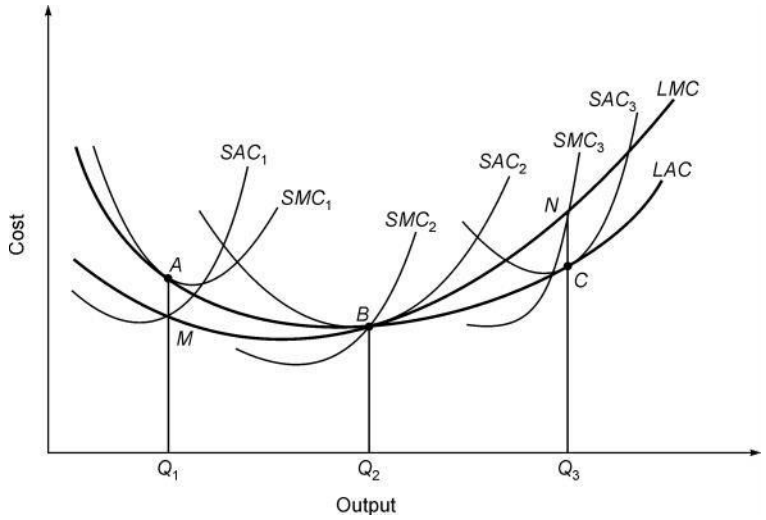


Fig. 3.15 *Derivation of LMC*

Conceptually, the optimum size of a firm is one which ensures the most efficient utilization of resources. Practically, the optimum size of the firm is one that minimizes the *LAC*. Given the state of technology over time, there is technically a unique size of the firm and level of output associated with the least-cost concept. In fig. 3.15, the optimum size of the firm consists of two plants represented by SAC_1 and SAC_2. The two plants 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 less than optimal size. Similarly, expansion of the firm beyond production capacity OQ_2, causes a rise in *SMC* and, therefore, in *LAC*. It follows that given the technology, a firm aiming to minimize its average cost over time must choose a plant that gives minimum *LAC* where *SAC* = *SMC* = *LAC* = *LMC*. This size of plant assures the most efficient utilization of the resources. Any change in output level—increase or decrease—will make the firm enter the area of inoptimality.

3.4.2 Cost Minimization

A typical problem frequently encountered in managerial decision-making is minimization of cost, given the constraints. We now illustrate the application of linear programming technique to a simple cost minimization problem through a simple hypothetical example.

Let us suppose that an automobile manufacturing firm produces only trucks and cars for which it uses only three inputs—labour, machine and steel. The firm gets the contractual supplies of inputs and, for compliance with agreement, it is required to make use of a minimum quantity of inputs, i.e., 160 man-hours, 36 machine hours and 48 tonnes of steel. The overall production conditions are given in Table 3.4.

Table 3.4 *Production Conditions*

Inputs	*Necessary minimum supply of inputs*	*Input requirement per*	
		Car	*Truck*
Labour	160 (man-hours)	10	40
Machine	36 (machine-hours)	6	3
Steel	48 (tones)	4	8

NOTES

Given the price of inputs, the production cost per truck has been worked out at ` 60,000 and the cost per car at ` 20,000. The task before the decision-makers is to work out the combination of trucks and cars, that can be produced at the minimum total cost. The problem may be reformulated in the linear programming mode as follows.

$$\text{Minimize Cost} = 20{,}000C + 60{,}000T \quad \ldots(3.31)$$

Subject to

$$10C + 40T \leq 160 \text{ (Labour)} \quad \ldots(3.32)$$

$$6C + 3T \leq 36 \text{ (Machine)} \quad \ldots(3.33)$$

$$4C + 8T \leq 48 \text{ (Steel)} \quad \ldots(3.34)$$

where $T \geq 0$ and $C \geq 0$

Graphical Solution

The graphical solution to this cost minimizing problem is illustrated in Fig. 3.16. The first step is to graph the input constraints following the procedure elaborated in the profit maximization problem. Consider, for example, the labour constraint equation, given as:

$$10C + 40T = 160$$

If T is set to zero, $C = 16$, i.e., if the entire 160 man-hours are used to produce cars, 16 cars can be produced. This is indicated by point M in Fig. 3.16. Similarly, if C is set to zero, $T = 4$, i.e., 4 trucks can be produced as indicated by point N where $C = 0$. By joining points M and N, we get the labour-constraint line. This line indicates the possible combinations of cars and trucks that can be produced by using 160 man-hours, other things given. The same procedure is used to graph the other constraint equations as shown by the lines, JK and TR. These lines are called **isocosts**. The area to the right of the isocosts, i.e., the shaded area, is the feasibility plane for larger amounts of inputs.

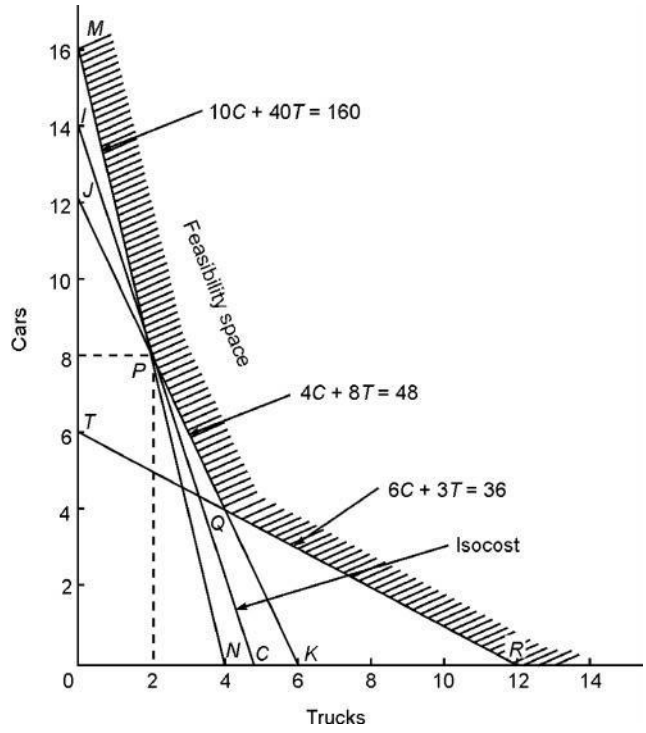


Fig. 3.16 *Cost-Minimisation: Graphical Solution*

The optimum solution to the problem lies either on the boundary line or in the shaded area. Thus, there are more than one solution. But only one of them is optimum. The optimum solution, i.e., the optimum combination of cars and trucks, can be obtained by drawing an isocost line *IC* having a slope of 3 (which is truck/car cost ratio = 60/30). The isocost (*IC*) which is tangent to the feasibility boundary at point *P*, offers the optimum solution, i.e., the combination of cars and trucks that minimize the total cost of production. The optimum solution is 8 cars + 2 trucks.

NOTES

Table 3.5 *Output-Mix and Total Cost*

Corner Points	*Combination*			*Total*			*cost*
	Cars	+	*Trucks*	*20,000(C)*	+ *60,000(T)*	=	*C*
M	16	+	10	20,000 (16)	+ 60,000 (0)	=	320,000
P	8	+	2	20,000 (8)	+ 60,000 (2)	=	280,000
Q	4	+	4	20,000 (4)	+ 60,000 (4)	=	320,000
R	0	+	12	20,000 (0)	+ 60,000 (12)	=	720,000

The optimality of the solution can be checked by comparing the total cost for the combination of cars and trucks at each corner point *M, P, Q* and *R* because only one of these points offers the optimum solution. As shown in Table 3.5, the optimum solution to the problem of cost-minimization lies at point *P*. Incidentally, the solution cannot be checked algebraically because three simultaneous equations in our example involve two unknowns, *C* and *T*.

4.5 PROFIT MAXIMIZING OUTPUT

Profit maximization has been the most important assumption on which economists have built price and production theories. This assumption has, however, been strongly questioned and alternative hypotheses suggested. This issue will be discussed in the forthcoming sections. Let us first look into the importance of the profit maximization assumption and theoretical conditions of profit maximization.

The conventional economic theory assumes profit maximization as the only objective of business firms. Profit maximization as the objective of business firms has a long history in economic literature. It forms the basis of conventional price theory. Profit maximization is regarded as the most reasonable and analytically the most 'productive' business objective. The strength of this assumption lies in the fact that this assumption 'has never been unambiguously disproved'.

Besides, profit maximization assumption has a greater predictive power. It helps in predicting the behaviour of business firms in the real world and also the behaviour of price and output under different market conditions. No alternative hypothesis explains and predicts the behaviour of firms better than the profit maximization assumption. Let us now discuss the theoretical conditions for profit maximization.

Check Your Progress

7. What are variable costs and what do they include?
8. What do the cost curves exhibit?
9. What do the fixed costs include?

4.5.1 Total, Marginal and Average Revenue

Marginal revenue is the revenue obtained from the production and sale of one additional unit of output and marginal cost is the cost arising due to the production of one additional unit of output. The marginal cost of production and marginal revenue

NOTES

are economic measures used to determine the amount of output and the price per unit of a product to maximize profits. A rational company always seeks to maximize its profit, and the relationship between marginal revenue and the marginal cost of production helps to find the point at which this occurs. The point at which marginal revenue equals marginal cost maximizes a company's profit.

Total revenue in economics refers to the**total** receipts from sales of a given quantity of goods or services. It is the **total** income of a business and is calculated by multiplying the quantity of goods sold by the price of the goods. It can be calculated as the selling price of the firm's product times the quantity sold, i.e. total revenue = price (average revenue) × quantity, or letting TR be the total revenue function:

$$TR(Q) = P(Q) \times Q$$

Average revenue is the revenue generated per unit of output sold. It plays a role in the determination of a firm's profit. Per unit profit is average revenue minus average (total) cost. A firm generally seeks to produce the quantity of output that maximizes profit. Hence, average revenue can be obtained by the following formula:

$$\text{Average revenue} = \frac{\text{Total Revenue}}{\text{Quantity}}$$

These concepts will be discussed further.

4.5.2 Profit Maximization Conditions

Total profit (Π) is defined as

$$\pi = TR - TC \qquad \ldots(3.35)$$

where TR = total revenue, and TC = total cost.

There are two conditions that must be fulfilled for $TR - TC$ to be maximum. These conditions are called: (*i*) necessary or the first order condition, and (*ii*) secondary or supplementary condition.

The necessary or the first-order condition requires that marginal revenue (*MR*) must be equal to marginal cost (*MC*). By definition, marginal revenue is the revenue obtained from the production and sale of one additional unit of output and marginal cost is the cost arising due to the production of one additional unit of output.

The secondary or the second-order condition requires that the necessary or first-order condition must be satisfied under the stipulation of decreasing *MR* and rising *MC*. The fulfilment of the two conditions makes it the *sufficient condition*.

The profit maximizing conditions can also be presented algebraically as follows.

We know that a profit maximizing firm seeks to maximize:

$$\pi = TR - TC$$

Let us suppose that the total revenue (*TR*) and total cost (*TC*) functions are, respectively, given as:

$$TR = f(Q) \text{ and } TC = f(Q)$$

where Q = quantity produced and sold.

By substituting total revenue and total cost functions in Eq. (3.35), the profit function may be written as

$$\pi = f(Q)_{TR} - f(Q)_{TC} \qquad \ldots(3.36)$$

Equation (3.36) can now be manipulated to illustrate the *first* and *second* order conditions of profit maximization as follows.

1. First-order condition: The first-order condition of maximizing a function is that its first derivative must be equal to zero. Thus, the first-order condition of profit maximization is that the first derivative of the profit function Eq. (3.36) must be equal to zero. Differentiating the total profit function and setting it equal to zero, we get:

$$\frac{\partial \Pi}{\partial Q} = \frac{\partial TR}{\partial Q} - \frac{\partial TC}{\partial Q} = 0 \qquad \ldots (3.37)$$

This condition holds only when:

$$\frac{\partial TR}{\partial Q} = \frac{\partial TC}{\partial Q}$$

In Eq. (3.37), the term $\partial TR/\partial Q$ gives the slope of the *TR* curve which in turn gives the marginal revenue (*MR*). Similarly, the term $\partial TC/\partial Q$ gives the slope of the total cost curve which is the same as marginal cost (*MC*). Thus, the first-order condition for profit maximization can be stated as:

$$MR = MC$$

The first-order condition is generally known as *necessary condition*. A necessary condition is one that *must* be satisfied for an event to take place. In other words, the condition that *MR = MC* must be satisfied for profit to be maximum.

2. Second-order Condition: As already mentioned, in non-technical terms, the second-order condition of profit maximization requires that the first order condition is satisfied under rising *MC* and decreasing *MR*. This condition is illustrated in Fig. 3.17. The *MC* and *MR* curves are the usual marginal cost and marginal revenue curves respectively. Incidentally, *MC* and *MR* curves are derived from *TC* and *TR* functions respectively. *MC* and *MR* curves intersect at two points, P_1 and P_2. Thus, the first order condition is satisfied at both the points, but the second order condition of profit maximization is satisfied only at point P_2. Technically, the second-order condition requires that the second derivative of the profit function is negative. The second derivative of the total profit function is given as:

$$\frac{\partial^2 \Pi}{\partial Q^2} = \frac{\partial^2 TR}{\partial Q^2} - \frac{\partial^2 TC}{\partial Q^2} \qquad \ldots (3.38)$$

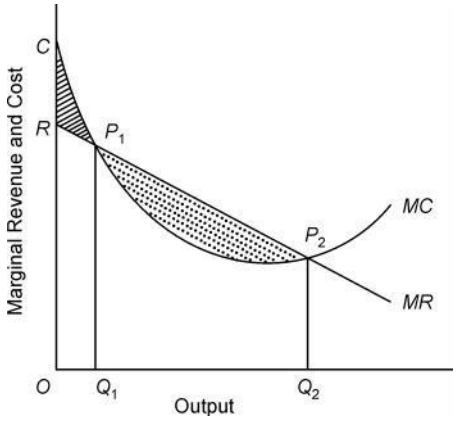


Fig. 3.17 *Marginal Conditions of Profit Maximization*

NOTES

The second-order condition requires that:

$$\frac{\partial^2 TR}{\partial Q^2} - \frac{\partial^2 TC}{\partial Q^2} < 0$$

$$\text{or } \frac{\partial^2 TR}{\partial Q^2} < \frac{\partial^2 TC}{\partial Q^2} \quad \text{...(3.39)}$$

Since $\partial^2 TR/\partial Q^2$ gives the slope of *MR* and $\partial^2 TC/\partial Q^2$ gives the slope of *MC*, the second-order condition may also be written as:

Slope of *MR* < Slope of *MC*

It implies that *MC* must have a steeper slope than *MR* or *MC* must intersect the *MR* from below.

To conclude, profit is maximized where both the first and second order conditions are satisfied.

We may now apply the profit maximization conditions to a hypothetical example and compute profit maximizing output.

We know that $TR = P.Q$

Suppose demand function for a product is given as $Q = 50 - 0.5P$. Given the demand function, price (P) function can be derived as:

$$P = 100 - 2Q \quad \ldots(3.40)$$

By substituting price function for *P* in *TR* equation, we get:

$$TR = (100 - 2Q)Q$$

$$\text{or } TR = 100Q - 2Q^2 \quad \ldots(3.41)$$

Let us also suppose that the total cost function is given as:

$$TC = 10 + 0.5\,Q^2 \quad \ldots(3.42)$$

Given the *TR* function (3.41) and *TC* function (3.42), we can now apply the first order condition of profit maximization and find profit maximizing output. We have noted that profit is maximum where:

$$MR = MC$$

$$\text{or } \frac{\partial TR}{\partial Q} = \frac{\partial TC}{\partial Q}$$

Given the total *TR* function in Eq. (3.41) and *TC* function in Eq. (3.42),

$$MR = \frac{\partial TR}{\partial Q} = 100 - 4Q \quad \text{...(3.43)}$$

$$\text{and } MC = \frac{\partial TR}{\partial Q} = Q \quad \text{...(3.44)}$$

Thus, profit is maximum where:

$$MR = MC$$

$$\text{or } 100 - 4Q = Q$$

$$5Q = 100$$

$$Q = 20$$

The output 20 satisfies the **second-order condition** also. The second-order condition requires that

$$\frac{\partial^2 TR}{\partial Q^2} - \frac{\partial^2 TC}{\partial Q^2} < 0$$

In other words, the second-order condition requires that:

$$\frac{\partial MR}{\partial Q} - \frac{\partial MC}{\partial Q} < 0$$

or $\frac{\partial (100 - 4Q)}{\partial Q} - \frac{\partial Q}{\partial Q} < 0$

That is, $-4 - 1 < 0$

Thus, the second-order condition is also satisfied at output 20.

4.5.3 Controversy Over Profit Maximization Objective: Theory vs. Practice

Arguments against Profit Maximization Objective

As noted above, traditional theory assumes profit maximization as the sole objective of a business firm. In practice, however, firms have been found to be pursuing many objectives other than profit maximization. It is argued, in the first place, that the reason for the firms, especially the large corporations, pursuing goals other than profit maximization is the dichotomy between the ownership and the management. The separation of management from ownership gives managers an opportunity and also discretion to set goals other than profit maximization. It is argued that large firms pursue such goals as sales maximization, maximization of managerial utility function, maximization of firm's growth rate, making a target profit, retaining market share, building up the net worth of the firm, and so on.

Secondly, traditional theory assumes full and perfect knowledge about current market conditions and the future developments in the business environment of the firm. The firm is thus supposed to be fully aware of its demand and cost conditions in both short and long runs. Briefly speaking, a complete certainty about the market conditions is assumed. Some modern economists question the validity of this assumption. They argue that the firms do not possess the perfect knowledge of their costs, revenue and future business environment. They operate in the world of uncertainty. Most price and output decisions are based on *probabilities*.

Finally, the equi-marginal principle of profit maximization, i.e., equalizing *MC* and *MR*, has been claimed to be absent in the decision-making process of the firms. Empirical studies of the pricing behaviour of the firms have shown that the marginal rule of pricing does not stand the test of empirical verification. Hall and Hitch have found, in their study of pricing practices of 38 UK firms, that the firms do not pursue the objective of profit maximization and that they do not use the marginal principle of equalizing *MR* and *MC* in their price and output decisions. Most firms aim at long-run profit maximization. In the short-run, they set the price of their product on the basis of *average cost principle*, so as to cover AC = AVC + AFC (*AC* = Average cost, *AVC* = Average variable cost, *AFC* = Average fixed cost) and a normal margin of profit (usually 10 per cent). In a similar

NOTES

NOTES

study, Gordon has found: (*i*) that there is a marked deviation in the real business conditions from the assumptions of the traditional theory and (*ii*) that pricing practices were notably different from the marginal theory of pricing. Gordon has concluded that the real business world is much more complex than the one postulated by the theorists. Because of the extreme complexity of the real business world and ever-changing conditions, the past experience of the business firms is of little use in forecasting demand, price and costs. The firms are not aware of their *MR* and *MC*. The *average-cost-principle* of pricing is widely used by the firms. Findings of many other studies of the pricing practices lend support to the view that there is little link between pricing theory and pricing practices.

The Defence of Profit Maximization

The arguments against profit-maximization assumption, however, should not mean that pricing theory has no relevance to the actual pricing policy of the business firms. A section of economists has strongly defended the profit maximization objective and 'marginal principle' of pricing and output decisions. The empirical and theoretical support put forward by them in defence of the profit maximization objective and marginal rule of pricing may be summed as follows.

In two empirical studies of 110 'excellently managed companies', J. S. Earley has concluded that the firms do apply the marginal rules in their pricing and output decisions. Fritz Maclup has argued in abstract theoretical terms that empirical studies by Hall and Hitch and by Lester do not provide conclusive evidence against the marginal rule and these studies have their own weaknesses. He argues further that there has been a misunderstanding regarding the purpose of traditional theory of value. The traditional theory seeks to explain market mechanism, resource allocation through price mechanism and has a predictive value, rather than deal with specific pricing practices of certain firms. The relevance of marginal rules in actual pricing system of firms could not be established for lack of communication between the businessmen and the researchers as they use different terminology like *MR*, *MC* and elasticities. Besides, businessmen, even if they do understand economic concepts, would not admit that they are making *abnormal profits* on the basis of marginal rules of pricing. They would instead talk of a 'fair profit'. Also, Maclup is of the opinion that the practices of setting *price* equal to *average variable cost* plus a *profit margin* is not incompatible with the marginal rule of pricing and that the assumptions of traditional theory are plausible.

While the controversy on profit maximization objective remains unresolved, the conventional theorists, the marginalists, continue to defend the profit maximization objective and its marginal rules.

Other Arguments in Defence of Profit Maximization Hypothesis: The conventional economic theorists defend the profit maximization hypothesis on the following grounds also.

- **Profit is indispensable for firm's survival:** The survival of all the profit-oriented firms in the long run depends on their ability to make a reasonable profit depending on the business conditions and the level of competition. Whatprofit is reasonable may be a matter of opinion. But, making profit is a necessary condition for the survival of the firm. Once the firms are able to make profit, they try to make it as large as possible, i.e., they tend to maximize it.
- **Achieving other objectives depends on firm's ability to make profit:** Many other objectives of business firms have been cited in economic literature,

NOTES

e.g., maximization of managerial utility function, maximization of long-run growth, maximization of sales revenue, satisfying all the concerned parties, increasing and retaining market share, etc. The achievement of such alternative objectives depends wholly or partly on the primary objective of making profit.

- **Evidence against profit maximization objective is not conclusive:** Profit maximization is a time-honoured objective of business firms. Although this objective has been questioned by many researchers, some economists have argued that the evidence against it is not conclusive or unambiguous.
- **Profit maximization objective has a greater predicting power:** Compared to other business objectives, profit maximization assumption has been found to provide a much more powerful basis for predicting certain aspects of firms' behaviour. As Friedman has argued, the validity of the profit maximization objective cannot be judged by a priori logic or by asking business executives, as some economists have done. The ultimate test of its validity lies in its ability to predict the business behaviour and the business trends.
- **Profit is a more reliable measure of a firm's efficiency:** Though not perfect, profit is the most efficient and reliable measure of the efficiency of a firm. It is also the source of internal finance. Profit as a source of internal finance assumes a much greater significance when financial market is highly volatile. The recent trend shows a growing dependence on the internal finance in the industrially advanced countries. In fact, in developed countries, internal sources of finance contribute more than three-fourths of the total finance.
- **Finally,** according to Milton Friedman, whatever one may say about firms' motivations, if one judges their motivations by their acts, profit maximization appears to be **a more valid business objective.**

4.6 SUPPLY CURVE AND SHIFT IN SUPPLY CURVE

A supply curve is a graphical presentation of the supply schedule. The supply curve *SS*' given in Fig. 3.18 has been obtained by plotting the data in Table 3.6. The points *S, P, Q, R, T* and *S*' show the price-quantity combinations on the supply curve *SS*'. The supply curve, *SS*', depicts *the law of supply*. The upward slope of the supply curve indicates the rise in the supply of shirts with the rise in its price and fall in the supply with fall in prices. For example, at price ` 200, only 35 thousand shirts are supplied per week. When price rises to ` 400, supply increases to 60 thousand shirts.

As shown in Fig. 3.18, *a supply curve has a positive slope*. The positive slope or upward movement of the supply curve is caused by the rise in cost of production and seller's effort to make a larger profit. The rise in cost of production results from the law of diminishing returns. In fact, supply curve is derived from the marginal cost curve.

Table 3.6 *Supply Schedule of Shirts*

Price (in `)	Supply (Shirts in '000)
100	10
200	35
300	50
400	60
600	75
800	80

Check Your Progress

10. Define marginal revenue.
11. What is average revenue?
12. What is a necessary condition?

NOTES

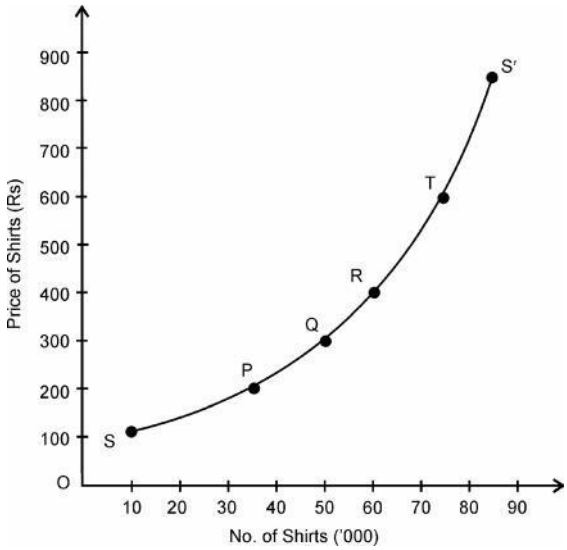


Fig. 3.18 *Supply Curve of Shirt*

4.6.1 Shift in the Supply Curve

We have shown above that a change in the price of a commodity causes a change in its quantity supplied along a given supply curve. Although price of a commodity is the most important determinant of its supply, it is not the only determinant. Many other factors influence the supply of a commodity. Given the supply curve of a commodity, when there is change in its other determinants, the supply curve shifts rightward or leftward depending on the effect of such changes. Let us now explain how other determinants of supply cause shift in the supply curve.

(i) **Change in input prices:** When input prices decrease, the use of inputs increase. As a result, product supply increases and the supply curve *SS* shifts to the right to *SS*", as shown in Fig. 3.19. Similarly, when input prices increase, product supply curve shifts leftward from *SS* to *SS'*.

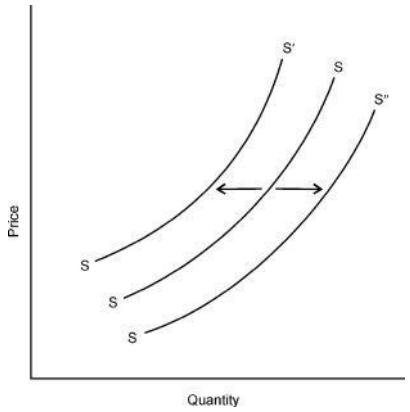


Fig 3.19 *Shift in the Supply Curve*

(ii) **Technological progress:** Technological changes that reduce cost of production or increase efficiency in production cause increase in product supply. For instance, introduction of high yielding variety of paddy and new techniques of cultivation increased per acre yield of rice in India in the 1970s. Such changes make the supply curve shift to the right.

(iii) **Price of product substitutes:** Given its technology and production capacity, a firm can produce more than one good which require a similar technology. For example, a refrigerator company can also produce ACs; Tatas, famous for truck production can also produce cars; Maruti Udyog can produce trucks, and so on. Fall in the price of one of the product substitutes may lead to the rise in the supply of other due to capacity utilization for profit maximization. This may cause shift in the supply curve.

(iv) **Nature and size of the industry:** The supply of a commodity depends also on whether an industry is monopolized or competitive. Under monopoly, supply is fixed. When a monopolized industry is made competitive, the total supply increases. Besides, if size of an industry increases due to new firms joining the industry, the total supply increases and industry supply curve shifts rightward.

(v) **Government policy:** When government imposes restrictions on production, e.g., import quota on inputs, rationing of or quota imposed on input supply, etc., production tends to fall. Such restrictions make supply curve shift leftward.

(vi) **Non-economic factors:** Factors like labour strikes and lock-outs, war, drought, flood, communal riots, epidemics, etc. also adversely affect the supply of commodities and make the supply curve shift leftward.

4.6.2 Supply Function

The *supply function* is a mathematical statement which states the relationship between the quantity supplied of a commodity and its price. Supply function is based on the law of supply. The law of supply states only the nature of relationship between the price and the quantity supplied. A supply function quantifies this relationship. A short-run supply function may be written as:

$$Q = 10\,P_x$$

(where Q_x denotes the quantity supplied of commodity X per unit of time and P_x denotes its price).

Given the supply function, a supply schedule can be obtained by substituting numerical values for P_x. For example, if $P_x = 2$, $Q_x = 20$ and if $P_x = 5$, $Q_x = 50$. By plotting the supply schedule, a supply curve can be obtained.

4.6.3 Elasticity of Supply

Like the law of demand, the law of supply states only the nature of relationship between the change in the price of a commodity and the quantity supplied thereof. The law does not quantity the relationship. The quantitative relationship is measured by the price elasticity of supply.

The price elasticity of supply is the measure of responsiveness of the quantity supplied of a good to the changes in its market price. The coefficient of price elasticity of supply (e_p) is the measure of percentage change in the quantity supplied of a product due to a given percentage change in its price. The formula of supply elasticity is given as:

$$e_p = \frac{\text{\% change in quantity supplied } (Q)}{\text{\% change in price } (P)}$$

$$e_p = \frac{\Delta Q/Q}{\Delta P/P} = \frac{\Delta Q}{\Delta P} \cdot \frac{P}{Q}$$

NOTES

Note that the formula for measuring the price elasticity of supply is the same as for the price elasticity of demand, without a minus sign. Given the formula, price elasticity of supply can be easily measured.

Example: Suppose that the supply curve for a commodity is given an SS' in Figure 3.20 and we want to measure the price elasticity of the supply for a price rise in price between points *J* and *P*. In that case:

$$\Delta Q = 60 - 100 = -40$$

$$\Delta P = 5 - 7.5 = -2.5$$

$$P = 5 \text{ and } Q = 60$$

By substituting these values into the elasticity formula, we get:

$$e_p = \frac{\Delta Q}{\Delta P} \cdot \frac{P}{Q} = \frac{-40}{-2.5} \cdot \frac{5}{60} = 1.33$$

Consider another example. Suppose we want to measure the price elasticity of supply between points *P* and *K*, i.e., for price rise from ` 7.5 to ` 10. Here,

$$e_p = \frac{100 - 120}{7.5 - 10} \cdot \frac{7.5}{100}$$

$$= \frac{-20}{-2.5} \cdot \frac{7.5}{100}$$

$$= 0.6$$

The price elasticity of a supply curve like the one given in Figure 3.20 may vary between zero and infinity depending on the levels of the supply. For example, as we have seen above, $e > 1$ between point *J* to *P* and $e < 1$ between point *P* to *K*. It can be noted that the price elasticity below point *P* is greater than unity and it is less than unity beyond point *K*. Thus, a supply curve is said to be (i) elastic when $e > 1$, (ii) inelastic when $e < 1$, and (iii) unitary elastic when $e = 1$. A perfectly inelastic supply has $e = 0$ throughout its length and is a straight vertical line. A perfectly elastic supply curve has $e = \infty$ all along its length is a straight horizontal line.

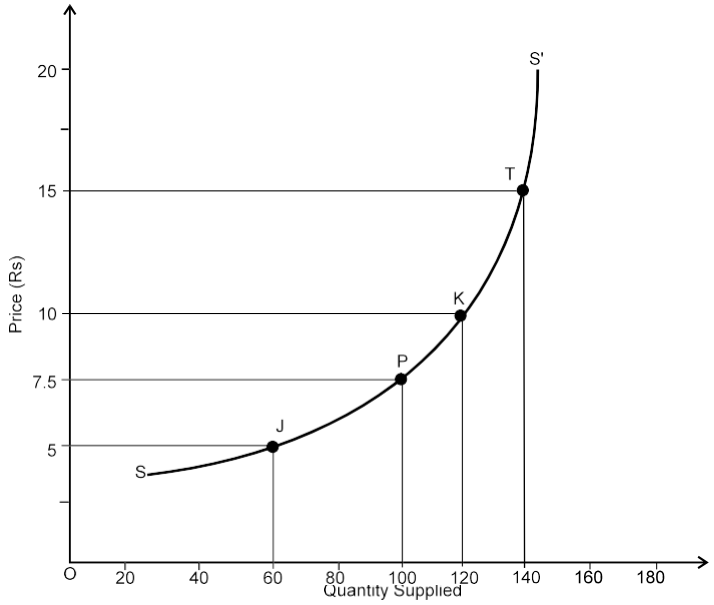


Fig. 3.20 *Price Elasticity of Supply Curve*

Determinants of the Price Elasticity of Supply

The price elasticity of the supply depends on the following factors:

Time Period: Time period is the most important factor in determining the elasticity of the supply curve. In a **very short period**, the supply of most goods is fixed and inelastic. In the short run, the supply tends to remain inelastic. In the **long run**, the supply of all the products gains its maximum elasticity because of increase in and expansion of firms, new investments, improvement in technology, and a greater availability of inputs.

It is important to note here that short and long periods are not fixed in terms of days, months or years. They vary, depending on the nature of the product. For example, for the supply of perishable commodities like milk and fish in a city, a week's time may be a short period. For agricultural products, 6 months may be a short period. But in regard to the local supply of petroleum products in India, a period of five years or even more may be regarded as a short period.

Law of Diminishing Returns: The other factor that determines the elasticity of supply is the Law of Diminishing Returns. If the law of diminishing returns comes in force at an early level of production, cost increases rapidly. As a result, supply tends to becomes less and less elastic.

4.7 SUMMARY

In this unit, you have learnt that,

- Supply of a product depends on the production of the commodity. Production of the commodity depends on the availability of inputs (labour and capital) and technology. Given the technology, output of a product depends on the relationship between the output and input. The input-output relationship is brought out by the theory of production.
- 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.
- 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.
- The economists have 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 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.
- An important concept used in discussions on production theory, though not much of theoretical importance, is average (physical) product.
- The law of diminishing returns states that when more and more units of a variable input are used with a given quantity of fixed inputs, the total output may initially increase at increasing rate and then at a constant rate, but it will eventually increase at diminishing rates.

NOTES

Check Your Progress

13. What is a supply curve?
14. What are the non-economic factors causing shift in the supply curve?
15. What is the supply function?
16. What is the price elasticity of supply?

NOTES

- The laws of returns to scale state the behaviour of output in response to a proportional and simultaneous change in inputs. Increasing inputs proportionately and simultaneously is, in fact, an expansion of the scale of production.
- When scale of production is expanded by increasing all the inputs, the productivity of indivisible factors increases exponentially because of technological advantage. This results in increasing returns to scale.
- When the increase in output is proportional to the increase in inputs, it exhibits constant returns to scale.
- The decreasing returns to scale are attributed to the diseconomies of scale. The most important factor causing diminishing returns to scale is 'the diminishing return to management', i.e., managerial diseconomies.
- Business decisions are generally taken on the basis of money values of the inputs and outputs. Inputs multiplied by their respective prices and added together give the money value of the inputs, i.e., the cost of production.
- Business costs include all the expenses that are incurred to carry out a business. The concept of business costs is similar to the actual or real costs.
- The cost-output relations are determined by the cost function and are exhibited through cost curves. The shape of the cost curves depends on the nature of the cost function.
- The short-run cost curves are helpful in showing how a firm can decide on the optimum utilization of the plant—the fixed factor, or how it can determine the output level that minimizes cost. Long-run cost curves, on the other hand, can be used to show how a firm can decide on the optimum size of the firm.
- Profit maximization has been the most important assumption on which economists have built price and production theories.
- The conventional economic theory assumes profit maximization as the only objective of business firms. Profit maximization as the objective of business firms has a long history in economic literature.
- Marginal revenue is the revenue obtained from the production and sale of one additional unit of output and marginal cost is the cost arising due to the production of one additional unit of output.
- Average revenue is the revenue generated per unit of output sold. It plays a role in the determination of a firm's profit. Per unit profit is average revenue minus average (total) cost. A firm generally seeks to produce the quantity of output that maximizes profit.
- The first-order condition of maximizing a function is that its first derivative must be equal to zero.
- Traditional theory assumes profit maximization as the sole objective of a business firm. In practice, however, firms have been found to be pursuing many objectives other than profit maximization.
- The separation of management from ownership gives managers an opportunity and also discretion to set goals other than profit maximization. It is argued that large firms pursue such goals as sales maximization, maximization of managerial utility function, maximization of firm's growth rate, making a target profit, retaining market share, building up the net worth of the firm, and so on.

NOTES

- A supply curve is a graphical presentation of the supply schedule. Asupply curve has a positive slope. The positive slope or upward movement of the supply curve is caused by the rise in cost of production and seller's effort to make a larger profit.
- Factors like labour strikes and lock-outs, war, drought, flood, communal riots, epidemics, etc. also adversely affect the supply of commodities and make the supply curve shift leftward.
- The supply function is a mathematical statement which states the relationship between the quantity supplied of a commodity and its price. Supply function is based on the law of supply.
- The price elasticity of supply is the measure of responsiveness of the quantity supplied of a good to the changes in its market price.
- Time period is the most important factor in determining the elasticity of the supply curve. In a very short period, the supply of most goods is fixed and inelastic. In the short run, the supply tends to remain inelastic. In the long run, the supply of all the products gains its maximum elasticity because of increase in and expansion of firms, new investments, improvement in technology, and a greater availability of inputs.
- The other factor that determines the elasticity of supply is the Law of Diminishing Returns. If the law of diminishing returns come in force at an early level of production, cost increases rapidly. As a result, supply tends to becomes less and less elastic.

4.8 KEY TERMS

- **Production:** 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.
- **Actual/Explicit costs:** They are those costs which are actually incurred by the firm in payment for labour, material, plant, building, machinery, equipment, travelling and transport, advertisement, etc.
- **Implicit/Imputed costs:** In contrast to explicit costs, there are certain other costs that 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**.
- **Out-of-pockets costs:** The items of expenditure that involve cash payments or cash transfers, both recurring and non-recurring, are known as **out-of-pocket costs**.
- **Variable costs:** They are those which vary with the variation in the total output.
- **Total cost (TC):** It is the total actual cost incurred on the production of goods and service. It refers to the total outlays of money expenditure, both explicit and implicit, on the resources used to produce a given level of output.
- **Marginal revenue:** It is the revenue obtained from the production and sale of one additional unit of output and marginal cost is the cost arising due to the production of one additional unit of output.
- **Total revenue:** In economics it refers to the total receipts from sales of a given quantity of goods or services.

NOTES

- **Average revenue:** It is the revenue generated per unit of output sold.
- **Supply function:** It is a mathematical statement which states the relationship between the quantity supplied of a commodity and its price.
- **Price elasticity of supply:** It is the measure of responsiveness of the quantity supplied of a good to the changes in its market price.

4.9 ANSWERS TO 'CHECK YOUR PROGRESS'

1. 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.
2. The economists have 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.
3. 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.
4. The law of diminishing returns states that when more and more units of a variable input are used with a given quantity of fixed inputs, the total output may initially increase at increasing rate and then at a constant rate, but it will eventually increase at diminishing rates.
5. There are three kinds of returns to scale:
   (i) Increasing returns to scale
   (ii) Constant returns to scale
   (iii) Diminishing returns to scale
6. A factor causing increasing returns to scale is higher degree of specialization of labour, manager and machinery, which becomes possible with increase in scale of production.
7. Variable costs are those which vary with the variation in the total output. Variable costs include cost of raw material, running cost of fixed capital, such as fuel, repairs, routine maintenance expenditure, direct labour charges associated with the level of output, and the costs of all other inputs that vary with output.
8. The cost-output relations are determined by the cost function and are exhibited through cost curves. The shape of the cost curves depends on the nature of the cost function.
9. The fixed costs include: (*i*) costs of managerial and administrative staff, (*ii*) depreciation of machinery, building and other fixed assets, (*iii*) maintenance of land, etc. The concept of fixed cost is associated with the short-run.
10. Marginal revenue is the revenue obtained from the production and sale of one additional unit of output and marginal cost is the cost arising due to the production of one additional unit of output.
11. Average revenue is the revenue generated per unit of output sold. It plays a role in the determination of a firm's profit. Per unit profit is average revenue minus average (total) cost. A firm generally seeks to produce the quantity of output that maximizes profit.

12. The first-order condition is generally known as necessary condition. A necessary condition is one that must be satisfied for an event to take place.
13. A supply curve is a graphical presentation of the supply schedule.
14. Factors like labour strikes and lock-outs, war, drought, flood, communal riots, epidemics, etc. also adversely affect the supply of commodities and make the supply curve shift leftward.
15. The supply function is a mathematical statement which states the relationship between the quantity supplied of a commodity and its price. Supply function is based on the law of supply.
16. The price elasticity of supply is the measure of responsiveness of the quantity supplied of a good to the changes in its market price.

NOTES

4.10 QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is a production function?
2. How does a production function serve a useful purpose in production analysis?
3. Suppose a production function is given as follows.

   $Q = 10L + 5L - L^3$

   Find the following:
   (a) *TP, MP* and *AP* schedules;
   (b) *TP* where *MP* = *AP*; and
   (c) Labour *(L)* required to maximize output.
4. State the factors behind the laws of returns.
5. What are the factors leading to increasing returns to scale?
6. Write a note on total, average and marginal cost.
7. How can the cost curves be applied to the law of diminishing returns?
8. What is the controversy over profit maximization objective?
9. What is the supply curve? What are the determinants of supply that causes shifts in the supply curve?

Long-Answer Questions

1. Discuss production function as a tool of analysis.
2. 'Returns to a factor relate to the short-period production function.' Describe.
3. Explain the laws of returns to scale.
4. Assess the cost-output relations through cost curves.
5. Discuss the problems of cost minimization in managerial decision-making.
6. Critically analyse the importance of profit maximization assumption and theoretical conditions of profit maximization.
7. Evaluate the supply curve and the shift in the supply curve.
8. What is elasticity of supply? What are the determinants of the price elasticity of supply?

4.11 FURTHER READING

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

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

Keating, B. and J. H. Wilson. 2003. *Managerial Economics: An Economic Foundationfor Business Decisions*, 2nd Edition. New Delhi: Biztantra.

Mansfield, E., W. B. Allen, N. A. Doherty and K. Weigelt. 2002. *Managerial Economics: Theory, Applications and Cases*, 5th Edition. NY: W. Orton & Co.

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

Salvantore, Dominick. 2001. *Managerial Economics in a Global Economy*, 4th Edition. Australia: Thomson-South Western.

Thomas, Christopher R. and Maurice S. Charles. 2005. *Managerial Economics:Concepts and Applications*, 8th Edition. New Delhi: Tata McGraw-Hill.

Adhikary, M. 2000. *Business Economics.* New Delhi: Excel Books.

Baumol, W. J. 1996. *Economic Theory and Operations Analysis,* 3rd edition. NewJersey: Prentice-Hall.

Chopra, O. P. 1985. *Managerial Economics.* New Delhi: Tata McGraw-Hill.Kautsoyiannis, A. 1991.

Modern Microeconomics. New York: Macmillan.

Institute of Distance Education
Rajiv Gandhi University

A Central University

Rono Hills, Arunachal Pradesh

Contact us:

+91-98638 68890 | Ide Rgu | Ide Rgu | helpdesk.ide@rgu.ac.in