UNIT 1 CENTRAL PROBLEMS

Structure

1.0 OBJECTIVES

This unit would enable you to understand

- what does constitute an economic problem to an individual and to the society
- what does constitute wants and the means to satisfy them
- what is Production Possibilities Curve (PPC)
- how PPC can be used to illustrate the problem of choice in Economics
- the distinction between Positive Economics and Normative Economics with illustrations.

1.1 INTRODUCTION

The purpose of this unit is to define the subject matter of economics, highlighting what constitute the central (essential) problems of an economy. The normal tendency is to start with a definition. However, in introducing the subject called economics, starting with a definition is an extremely difficult task, if not impossible, since there exists no single accepted definition. You will be surprised if we tell you that there are as many definitions of economics as there are number of economists in this world. Each will claim that his or her definition is superior and more relevant and appropriate than that of others. Hence there is a dilemma as well as an ego problem. However, to do justice to the beginner students who will be reading our unit in economics, we have to start with an analysis of what the subject of economics is all about. It is in this context that we choose Lionel Robbin's definition of economics in preference to Adam Smith's, Alfred Marshall's or J.M. Keynes'.

1.2 DEFINING ECONOMICS.

According to Lionel Robbins, "Economics is a science which studies human behaviour as a relationship between ends and scarce means which have alternative uses" [An Essay On The Nature And Significance Of Economic Science 1933]. What does Robbins stress in his definition of economics? Firstly, that economics as a subject deals with human behaviour. A critic can say that a study of human behaviour is not a prerogative of economics only. There are other social sciences, like sociology, psychology, political science etc., which also deal with human behaviour. Like economics these subjects also deal with the behaviour of people in their individual

as well as. Economics, however, deals with the behaviour of people in the pursuit of economic activities. As one author puts it, "What distinguishes economics from the other social sciences, however, is the manner in which it studies people, and Robbins' statement makes this clear. Economics interprets human behaviour as the conscious attainment of objectives, of ends".

1.3 UNLIMITED WANTS AND LIMITED MEANS

The starting point of economics is human wants, needs and desires. Human beings living in any society at any stage of historical development of that society have wants and needs. Some of these needs like the need for food, clothing, shelter are biological in character, needed to sustain life in this universe. The origin of needs is, therefore, biological. However, the majority of human needs arise from the fact that people live together in a society. It is the existence of human society, which accounts for a large chunk of our needs. Such needs are determined by a complex set of factors called the culture of a given society. Even biological needs (like food) are also influenced by the culture of the society at any particular stage of historical development of that society. Hence, we find that the origin of needs is the biological necessity of sustaining life, while the existence of human society would determine its form, nature, and structure. One important characteristic of human wants is that they keep on occurring again and again. For instance, eating some thing may satisfy one's hunger for the moment but after some time, one has to eat again. Now people would like to see their needs, wants, desires, aspirations to be realised and it is towards this realisation that human activities are directed. Economics is concerned with human (individual, collective) goals, objectives, ends to be achieved and realised through the use of certain means or resources at their disposal. For instance if you want to have a cold drink, you have to have purchasing power to buy it. To produce wheat you must have a piece of land, seeds, fertiliser and irrigation water. To build a house you need bricks, cement and steel, glasses, woods etc. In these examples can you identify the objectives (ends) and the means (resources)? Consider another example: you might drive your car with the objective of getting to work; you might go to work with the objective of earning money (purchasing power); and you might earn money because you have the objective of buying a new car. From these it is clear that to realise goals, objectives, ends, one requires means or resources.

Check Your Progress 1

(1) If you have Rs. 1000 income per month, name some biological needs that you will satisfy first. (In other, words name some commodities which you regard essential for survival)

.....

.....

.....

.....

(2) If you want to grow sugarcane on a piece of agricultural land that you own, what are the resources that you will employ as means to achieve a certain level of output of sugarcane?

.....

.....

.....

1.4 SCARCITIES AND CHOICE

Recall that Robbins' definition stresses the fact that means which human beings use to satisfy their wants or fulfil their needs are scarce. But then one can ask what is so important about scarcity? In economics scarcity has to be understood in the relative sense, the scarcity of means in relation to ends. It is the imbalance between ends and means (whether to an individual or to a society) that gives rise to scarcity and hence an economic problem. Here we must point out another aspect of Robbins' definition. The means that one uses to satisfy needs are not specific to any particular end. The means have alternative uses. They may be employed (at least potentially) to attain any one of a number of ends. For instance, while writing this piece, I have so far used up about one hour of my precious time. Time represents a general means and each one of us is allocated only a fixed amount (24 hours a day or 60 years of life on an average). If I was not writing this unit, I might have used my time in some other ways - I might have gone to the library (Ratan Tata at Delhi School of Economics) to read a new book on Micro-Economics, or I might have gone to a book shop to scan the latest arrivals, or I might have just watched television in the hope of being entertained and educated on the latest trend in the movie world. Well, the alternative use of time (means) is endless. One can go on expanding the list. However, the point is clear, any scarce means have more than one-way of using them. Hence, a meaningful choice exists. Once you have committed your resources to one use (towards achieving a particular end) you deny yourself the benefits, which you could have got, if you had used the same resources to achieve some other end. Economics not only deals with scarcity but also helps us to exercise meaningful choices, since scarce means can be put to alternative uses. For instance, if you spend your income on bread and butter, you forfeit the opportunity to spend your income on fruits and ice creams; if the economy uses oil to energise power plants, the same oil cannot be used to run locomotives; if you use your time to study economics, you cannot at the same time doing a job and earning an income.

"Herein lies the essence of economics. Economics exists when the resources of an individual or group are insufficient to meet all the demands, objectives of that individual or group at the same time. Economics is concerned with the choices that people make about how best to employ a scarce resource - if scarcity does not exist than neither does economics. Note too, that it is not necessary, as sometimes thought, to assume that human wants are `unlimited' or `infinite'. Possibly they are; the question of whether this is true could make an interesting research topic. All that is required for an economic problem to exist is for scarcity to exist - that is for ends to exceed means". (David Whynes: Invitation to Economics, page 15).

Check Your Progress 2

(i) If you can achieve whatever you wanted to achieve, does a meaningful choice exist?

.....

.....

.....

.....

(ii) If you are traveling from home to your place of work, what kind of choice exists so far as the mode of transport to be used is concerned?

.....

.....

1.5 THE PROBLEM OF CHOICE: AN ILLUSTRATION

It is the conflict between means and ends (including conflict between individual's goals) which gives rise to an economic problem. In such a situation the problem arises as to how best to use the resources to achieve the ends. The central problem of an individual as well as for a society is the allocation of scarce means among competing ends. Economic problems for a society do not exist when either human goals are not in conflict or economic resources (means) are so abundant that there is no need to conserve them. However, the real world is not blessed with superabundant resources or harmonious goals. This is true of the most developed countries of the world like U.S.A., Germany, Japan etc., to the most backward countries like Somalia, Ethiopia, Bolivia, Bangladesh, Albania etc. In other words, scarcity is global as well as local, though some societies could be relatively better off than some others giving the impression that for such societies scarcity does not exist. Hence, every society has to develop a set of institutions to direct the way in which scarce resources are to be allocated. These institutions, which form the society's economic system, would determine *what goods are to be produced? how are they to be produced? and who should obtain them?*

A production possibilities curve (PPC) shows the various combinations of two goods (X1 X2) which the firm can produce using technically most efficient methods of production and allocating resources in an economically efficient manner, with its resources being always fully utilised. It shows, given scarcity of resources and given technology, the maximum output produced of one good, given the output of the other good. It shows how one good can be transformed into another good not physically but via the transfer or shifting of resources from one line of use to another. It shows how food is transformed into clothing or from clothing to food by diverting resources from one use to another. Hence PPC is also called *the transformation curve (see Fig.1.1).*

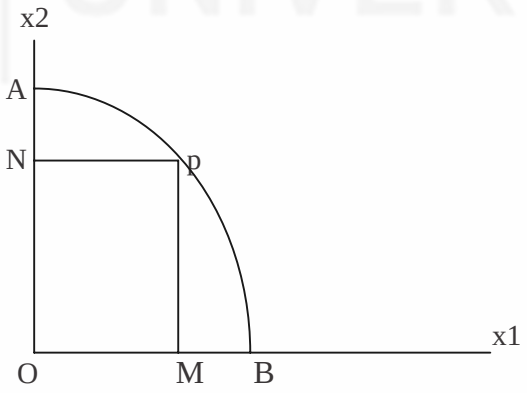


Fig. 1.1: If society devoted all its productive resources in most efficient manner to produce only X_1 commodity it could produce the same in OB amount. If on the other hand, all the resources were devoted to production of X_2 at the most OA amount of the same could be produced. The curve AB shows all the combination of X_1 and X_2, which can be produced by the society. For example, if society decides to produce at point P then, OM quality of X and ON quantity of X_2 is produced. Alternatively, we may say that if society decides to produce OM of X_2 than, maximum amount of X_1 it can produce will be ON (and vice verse).

The technique of PPC can be used to illustrate the various central problems of an economy. The first central problem is the problem of what goods are to be produced by the economy. Let us suppose the two goods are X_1 (consumer good) and X_2

(capital good). Following Paul Samuelson we can say that the economy's PPC describes the society's "MENU of choice". Since the resources needed to produce X_1 and X_2 are scarce and limited, hence every society has to make a choice (take a hard decision) whether to use current resources to make more consumer goods (X_1) or more capital goods (X_2) or seek a proper balance between the two. It is quite possible that if the economy is a developing economy like India, to build up the capital base of the economy it gives more importance to the production of capital goods (X_2), where an economy like Japan with a strong capital base, might give more weightage to the production of consumer goods (X_1). (see Fig. 1.2)

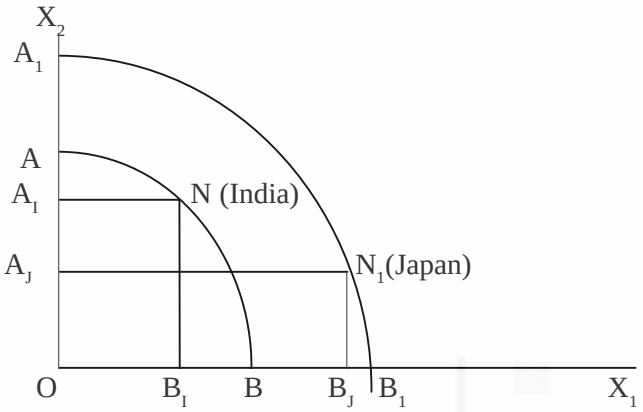


Fig. 1.2: India may produce A_I B_I combination of capital and consumer goods while prosperous Japan may produce A_J B_J Bundles which has a very large amount of consumer good X_1 and small amount of capital good X_2.

Similarly, PPC can be used to illustrate the second central problems of an economy - the problem of how to produce. How resources are to be combined to produce the goods that is the choice of methods of production. However, this cannot be discerned directly from the PPC (we have to go back to the efficiency locus from which PPC is derived). If we assume that X_1 is labour intensive and X_2 is capital intensive, then at point N on India's PPC AB relatively more of capital goods are produced and since they require the use of capital intensive techniques, hence at the product mix N there is preponderance of capital intensive methods of production. However, it must be noted that for the Japanese PPC A_1B_1 the choice of N_1 represents the dominance of consumer goods, which need not imply the preponderance of labour intensive methods. Quite the contrary, in a developed economy like Japan consumer goods are normally produced with relatively capital intensive techniques. Hence PPC can be misleading in indicating the choice of methods of production.

The third central problems of an economy is the problem of for whom to produce - the problem of distribution. The PPC can be used to illustrate it but not directly. If we assume that the more unequal the distribution is, the greater will be the share of luxury consumption (like cars, colour TVs, refrigerators etc). Luxury consumer goods are more capital intensive than the basic consumer goods. Hence Indian economy would be operating at point N on its PPC. But remember, we tried to justify production at N is the name of India being a developing economy a short while ago. Thus, unless we also specify which capital intensive goods are shown on vertical axis, we will not be able to see developing or it is a symptom of very high degree of inequalities in distribution of income. But again we have to be cautious in interpreting the choice of N_1 by the Japanese economy.

Yet if one shows a commodity such as bread on horizontal axis and air conditioners on the vertical axis, the choice of production point will clearly indicate the distributional objectives of those who make the production decisions.

1.5.1 A Digression: Meaning Of Goods In Economy

The material objects (like bread, butter, shirts, trousers, skirts, pens, pencils, chairs, tables, bicycles, cars, watches, bathing and washing soaps etc.), which are used to satisfy human wants are called goods in economics. Goods are material means of satisfying human wants. These means are drawn from nature, by changing their physical, chemical and biological properties, and by conveying them in space and conserving them overtime. We will use the terms goods, products, commodities interchangeably and as synonymous.

Goods are characterised by three things, (i) their physical attributes, the way they satisfy human wants, (ii) the dates at which they are available (a car available this year is not the same commodity as a car available last year even of the same make); (iii) the location at which they are available (a shirt available in New Delhi is not the same as a shirt available in Paris.)

Check Your Progress 3

1. A PPC provides an economy's 'Menu of Choice'. If resources are scarce and fully employed, will the economy be operating on PPC or inside it?

   .....

   .....

   .....

   .....

   (Hint: If you are unemployed you are not producing any output. If you are under-employed you can produce some output but not your potential output that is maximum output).

2. If your resources (means) are expanding (growing) what happens to the PPC? shifts outwardly or inwardly.

   .....

   .....

   .....

   .....

1.5.2 Another Digression

Another aspects of Robins' definition:

He considers economics to be a science. The question is how is it a science? Is it science in the same way *as physics or chemistry*? Not really. Economics is not an exact science like physics or chemistry, since it deals with human behaviour and human behaviour cannot traverse a predetermined fixed pattern, unchanging over time and space. There is no unique human response to a given stimuli. Hence, prediction of human behaviour is much more difficult than prediction of some natural phenomenon. The regularities in natural phenomena are more exact, and more easily quantifiable, representable and measurable than regularities in economic phenomena.

Economics is a science as well as an art. Since it deals with real world phenomena of human activities in production, distribution and consumption, economics is a liberal art. However, the methodology used to derive economic propositions is the methods of science. Now what constitutes a scientific method of inquiry? A

scientific method consists of a set of hypotheses or assertions on objects and phenomena as they are found in reality, and a model which establishes a set of relations among variables or objects or phenomena using the method of logical reasoning (deductive logic). By applying the model to the hypothesis, a set of propositions is derived, which are simplifications or conclusions derived from the hypothesis. At the last step the propositions are put to test by confronting them with the observations on the objects as they are found in reality.

Check Your Progress 4

1. If you throw an apple from a tree top it will always fall on the ground. Your prediction will always be 100 percent correct. However, if you reduce the price of an apple, will people always consume more of it? Do you think that your prediction will always be correct in this case?

   .....

   .....

   .....

   .....

2. In what sense is economics a science?

   .....

   .....

   .....

   .....

1.6 POSITIVE VERSUS NORMATIVE ECONOMICS

In economics we often make a distinction between *positive* issues and *normative* issues.

In positive economics we deal with scientific issues and questions. For instance how is price of a commodity determined? What determines the level of employment in a firm, in an economy? How does a firm choose an optimal input combination? They are all value-free questions and issues, and no ethical issues are involved. In contrast in normative economics we deal only with ethical issues. For instance, what is the best way of allocating resources in an economy? How should National Income be distributed in a society? Is inflation good or bad? The economy ought to operate with full employment of labour! The answer to such questions involves value judgments, which would vary from analyst to analyst. On ethical issues, the stand one takes depends on one's ideology, including one's class affiliation. Since no scientific issues are involved, the axioms as well as the propositions derived from them are ethical in nature. One is free to accept or reject the conclusions of normative economics by just accepting or rejecting the axiom system.

In normative economics, disagreements among economists can never be resolved, since both the hypothesis (that is the axiom system) as well as the propositions derived from it are ethical in nature, involving value judgments. One can go on arguing without coming to an end when one is dealing with ethical questions. Ethical propositions cannot be empirically tested either.

On the other hand though positive economics deals with scientific questions, and hence one can use a scientific method of inquiry to develop a theory, disagreements can still arise among economists. Economists can disagree on the model or the theory to be used to analyse a particular problem. In positive economics since an

economist's perception of a phenomenon or process might differ from other economist's perception, hence disagreements can arise. However, in positive economics such disagreements can be resolved by an appeal to facts, that is, by empirically testing the theory.

Differences can also arise in the nature of the data, which would form the basis of empirical testing of any theory, be in economics or physics. In natural sciences one can make use of controlled experiments in laboratory conditions to generate the database. This is not possible in economics where the analyst will have to rely on using imperfect statistical tools to generate the data required. Hence, the final outcomes of two positive economics theorists handling the same problem may differ.

Check Your Progress 5

(i) For the following statements, indicate which are positive and which are normative.

(a) Unemployment ought to be reduced in the economy.
(b) India's balance of payments is not in equilibrium.
(c) Investment is determined by income.
(d) Inflation ought to be controlled.
(e) It is not desirable to have public sector in Indian economy.
(f) Prices must not be controlled.
(g) Rationing reduces efficiency.
(h) India imports as well as exports petrol.
(i) India is a poor country.
(j) Inequality should be reduced.

1.7 LET US SUM UP

We have discussed in this unit what constitute the central problems of an economy. All of us living in any society at any stage of its historical development, society always face an economising problem. Our desires and wants are unlimited in the sense that even if certain desires are fulfilled some new desire keeps arising with the progress of our society through time. Creation of desires and wants are endless continuous process. However, even if some desires are fulfilled, because of limitations of means we cannot satisfy all our desires at the same time. This is true for any individual in isolation as well as for the society of individuals living in a (given) milieu. Hence Lord Robbins' definition of economics not only attains significance but also is also very appropriate. Economics deals with the act of economising. It is the conflict between limited means and unlimited wants that give rise to scarcity, and hence the economic problem of making a meaningful choice from among the alternative uses of means to satisfy a part of the unlimited human desires. Scarcity is a relative concept and is faced not only by poor country but even the most developed countries of the world. There is scarcity of means in relation to wants in India as well as in U.S.A. But the nature of wants as well as the means could be different. The exercise of a meaningful choice could be illustrated by using the technique of a production possibility curve. The conflict between means and ends (i.e., the wants) gives rise to the following central problems: -

(i) *What goods are to be produced with the limited means?*

Should the economy produce more of food or more of luxury motorcars? Should the economy produce more for exports or more for domestic consumption?

(ii) *How the goods are to be produced?*

In other words how the scare resources are to be combined to produce goods required to satisfy human wants?

This is something to do with the choice of production methods. For instance between labour intensive methods and capital intensive methods.

(iii) *For whom the goods are to be produced*? In other words once the goods are produced for final consumption the problem remains as to how to distribute them to the consumers. Should the goods produced be distributed more to the poorer consumers or more to richer consumers?

We have concluded our discussion of Unit 1 with the distinction between positive economics and normative economics. Just to recollect, positive issues deals with scientific issues, like how price of a commodity is determined? While normative economics deals with ethical issues, like how should a fair wage be determined in the economy? In positive economics no value judgments are involved. In normative economics we essentially deal with the value judgments, which are subjective in nature. Value judgments are ultimately derived from one's ideological position in the society.

1.8 SOME KEY WORDS

Capital : All man made non-labour resources (reproducible) are included in the factor capital. It represents part of society's output, which is not consumed currently but set aside to be used in the productive process sometimes in the future. It includes the stock of man made real physical goods, like machines, plants, equipments, buildings etc.

Economics : The word economics has Greek origin. Oikos Plus Nomos meaning Home and Law. The principle of household management. The word economics has something to do with economising on the use of means to attain ends, since means are scarce and limited.

Ends : The objective pursued by human beings while being engaged in economic activities.

Enterprise (or a firm) : A firm is an organisation formed for the purpose of producing goods and services to be sold in a market. Firms purchase resources (inputs), organise their use in production and market the goods produced. In the process firms take risks. Activities within a firm are based on the principle of co-operation and division of labour.

Goods : Material means of satisfying human wants, desires and needs. The terms goods, products and commodities are used interchangeably and as synonymous in economics.

Labour : Labour as a factor of production includes the mental and physical powers (abilities) of human beings.

Land : All free gifts of nature and natural resources are included in the factor of production land.

Means : The instruments or resources used in attaining the perceived objectives.

Normative Economics : Deals with ethical issues, questions and problems. For instance should unemployed people be given doles? Should government control prices?

Positive Economics : Deals with scientific issues and questions. For instance, what goods are to be produced and how they are to be produced. It analyses and solves economic problems without bringing in value judgments.

Production Possibilities Curve (PPC) : The locus of output combinations which a firm or an economy can produce using technically most efficient methods of production and allocating resources in an economically efficient manner, given production functions, input prices and endowment of resources.

Scarcity : Human needs outstripping availability gives rise to scarcity i.e., the imbalance between ends and means. Scarcity is a relative concept and not an absolute concept.

Wants : Desire for goods, material objects and services. Wants become demand when they are expressed in the market backed by willingness and ability to purchase.

1.9 SOME USEFUL BOOKS

Begg, D.R. Dornbusch, S. Fischer (1991), *Macroeconomics* (4th Edition), McGraw-Hill Book Co. New York

Lipsey, Richard (1997), *Introduction to Positive Economics* (8th Edition), Oxford University Press (ELBS Edition), London

Nicholson, W. (1995), *Intermediate Micro Economics* (VIth Edition), Dryden Press, New York.)

Roychoudhry, Kalyanjit (1999), *Modern Microeconomics* (II Edition), Book Land, Delhi

Salvatore D. (1996), *Micro Economic Theory* (Schaum series 3rd Edition), McGraw-Hill Book Co., New York.

Salvatore D. (1995), *Micro Economics* (2nd Edition), Harper Collins Publishers, New York

Treatment, Timothy (1996), Micro *Economics*, (Its Edition, 1996) McMillan, New York

1.10 ANSWERS OR HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1 Some essential food items milk, cereals, oil, pulses, vegetables, clothes and a small room for living. Since Rs.1000 is a very small income you can afford only small quantities of the above goods. Your consumption habits will determine the weights to be given to the various food items in your budget.

2 Some seeds, some fertilizer, certain amount of water, some labourers, may be some agricultural tools.

Check Your Progress 2

1 Since wants are unlimited and means are limited, all our wants cannot be satisfied simultaneously at any time. Hence you are required to establish

priorities as to which wants are to be satisfied with your limited means, given that means have alternative uses. For instance, on a given piece of land you can grow Rice or Vegetables or grow some cash crops. If means are specific to a particular use then no meaningful choice exists

2 You can either use public transport, you can cycle your way, you can use two wheelers and if you can afford, your motorcars. In some cities people even use local trains, trams, even waterways.

Check Your Progress 3

1. Hint already provided

2 With more resources the economy can produce more of both the goods (in a two goods situation). The PPC will shift outwardly. Of course it would ultimately depend on the technology used, the relative change in factor endowments and the efficiency with which resources are used in the production of the two goods.

Check Your Progress

1 Economics is not an exact science unlike physics and other natural sciences. Human behaviour cannot reveal any regularity unlike natural phenomenon. To the same stimuli the reactions of human beings can be different in different situations and varies from individual to individual. Hence one cannot say with 100% certainty that the quantity demanded of apples would increase when price of apples falls.

2 In the sense of using scientific methods of enquiry. Refer to the diagrams in section 1.5(b)

Check Your Progress 5

1 (a) Normative; (b) Positive; (c) Positive; (d) Normative; (e) Normative; (f) Normative; (g) Positive; (h) Positive; (i) Positive; (j) Normative

UNIT 2 BASIC ECONOMIC CONCEPTS

Structure

2.0 OBJECTIVES

After going through this unit you will be able to explain

- the basic differences between two branches of economic theory: Micro and Macro
- the concept of market and its functioning
- demand function and supply function
- the concepts of equilibrium and disequilibrium and their significance in economics,
- the determination of the price of a good and its quantity traded by using demand-supply model.

2.1 INTRODUCTION

In this unit we will first spell out the distinction between micro and macroeconomics. Since in this block we will be concerned only with microeconomics, which deals only with markets, we start with the concept of market. Demand for and supply of a good form the basic tools of economic analysis when dealing with markets. If a good has market it also has a price attached to it. Hence, the next step will explain the concept of price. In market, transactions take place only at a price. The role of price in the market is to equate demand for and supply of a good so that an equilibrium can be attained. Hence in this unit the last concepts to be introduced are equilibrium and disequilibrium.

2.2 MICRO VS MACRO ECONOMICS

The words *Micro* and *Macro* have Greek origins Mikros and Makros. Mikros implies small and Makros large. Microeconomics is concerned with the most 'Elemental' economic units, like consumer, firm, input, market and industry. In other words, micro-economic theory analyses the behaviour of a consumer or a group of consumers; a firm, an industry, a market; a supplier of an input etc. The unit of analysis is small. In contrast to this, with macro-economic theory the unit of analysis is large. In fact, both micro and macroeconomics are two ways of looking at the same thing, the functioning

or the working of an economy. They are two starting points in analysing how an economy functions or operates. Micro-economic theory focuses attention on individual markets (like the grain market), consumers (say of wheat), firms, industries. It is an in-depth study of how these individual economic units or agents operate or function or make decisions, as well as how they interact with each other. Macro-economic theory, on the other hand, deals with broad aggregates like national income, national expenditure, aggregate consumption expenditure, aggregate investment expenditure, the level of employment, the general price level and so on. It analyses how the economy functions through the interactions of these broad aggregates; how these aggregative variables behave and how they are determined. The distinction between micro and macro is made in terms of the level of aggregation and disaggregation used in analysing the functioning of an economy. Microeconomics uses more disaggregative variables than macroeconomics. Together they form the two sides of the same coin. However, it must be noted that economic decisions are ultimately taken at the micro level, and the conjunction of all micro decisions have important ramifications at the macro level. For instance, when we add consumers' expenditures on all goods and services, we get the aggregate consumption expenditure for the economy as a whole, which is a macro concept. Similarly, the functioning of the economy at the macro level will have bearings for decision-making at the micro level. When income tax is raised, disposable income of households falls, firms will experience a decline in sales and as a result will cut back output. Hence, a macro level event will generate a micro manifestation.

Micro-economic theory will be basically concerned with relative prices of commodities and factors of production (inputs). Hence, we will not be incorrect if we call Micro-economic theory, a price theory. Following *Quirk* we can say, "micro-economic theory provides the framework within which the economist describes and analyses the behaviour patterns and inter-relationships of the elemental economic units like consumer, firms, industries, commodities and markets. The main objective of micro-economic theory is to explain and predict how production, exchange and distribution of goods and services respond to the incentive structure operating in a given society." (James Quirk, - **The Intermediate microeconomics**, Ist Edition).

Check Your Progress 1

1. Which of the following statement pertains to microeconomics?

   (a) When inflation occurs, the Indian economy is able to export smaller volume of textiles.

   (b) When supply of wheat goes up, its price falls.

   (c) When there is strike in Mother Dairy the price of milk goes up.

   (d) An increase in investment increases employment.

2.3 THE CONCEPT OF MARKET IN ECONOMICS

In ordinary language, by market we mean the place where goods and services are bought and sold. Remember the places like Sabzi Mandi at Azadpur and Okhla; Timber Market at Paharganj or Kotla; Grain Market at Khari Baoli; Cloth Market at Chandni Chowk, Car Bazar at Lajpat Nagar, Paper Market at Chawri Bazar and Furniture market at Kirti Nagar all situated in Delhi. However, the meaning of market in economics is different. By market an economist would mean a complex set of activities by which potential buyers are brought into contact with potential sellers in the process of purchase and sales of goods and services. When two or more individuals undertake exchange transactions, a market is established irrespective of time and place. The physical presence of buyers and sellers are not at all relevant in the economist's conception of market. Due to the development of modern mode of communication based on computer technology, physical presence is not required for

establishing a market. What is more relevant is that there must exist two groups of transactors who are willing to undertake exchange transactions.

The essential characteristic of a market is the confrontation between potential buyers and potential sellers, making bids and offers in the process of determining the terms at which exchange would take place. The emphasis is on the word Potential. Every one comes to the market with his or her notion of how much to buy or sell depending on the prevailing prices. If the prevailing price of a good is high, those who have made plans to buy at a lower price will be priced out of the market. Similarly, if price is low those sellers who planned to sell at higher prices will be eliminated from the market. Activities on a market would determine what the price will be, what quantities will be bought and sold, and who the buyers and sellers are. In a market-based economy, it is through the market that resources get allocated among competing ends.

Markets in economics vary from perfectly competitive, to monopoly, to imperfectly competitive, to monopolistically competitive, to oligopolistically competitive.

The existence of markets provides efficiency in the use of resources. In economies where markets do not exist, resources cannot be used efficiently. Markets provide information required for making optimal decisions. The kind of information provided by a market would pertain to the nature of goods traded, the prices prevailing, and who the transactors are. No other information is transmitted by the market. Also information transmitted by markets would be costless, though for many goods one has to spend money, time and efforts to acquire the relevant market information. Modern modes of tele-communication make information gathering, storing and processing much easier. This surely must have improved the quality of decision-making.

Markets can exist and function efficiently if and only if there exist ownership rights in commodities (property rights) which are well defined, transferable and are protected by the laws of the country. Ownership rights or property rights imply that the individual owning a commodity can take certain actions with respect to that commodity. Such rights include the right to allow access to the commodity on the part of any other individual(s), and also include the right to charge others for the use of the commodity or service as well as transferring the ownership of the commodity to someone else. As Quirk says, "What is produced, exchanged and consumed in a society are bundles of property rights that we call commodities. The pattern of economic activity within a society is closely linked to the structure of property rights of that society, because it is through the acquisition of property rights that self-interest is expressed and incentives operate."

2.4 WHAT IS MEANT BY DEMAND?

In economics, by demand we mean effective demand and not absolute demand. The demand for a good by an individual consumer means individual's desire for the good backed by a capacity to pay. Human desires backed by purchasing power constitute demand. In other words, an individual's desire for a good to satisfy a particular want backed up by his/her willingness and ability to pay gives rise to demand for that good. If and only if individuals have means to pay that demand becomes effective in the market for goods. An individual's income measures his/her capacity to pay, purchasing power or means to pay for the goods desired.

Example: A beggar desires milk, but has no purchasing power. Hence a beggar's desire for milk does not constitute an effective demand for milk. As a result a beggar cannot participate in market activities. However, suppose this beggar becomes successful in getting a job, becomes a helper in a shop and for his work as helper gets paid for in money. The beggar who is now a helper earns an income, with which she can buy milk. The beggar's demand for milk, which earlier constituted only an absolute demand, has now become an effective demand. This particular beggar's demand for

milk now adds to the market demand for milk, which a potential supplier of milk has to take into account in deciding how much milk must be produced and supplied. Note that the person's demand will be effective demand even if she had not picked up a job and had paid for milk out of her collection of alms. Hence, for demand (for a good like milk) to exist two conditions must be fulfilled:

(i) individuals must have a desire for that good, and

(ii) their desire must be supported by income or purchasing power or means to pay.

Check Your Progress 2

1. Your monthly income from salary is Rs. 3000. You desire to travel by taxi to and fro from office every day. Does your desire constitute demand for taxi?

2. You are fond of tea. Analyse what happens when price of coffee rises; your income goes up and suddenly some guests come to stay with you?

2.5 THE DEMAND FUNCTION FOR A GOOD

In Section 2.4 we have considered various influences on the demand for a good. Here, it must be pointed out that there could be many other influences working on the demand for a good, which we have not considered. However, the above four factors are the major influences on the demand. In any particular situation if we keep factors other than own price as constant, we can derive a demand schedule, a demand function, and a demand curve. A demand schedule lists the various quantities of a good, which a potential consumer buys from the market at different prices, observed at a given moment of time. Its tabular representation would be as follows:

Price of Milk (In Rs. per litre)	20	15	10 5
Quantity Demanded (In liters)	1	1.5	3 6

A demand function for a good expresses a causal relationship between quantity demanded of the good and its own price. In other words, it is a functional relationship between demand and price. If the good is X (milk in our case), Q^d_x is quantity demand and P_x is the own price of good X then the general form of the demand function will be

$$Q^d_x = f(P_x)$$

What it says is that quantity demanded depends on price. Price is the *cause* and quantity demanded is the *effect*. Stated alternatively, price is the independent variable while quantity demanded is the dependent variable. In technical terms, independent variables are also called *exogenous variables* while dependent variables are called *endogenous variables*. The demand function which expresses the functional relationship between the quantity demanded of a good and its own price is based on *ceteris paribus* assumptions. That is, we only allow the own price of commodity to vary with demand with everything else held constant at their pre-assigned values. In other words, when we try to capture the relationship between quantity demanded of a good and its own price, we ignore all other influences on the demand for that good (like prices of substitute goods, complementary goods, the households income level, tastes and so on).

The graph of the demand function on a two dimensional Euclidean space with horizontal axis measuring the quantity demanded and vertical axis measuring the price, we get the demand curve for the good. Since the normal behaviour is one of an inverse

relationship between quantity demanded of a good and its own price, the demand curve for a good will be downward sloping. For convenience we can assume the demand curve to be a downward sloping straight line. This is illustrated in Fig. 2.1.

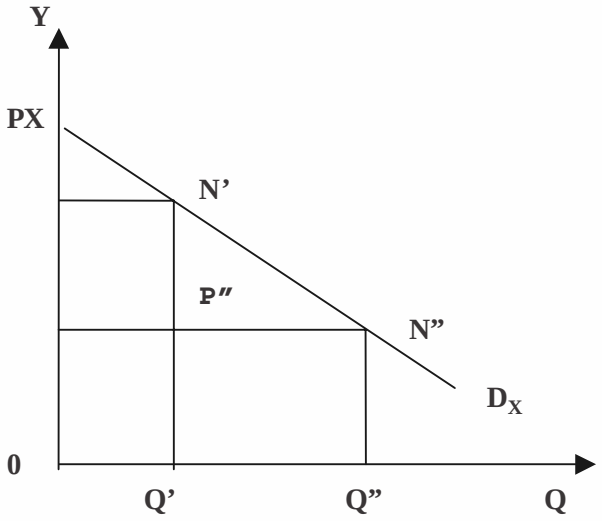


Fig. 2.1: We have shown price of milk in rupees on vertical axis and quantity in litres on the horizontal axis. The consumer buys OQ' quantity at OP' price. As the price falls to OP", she is willing to buy OQ" amount. N' and N" show there are two situations on the demand curve DX.

In the figure the quantity demanded of good X (milk) is plotted on the horizontal axis (abscissa) and the price of good X on the vertical axis (ordinate). The quantity demanded is measured in physical units of the good. In case of milk, physical unit is litres. Price of the good is, however, expressed in monetary units (in rupees or paise). In case of milk, the price of milk would read as Rs. 8 per litre or Rs. 12 per litre etc. A downward sloping demand curve for good X would imply that when price is lowered, the quantity demanded would tend to increase. And when the price is raised, the quantity demand would tend to be reduced. In the diagram when price is OP^1 the consumers of good X operate at point N^1, demanding (purchasing or buying) OQ^1 quantity of good X. When price falls to OP^{11}, the consumers operate at point N^{11} on the demand curve Dx, demanding OQ^{11} quantity of X. A lowering of price induces a larger quantity of it being demanded. This is supposed to be a commonly observed aspect of consumer behaviour. As to the question why demand curve is downward sloping, we defer the explanation to Units No. 4 and 5. *A downward sloping demand curve reflects the law of demand. This law says that, other things remaining the same, a consumer (or, in general consumers) would tend to buy more when price falls and buy less when price rises.*

Here, it can be pointed out that when the demand curve for a good is a straight line, the corresponding demand function will have a linear equation of the type:

$$Q^d_x = a - bP_x$$

where 'a' is the quantity intercept and 'b' is the slope. The slope, b, expresses the rate at which quantity demanded changes (increases or decreases) when price changes (falls or rises). That is,

$$b = dQ / dP$$

When we plot the demand curve, we actually plot the inverse demand curve

$$P_x = \alpha - ß\, Q_x$$

Where $\alpha = a/b$ is the price intercept and $ß = 1/b$ is the slope of the inverse demand curve and equals dP/dQ.

Note: Here it must be pointed out that in economics when we plot diagrams the convention is to plot the independent variable on the vertical axis and the dependent variable on the horizontal axis. In mathematics we do just the opposite.

In its normal form $Q^d_x = a - bP_x$, the demand curve would indicate the maximum quantity demanded at any given price of the good whereas in its inverse from $P_x = \alpha - ß\ Qx$, it would indicate, for each given quantity demanded, the maximum price a consumer (or consumers) would be willing to pay rather than doing without that quantity. The demand curve (or its inverse form) always indicates the maximum boundary to consumers. No consumer will be willing to pay for **OQ^1** quantity a price higher than **OP^1** ($=N^1Q^1$) in the above diagram. Also since price paid by buyers reflects the average revenue (AR) earned by sellers, the inverse demand curve is also known by the name AR curve. Note that the demand curve touches the price axis. It gives us the price at which quantity demanded falls to zero. Such a price is called the prohibitive price **(OP).**

Check Your Progress 3

1. For the following demand function determine what will be the quantity demanded when the price assumes the values Rs. 5, Rs.4, Rs. Rs.3, Rs.2, Rs.1.

   Q = 40 - 0.5 P.

2. How do you interpret the following two functions?

   (a) Q = 100 - 2 P
   (b) P = 50 - 0.5 Q

2.6 WHAT IS MEANT BY SUPPLY

By supply of a good we mean a producer's desire to produce (manufacture) a good backed up by her capacity to produce as determined by technological knowledge and command over employable resources. A commodity (good) can be supplied if and only if it is produced. The concept of supply is, therefore, related to the concept of production.

2.7 THE SUPPLY FUNCTION OF A GOOD

How much of a good gets produced and supplied will depend on many factors. Some of these factors are the own price of the good, the prices of inputs used to produce the goods, the technology used etc. Of these factors the own price of the good is an important factor determining how much of it will be supplied. The supply function captures that relationship. The general form of supply function of a good is the following:

$$Q^s_x = F\ (P_X)$$

Where X is the good. Recall that a function expresses a cause-effect relationship. The above relationship indicates that the quantity supplied varies directly with the price of the good under consideration and the relationship is one of a positive correlation between quantity supplied and own price. As price of good X rises the quantity supplied tends to increase. And when own price of good X decreases the quantity supplied tends to decline. This is supposed to be a normal behaviour on the part of suppliers. Of course, such a relationship is based on *ceteris paribus* assumption. All other influences like technology, prices of inputs and so on are held constant.

A supply schedule of good X indicates the quantities that will be supplied by potential sellers at its prevailing prices. In other words, it lists the various amounts of a good that potential suppliers will put up for sale at the alternate prices prevailing. A supply schedule will look something like the following:

Supply Schedule

Price of Good X (In Rs.)	Quantity of Good X Supplied (In the physical unit of the good)
9	100
8	80
7	75
6	65
5	45

When the above schedule is plotted on a two dimensional graph with price of good X measured on the *ordinate* and the quantity supplied on the *abscissa*, we get the supply curve of good X which will be upward sloping. This is shown below (Fig. 2.2).

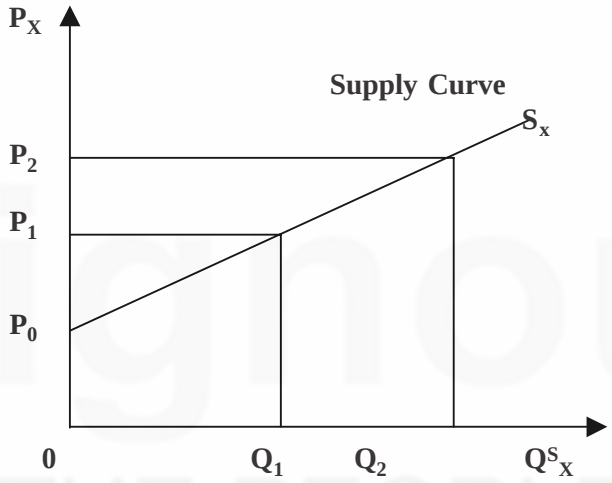


Fig. 2.2 At OP0 price, the producer does not supply anything. As price rises to OP, she will be ready to supply OQ1. A further rise in price to OP2 makes the producers willing to supply OQ2 quantity.

The supply curve can be linear (straight line) or non-linear (curve). In the diagram below we have depicted both the cases. An upward sloping supply curve implies that as price rises the quantity supplied tends to increase. A higher price induces a larger supply.

The supply curve or a **market supply curve** is an aggregate of individual supply curves. Since the supply curve is the graph of supply function, it is also based on c*eteris paribus* assumption, and considers only the relationship between quantity supplied of a good and own price.

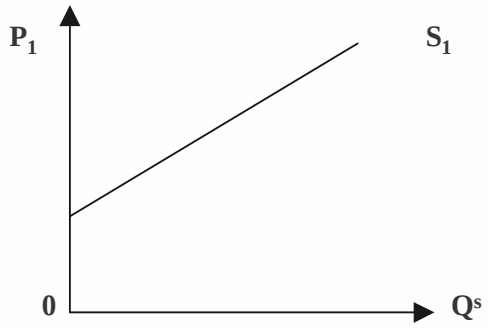


Fig. 2.3a

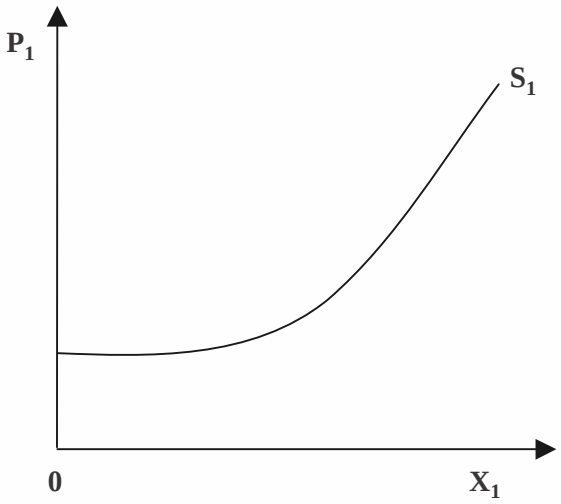


Fig.2.3b

Note that in defining the concepts of demand and supply the stress is on the word potential. Each buyer or seller comes to the market with her notion of how much to buy or sell. If some buyers have made plans to buy at a given price will leave the market if price of the good rises. They will be priced out of the market. Similarly, if market price is low, those sellers who had made plans to sell at a higher price, will be priced out of the market. Hence, the emphasis is on the word potential in the definitions of demand and supply.

Like the demand curve, the supply curve also indicates the planned (expected) behaviour of sellers (or, producers-cum- sellers). At any given price it would indicate the maximum quantity produced and supplied. To put it in a different way, if a given quantity of a good is to be produced and supplied, the supply curve would indicate the minimum price of the good expected by each seller (in order to cover the cost of production and to earn a normal rate of profit).

In the Fig. 2.4, if OQ_1 quantity is to be forthcoming in the market, OP_1 price must prevail. However, if the price is $E_1Q_1 = OP_2$ then an output OQ_1 will not be produced and supplied. At such a price if OQ_1 is produced and supplied the producer-cum-supplier will suffer losses.

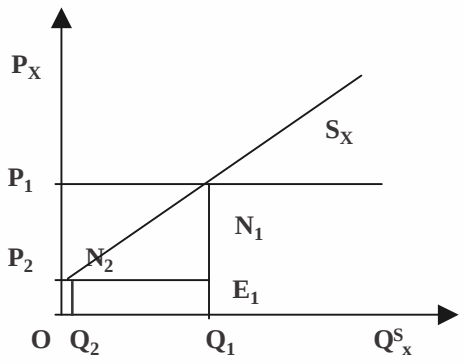


Fig. 2.4 Producer expects a minimum price OP1 for quantity OQ1. At price OP2, if she is asked to supply OQ1, she should refuse as it involves suffering a loss equal to triangular area N1 E1 N2

Check Your Progress 4

1. For the following supply function determine what will be the quantity supplied when price assumes values 1, 2, 3, 4, 5, 6.

   Q = -4 + 4P

2. How do you interpret the following two functions?

   (a) Q = -20 + 4P
   (b) P = 5 + 0.25 Q

2.8 MEANING OF PRICE IN ECONOMICS

In economics we always talk of prices, the price of wheat, the price of milk, the price of a car, prices of vegetables, input prices like wages, interest and rent. If a commodity has a market it will have a price associated with it. A price can be expressed in one of the following two ways. Firstly, we have price expressed in terms of a n*umeraire commodity* (NC). Suppose, we choose gold as the *numeraire* commodity, then the price of a commodity, say, wheat, is the number of units of gold (of a particular specification), which would exchange for one unit of the commodity (that is, one kilogram of wheat). It is to be noted that the price of gold in terms of itself always equals one. In theory any commodity can be chosen as a *numeraire*. However, in practice the choice would depend on a number of attributes, which a *numeraire* commodity must possess. They are: the commodity chosen must be finely divisible, must not be bulky, must not be subject to physical decay, and must be easily transportable. It must be noted that the *numeraire* commodity is not intended to perform the function of money as a medium of exchange, but supposed to function only as a unit of account, in terms of which all values are expressed and measured. Expressed in terms of a *numeraire, prices are commodity rates of exchange-the rate at which each commodity exchanges per unit of the numeraire good. Prices are ratio of the form* X_g/X_i*, where* X_g *is the units of gold (the NC) and* X_i *is the units of commodity i.* As rates of exchange, prices are measured in the physical units of the commodities.

The second way in which prices can be expressed is in terms of an *Abstract Unit* of account used in a bookkeeping sense. It has no physical substance attached to it (unlike the first case). When a unit of commodity is sold, certain number of units of account is credited to one's account, while if a unit of commodity is bought, the same number of units of account is debited from one's account. *The price of the commodity is then the number of units of account credited or debited per unit of the commodity. In the Indian context Rupee is such an abstract unit of account. Price of a commodity will then be the number of Rupees per physical unit of the commodity.* "Notes and coins have no intrinsic worth, but are simply tokens representing number of units of account which are passed around directly and form part of the credit side of one's account. As between the two ways of expressing prices the second way represents the way prices are expressed in reality" (Gravelle and Rees, ***Micro Economic Theory***, 2nd Edition)

Check Your Progress 5

1. If four kg. of wheat exchanges for one kg. of rice, which concepts of price we are referring to?

2. The price of a bicycle is Rs. 1000/-, which concept of price we are referring to?

2.9 MEANING OF EQUILIBRIUM OR DISEQUILIBRIUM

As with many other concepts in economics, the concept of equilibrium is also borrowed from physics, to be precise, mechanics. Equilibrium is described to be a situation

where a body subjected to actions of opposing forces attains a position of rest. That is, the state when opposing forces are in balance. In the same way, *in economics equilibrium is said to exist in a market where the forces operating from the side of potential buyers exactly offsets the forces operating from the side of potential sellers.* This means that when quantity supplied balances (matches) the quantity demanded, the market for that commodity reaches equilibrium. Once again, *for equilibrium to exist, opposing forces must be in balance.* The price at which demand equals supply is the equilibrium price and the corresponding quantity traded, the equilibrium quantity for that price. Such an equilibrium is a static concept, describes the position of the market when it is at rest. So long as the factors on which demand and supply depend do not change, equilibrium will be maintained period after period. This is one definition of *equilibrium,* describing it *as a position of rest.*

There is another concept of equilibrium, which is used in neo-classical economics, that is, equilibrium defined as a chosen position of individuals (the participants in a market transactions). As Gravelle & Rees put it, "*Equilibrium as being that state in which economic agents find themselves in those positions in which they wish to be*" (Gravelle & Rees). According to this definition, a market is in equilibrium when at a given price all economic agents can buy what they had planned to buy and all sellers can sell what they had planned to sell and the chosen position of buyers and sellers coincide.

It is quite possible that the two ways of defining equilibrium may or may not coincide. In the demand-supply model of price determination under perfect competition that we will be developing later, the point where demand schedule and supply schedule intersects is the point where both the definitions of equilibrium will coincide. In static micro theory we will be concerned with equilibrium configurations of markets.

When the market is not in equilibrium, it must be in disequilibrium, which reflects a situation in which demand and supply forces are not in balance. What the buyers plan to buy do not match what the sellers plan to sell. So the market is not cleared. Either unsold stocks of inventories remain or a shortage develops which requires interventions by authorities with previously accumulated stocks. When the market is in a disequilibrium situation, the market adjustment process or corrective process must be brought into the picture, to analyse how the situation can be corrected. Such an analysis must be dynamic in nature, i.e., it involves movement of the market through time. When the market is not in equilibrium the role of economic agents and decision-makers become very important and relevant. In actual situation, equilibrium is never attained. As the market tends towards equilibrium, disturbances occur due to dynamic changes always taking place in the economy, which prevents equilibrium being attained. Hence, disequilibrium analysis is more relevant. However, equilibrium analysis is simpler and easier to handle.

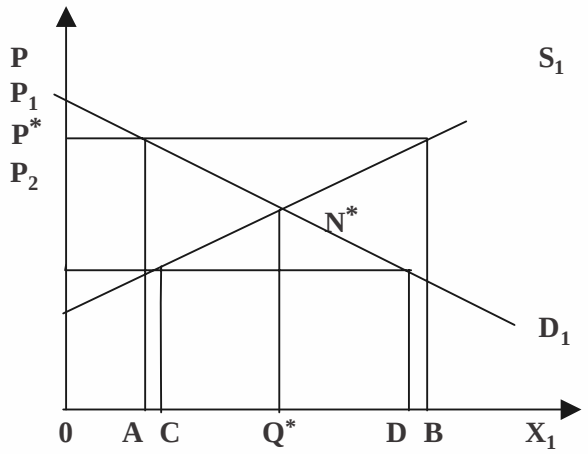


Fig. 2.5 Equilibrium price and quantity are OP* and OQ* respectively. At a higher price OP1, producers are ready to supply OB units while buyers demand only OA. Thus, excess supply equal to A, emerges. On the other hand, if price falls to OP_2, buyers demand OD but sellers are willing to supply only OC. Hence, there emerges an excess demand in the market (=CD).

Check Your Progress 6

1. Consider the following demand and supply schedules

Price (in Rs.)	Quantity demanded (in kg.)	Quantity supplied (in kg.)
0	100	0
10	90	0
20	75	40
30	65	65
40	40	80
50	30	90
60	15	110

What is equilibrium price and quantity?

1.6 LET US SUM UP

We have started this unit with a distinction between microeconomics and macroeconomics. They represent two different ways of looking at the functioning of a market economy. Only the degree of aggregation or disaggregation is different. Microeconomics is a more disaggregated study of economics where individual economic agents and individual markets are the focus of attention. In Macroeconomics individual agents as well as individual market are aggregated into broad categories like consumption, investment, employment, and markets like goods market, money market and factor markets. Microeconomics is based on partial equilibrium analysis where each market is studied in isolation of the rest of the economy. Macroeconomics, on the other hand, is based on general equilibrium analysis where all markets are interrelated (see Unit-4).

Since economics, whether micro or macro, deals with markets, we have tried to explain what the concept of market means to an economist as different from that to a layman. Market does not refer to any physical area where buyer or seller congregate to bargain and haggle to obtain better terms for themselves. To a layman that is what market means. To an economist a market for good exists whenever two or more persons are willing to undertake exchange transactions irrespective of place and context. In this sense market can have no physical boundaries. This is made possible by the development and use of modern technology in transport and communication.

In markets, transactions are made by potential buyers and sellers. As a result the buyers demand function and the sellers supply function constitute two market forces to determine the equilibrium market price and the equilibrium market quantity traded (bought and sold). We have also defined the concept of price, which is always expressed in items of a unit of account whether real or abstract. In the Indian context price of a good is expressed in terms of Rupees. The number of rupees given up to acquire a unit of a good is defined as a price of a good.

In economics, equilibrium price is determined whenever quantity demanded of a good equals quantity supplied. At the equilibrium price the market is cleared in the sense that whatever buyers wish to buy matches whatever sellers wish to sell. Given that market demand curve for a good is downward sloping and market supply curve of the good is upward sloping, at any other price the market for the good will be in

disequilibrium. At such a price either demand exceeds supply or the supply exceeds demand. Since demand and supply do not balance when the market is in disequilibrium, an excess demand pressure or an excess supply pressure will develop which will move the market price towards the equilibrium price. This happens only when demand and supply curves are normally shaped. Otherwise, even if equilibrium exists, it may not be a stable equilibrium.

1.7 SOME KEY WORDS

A Demand Curve : Captures the relationship between the quantities of a good which consumers would be willing to purchase at alternative prices, *ceteris paribus*.

Demand Function : The demand function of a good expresses a relationship between the quantity demanded of a good and its own price *ceteris paribus*. It is derived from the utility function by applying an optimisation rule. It reflects the consumer's maximising behaviour.

Demand : Means human desires backed by willingness and ability to pay. It means effective demand and not absolute demand.

Disequilibrium : Defined as a situation in which opposing forces are not in balance, that is, when the forces of demand and of supply are not in balance.

Equilibrium : Defined as a situation in which opposing forces are in balance. A market attains equilibrium whenever quantity demanded equals quantity supplied. The forces of demand and supply are in balance.

Macro : A large unit of analysis.

Market : When two or more individuals undertake exchange transactions a market is established irrespective of time and place. It is not restricted to a particular area or a place. The development of modern communication technology has extended the spatial limits of a market.

Micro : A small unit of analysis.

Price : The rate at which each and every commodity exchanges per unit of the *numeraire* good (unit of account).

Supply and Supply Function : By supply we mean how much of a good will be produced and offered for sale at alternative prices given technology, input prices and resources. The supply curve captures such a relationship given *ceteris paribus* conditions.

2.12 SOME USEFUL BOOKS

Same as in Unit-1.

2.13 ANSWER OR HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1. (a) Micro Economics

(b) Micro Economics
(c) Micro Economics
(d) Macro Economics

Check Your Progress 2

1. With a monthly salary of Rs. 3000 you cannot afford to travel in Taxies. Your desire for Taxi ride does not constitute effective demand.

2. The demand curve for tea *shifts* upward to the right. In other words, at the same price you tend to buy more. This phenomenon is known as *increase in demand*.

Check Your Progress 3

1.

Price (in Rs.)	Quantity demanded Q = 40 - 0.5P. (in physical unit like Kg.)
5	37.5
4	38.0
3	38.5
2	39.0
1	39.5
0	40.0

2. (a) At any given price this equation indicates the quantity demanded. Hence, it is called the demand equation (function).

(b) This would indicate for any given quantity demanded the maximum price the buyers would be willing to pay rather than doing without the units. Hence, the equation would indicate the inverse demand curve.

Check Your Progress 4

1.

Price (in Rs.)	Quantity Supplied Q = - 4 + 4P (in physical units like kg.)
1	0
2	4
3	8
4	12
5	18
6	20

2. (a) This equation would indicate at any given price the quantity supplied by the sellers. Hence, it constitutes the supply curve.

(b) This is the inverse supply curve. It would indicate the minimum price at which each unit of the good will be supplied in the market, (for instance to cover the cost of production).

Check Your Progress 5

1. Price as an exchange ratio. This is expressed in terms of a unit of account. (Wheat is the unit of account)

2. It is the money price of the good. It is expressed in terms of an abstract unit of account called Rupee.

Check Your Progress 6

1. Equilibrium Price is Rs. 30; the equilibrium quantity demanded and supplied is 65 kg.

UNIT 3 METHODS OF ECONOMIC ANALYSIS

Structure

3.0 OBJECTIVES

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

- difference between partial and general equilibrium approaches;
- the difference between static and dynamic methods of analyses;
- the difference (if any) between an economic theory and an economic law;
- the difference between a stock variable and a flow variable and
- the art of constructing economic theory by using the principles of scientific method of enquiry.

3.1 INTRODUCTION

Continuing with the discussion on methodology of economic analysis, we, in this *unit go through the discourse on the distinction between partial and general equilibrium analysis. The focus of partial approach is on single market in isolation, while in the general equilibrium approach it is on simultaneous working of all markets in an economy in an interwoven manner*. An economist has to take recourse to both the methods depending on the types of problems and issues that are to be analysed. As we proceed in the text we will have more to say on this subject. On the question of methodology itself, an economic analyst would also like to make a distinction between static and dynamic methods of analysis. Like partial equilibrium approach, the static method is easier to handle than dynamic analysis. We will complete our discussion on methodology by describing how economic theories are constructed and verified. Remember that we described economics as a science in Unit 10. Here, a scientific method of enquiry will be used to construct and verify an economic theory. Subsequently, we will explain the difference between a theory and a law. Are economic laws immutable? I can be pointed out that the purpose of constructing a theory is both to provide an explanation of the phenomenon that is under study and to using it for making predictions of events which have not yet occurred in the economy. The discussion in this unit is closed by describing the difference between a stock variable and a flow variable. Most of the variables we deal with in Micro and Macroeconomics are flow variables. However, market exists for flow as well as stock variables.

3.2 PARTIAL VERSUS GENERAL EQUILIBRIUM ANALYSIS

In partial equilibrium analysis, we concentrate on a single market, in isolation from the rest of the economy. We analyse in detail a particular market or a set of markets neglecting everything else. For instance, when we want to study the market for wheat in detail, we do not bother about other markets in the economy. Such an analysis is based on c*eteris paribus* assumption. Demand and supply models of price determination of a good is based on partial equilibrium analysis. It ignores various linkages and inter-relationships that might exist between different markets. On the other hand, in general equilibrium analysis we analyse simultaneously all the markets in the economy. The basic premise in such an analysis is that, "everything depends on everything else". *All the markets of the economy are interdependent and interrelated so that a disturbance originating from any one market will have repercussions throughout the economy*. In such a situation general equilibrium analysis is the correct approach for analysing the functioning of the economy. In fact, partial and general equilibrium analyses are two ways of looking at the functioning of the economy.

Partial equilibrium analysis is appropriate when we want to analyse in detail the functioning of a particular market or a particular sector of the economy. It is used when a market is self-contained or insulated from other markets or when the market in question is relatively small, relative to the size of the economy, or when the cross-effects generated by this particular market are negligible and hence can be ignored. Partial equilibrium analysis makes the analysis of a problem more manageable, unlike general equilibrium analysis which is often difficult to comprehend. Reality is so complex that one needs a process of simplification (abstraction) to understand it. Partial approach is one such form of simplification, where each market is viewed in isolation. Partial equilibrium analysis was championed by Alfred Marshall (1890) and is based on "*ceteris paribus*" assumption. Such an assumption abstracts from all interconnections and inter-links that exist between the market under study and the rest of the economy. For instance, we use demand-supply model to show how equilibrium price and quantity is determined in each market, independently of other markets. However, we know very well that a change originating from any market has spillover (repercussions) effects on other markets. When these changes in other markets (sectors or industries) are significant, the partial equilibrium analysis is inappropriate and inadequate. By taking into account only the direct effect on price and quantity, partial equilibrium approach, "provides a misleading measure of the total, final effect, after all the repercussions or feedback effects from the original change have occurred." If and only if the market or the sector (industry) from which the original change occurs is relatively small and has very few linkages with the rest of the economy, the partial equilibrium analysis would be the right approach to study the operation of market system . Otherwise a general equilibrium approach is needed.

When market (economic) interdependencies or interrelationships are not taken into account, or do not exist, partial equilibrium analysis is the correct approach. However, when such interrelationships and interdependencies exist and are important, and the ignorance of which will have serious consequences or will prove costly in terms of the quality of economic predictions, a general equilibrium analysis must be used. It must be used whenever an event has all pervading effect.

Check Your Progress 1

1) If you want to study in details the working of the market for milk in your city, which methodology will you use?

2) As demand for automobiles goes up, the demand for steel goes up, which in turn

increases the demand for aluminum and so on. Is this an example of partial equilibrium approach or general equilibrium approach ?

3.3 STATIC VERSUS DYNAMIC METHOD OF ANALYSIS

Economic analysis can be conducted either by using a static framework or a dynamic setting. Static and dynamic modes of analysis can be differentiated in more than one ways. According to one definition, in a static model (theory) the variables (cause-effect) are not dated. The demand-supply model of market behaviour is a static model. The model that demand depends on own price, supply depends on own price, with an equilibrium condition that demand must equal supply, time does not enter into the picture at all and the variables are all undated. According to this definition, a dynamic model would be one where the relevant variables are dated. If the demand-supply model is restructured as follows, then the model would become dynamic according to this criterion.

$$D_t = f(P_t)$$

$$S_t = g(P_t)$$

$$D_t = S_t$$

where 't' is the relevant time unit.

However, according to some economists, even if the variables are dated the model does not become dynamic. A dynamic model according to this definition would be one where the variables must be dated and a time lag must exist in their relationships According to this criterion the following would be a dynamic model.

$$D_t = f(P_t)$$

$$S_t = g(P_{t-1})$$

$$D_t = S_t$$

There is **no lag** in the demand relationship. Demand in period 't' depends on own price of the same period. However, in the supply relationship a gestation lag exists which makes the model dynamic. Supply in period 't' depends on price prevailing in the previous period (t-1). The price level in previous period (t-1) would have induced the producers to increase or decrease the supply, full impact of such decisions are visible in time period 't' only. For market to attain equilibrium, demand in period 't' must equal supply in period 't'.

It must be noted that if one is concerned with the equilibrium configurations of a market for a good, one has to take recourse to a static methodology. *Equilibrium is a static concept*. It describes the position of a market at rest. In constrast, disequilibrium analysis must pertain to dynamics. It brings into focus the market adjustment process (or, market corrective process), the interplay of which would move the market back to equilibrium. One has to analyse how the market moves through time during the period the adjustment process is working. In a static framework, we implicitly assume that market adjustment is instantaneous, and without any loss of time, equilibrium is or is not restored. How the economic agent behaves in the disequilibrium situation is not the concern of static analysis. This is where dynamic analysis sets in. It must be noted that in a static framework one might be interested in comparing (or evaluating) two or more equilibrium positions before and after a change in some exogenous forces. Such a method is known as **comparative static**. For instance, consider analysing the

effect on price of cars when demand increases. We concentrate on two equilibrium positions, one before change and another after the change in demand has taken place. What happens in the interim period is not the concern of static analysis.

Check Your Progress 2

1) Consider the following demand/supply functions

A) i) $Q^D_X = a - bP_x$;

ii) $Q^S_t = f(P_t)$;

B) i) $Q^D_X = A-BP_x$

ii) $Q^S_t = f(P_{t-1})$.

Are these models Static or Dynamic ?

3.4 CONSTRUCTION AND VERIFICATION OF ECONOMIC THEORIES

Construction of economic theory consists of the following stages:

1) We choose a hypothesis or a proposition about the phenomenon or events we are interested in investigating. This is derived from the empirical observation of the phenomenon under study. For instance, we choose as a hypothesis the inverse relationship between the quantity demanded for tea and the market price of tea. This hypothesis is thrown up by market observations (or data).

2) At the next step, we formulate an economic model (which is our theory) about the causal relationship among the relevant variables embedded in the hypothesis. The method *of deductive reasoning* is used to develop such a cause-effect relationship. It deals with an '*if-then'* kind of argument in terms of why such as relationship exists. To develop such a model or theory a process of abstraction has to be applied. Since reality is infinitely complex, to make any headway in explaining a phenomenon simplification process has to be used. In other words, the inessentials are discarded while the essentials are incorporated into the model. Such type of simplification is required in order to keep the model manageable. One such model is the cardinal utility theory of demand.

3) Next, the model is applied to the hypothesis or the proposition to derive the implications or the conclusion as regards the phenomenon under study. For instance, when we use *the cardinal utility theory of demand,* we find that when a consumer consumes larger quantity of a good, the marginal utility tends to fall. Hence, unless the price is lowered, larger quantity will not be consumed (since the theory would postulate that price reflects marginal utility). Thus, the conclusion that when price is lowered a larger quantity is demanded (which is our hypothesis to start with).

4) The last step relates to conclusion derived from the model is put to empirical testing. In other words, the conclusion regarding the phenomenon or the object under study is set against or confronted with the observations on the phenomenon or object as it is found in reality. In order to test empirically the conclusion of a theory we have to take recourse to statistical or econometric methods to scan the empirical data for relationship it is trying to establish. If such a testing confirms the relationships established by our model, then we accept the theory as providing a logically valid explanation for the phenomenon or objects as observed in reality. However, if empirical testing contradicts our conclusions about the object under study then either we discard the theory altogether or modify the hypothesis (going

back to the first step and starting the enquiry process once again) as well as the model. This process of going back and forth from hypothesis to empirical observation and testing is repeated till it is possible to find a hypothesis which agrees with observations on the phenomenon as they are found in reality.

3.5 ECONOMIC THEORY AND ECONOMIC LAWS

A model is one, which is used to describe a set of causal relations among economic variables or economic objects. *A hypothesis, which is successfully tested, is called a theory. The purpose of a theory is to explain and predict.*

An economic theory, which is true under similar set of circumstances, is called a law. For instance, the Law of Demand.

Check Your Progress 3

Consider the statements:

i) Consumption depends on income

ii) Population rises when per capita income goes up .

Which of the above statement is a theory or a law?

3.6 STOCK VARIABLE AND FLOW VARIABLE

In economics we deal with two types of variables, *flow variables* and *stock variables*. Anything, which varies, is a variable, (For instance price, quantity demanded and supplied, income, investment, exports, imports, employment, cost of production, profits etc.). However, for variation to be observed we will have to specify the period of time. Whether it is a week, a month, or a year or a longer period of time. Now in each of these periods the relevant variable may be a stock variable or a flow variable. *Both stocks and flows are expressed at a precise moment in time. However, a flow variable has both a time dimension and a time reference,* while a stock variable has only a time reference. Though both are measured at fixed points in time, flow variables are measured in "temporally determined units". *In other words, flows are always expressed per unit of time. While stocks are always expressed at a point in time.* For instance, capital is a stock variable, since it has no time dimension but has only time reference, like stock of capital on 1st January 2000. Investment, however, is a flow variable since it is expressed per unit of time, like 10 per cent per annum. *Wealth is a stock magnitude while income is a flow magnitude.* Stocks and flows are related to each other in the sense that differences in stocks in the two periods (say 1-1-2000 and 31-12-2000) will constitute flows during the period (the time interval 1-1-2000 to 31-12-2000). It is through flows that stocks are added (to, or depleted). For instance, capital stock on 1-1-2000 plus the net flows of investment (net of depreciation)during the period 1-1-2000 to 31-12-2000 will make the stock of capital on 31-12-2000. The importance of stocks and flows is that there exists markets for both stocks and flows. In micro-economic theory we will be concerned more with flows than with stocks.

Check Your Progress 4

Identify the 'stock' and 'flow' among the following:

i) Inflation rate, rate of interest, money supply, population

ii) Demand for wheat and supply of wheat

3.7 LET US SUM UP

To recapitulate, in this unit we concentrated once again on methodological issues in economics. Essentially, microeconomic theory is based on partial equilibrium approach while macroeconomic theory is based on general equilibrium approach. The distinction between the partial and general equilibrium analysis has been spelt out in great details. Partial equilibrium analysis is based on *ceteris peribus* assumptions, while general equilibrium is characterised by *"everything depends on everything else"*. You must be often hearing the terms static and dynamic. We focussed on this distinction as well in this unit. In a very pedestrian manner we found that '*static*' deals with timeless situation while '*dynamic*' deals with changes over time. Equilibrium analysis is supposed to be static in nature while disequilibrium analysis pertains to dynamics. A description or the characterisation of equilibrium constitutes static framework.

We have concluded this unit with a discussion on how to construct and verify an economic theory. We have already tackled this question in the very first unit itself while expanding upon Lionel Robbins' definition of economics. One uses the scientific method of deductive logic to construct economic theories. The conclusions derived from an economic theory must be tested empirically using the scientific method of econometrics. One cannot, however, be 100 per cent sure as to the success of such tests.

3.8 SOME KEY WORDS

Dynamics : When variables are dated. When changes take place over time.

Economic Laws : An economic theory, which is true under similar set of circumstances, is called a law.

Economic Theory : A model, which is used to describe a set of causal relations among economic variables or economic objects. A hypothesis, which is successfully tested, is called a theory. The purpose of a theory is to explain and predict.

Flow : A variable expressed per unit of time. It has both time dimension and time reference.

General Equilibrium : An equilibrium analysis where we analyse simultaneously all markets in the economy. All markets are considered to be inter-dependent and inter-related. It represents a more complex analytical framework than partial equilibrium analysis.

Partial Equilibrium : An equilibrium analysis pertaining to a particular market with everything else ignored. In particular, the interactions between various markets of the economy are not taken into account.

Static : When variables are not dated. Refers to 'timeless' situations.

Stock : Variable measured at a point of time. It has only a time reference.

3.9 SOME USEFUL BOOKS

Same as in Unit No. 1.

3.10 ANSWERS OR HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

i) Partial equilibrium

ii) General equilibrium analysis

Check Your Progress 2

1) a) Static

b) Dynamic

Check Your Progress 3

i) Statement of a theory

ii) More of a law

Check Your Progress 4

i) flow, flow, stock, stock.

ii) flow as well as stock.

Indira Gandhi
National Open University
School of Social Sciences

EEC-11
Fundamentals of Economics

Block

1

Introduction to Economics

Expert Committee

Mr. Kalyanjit Roy Chaudhary
St. Stephens College
University of Delhi
Delhi

Mr. B.S. Bagla
PGDAV College
University of Delhi
Delhi

Mr. R.S. Malhan
St. Stephens College
University of Delhi
Delhi

Mr. R.S. Bharadwaj
Shivaji Collage
University of Delhi
Delhi

Dr. Gopinath Pradhan
IGNOU

Dr. Narayan Prasad
IGNOU

Dr. Madhu Bala
IGNOU

Dr. Kaustuva Barik
IGNOU

Mr. Saugato Sen
IGNOU

Prof. S.K. Singh (retd.)
IGNOU

Course Editor	Course Coordinator	Block Coordinator
Mr Bhawani Sankar Bagla	Dr. Gopinath Pradhan	Dr. Gopinath Pradhan

Block Preparation Team

Unit No.	Unit Writer	Unit Editor (IGNOU Faculty)
1	Mr. Kalyanjit RoyChaudhary	Dr. G. Pradhan & Dr. N. Prasad
2	Mr. Kalyanjit RoyChaudhary	- do -
3	Mr. Kalyanjit RoyChaudhary	- do -

Production	Cover Design	Word Processing
Mr. Arvind Kumar	Ms. Arvinder Chawla	Mrs. Rekha Mishra
Mr. Manjit Singh		Ms. Daisy Lal

March, 2003

ISBN-81-266-

Further information on the Indira Gandhi National Open University courses may be obtained from the university's office at Maidan Garhi, New Delhi-110 068.

Printed and published on behalf of the Indira Gandhi National Open University, New Delhi, by Professor Kapil Kumar, Director, School of Social Sciences.

Laser typeset by : HD Computer Craft, EA1/75, Main Market Inderpuri, New Delhi-110012. Ph: 25831437; 3108 1117

Printed at :

BLOCK 1 INTRODUCTION TO ECONOMICS

This block introduces you to the subject matter of economics and underlying need for studying the economic problems. It has three units:

Unit1 deals with the nature, significance and definition of economics. **Unit 2** discusses the basic concepts used in economic analysis while **Unit 3** introduces the learners to methods available for analysing the economic problems. Employment of the techniques of partial and general equilibrium framework for quantification of the impact of independent variable(s) on a (or, set of) dependent variable(s) is the core methodology you will be introduced to.

INTRODUCTION TO ELECTIVE COURSE IN ECONOMICS EEC-11

Economics is no longer the preserve of those who practiced statecraft and regarded it to be confined to the ways and means of raising finances to meet the "requirements" of the ruling elite. The discipline has moved from such confines to the domain of the common man. It is now concerned with our day-to-day decisions such as: Which commodities to produce? How to produce? Which techniques to use? Which factors or resources to use in which combinations to produce what quantity of a commodity? Not only this, it shows which consumer may gain access to what specific amounts of different goods ? How to increase/decrease production of which good(s) in future? In other words, economics has moved away from financing the activities of state to helping the common man in the street to make many a crucial decisions impinging on their day-to-day life.

It must be remembered, however, that we have not moved from one extreme to another— from the state to the street. We, today incorporate a rather wide spectrum of activities in the domain of economics. These activities are: (a) consumers' behaviour or choice process; (b) producers' behaviour or how is the production organised and carried on, what is the special role of cost functions therein and also the different forms of market organisations; (c) different individuals co-operate in the process of production to contribute factors owned by them. How do we determine their 'rewards'? Or, how do we distribute aggregate output among the members of society? (d) estimation of national (social) product and various aggregates, determination of level of income, employment and interest and also the relationship between money supply and prices; (e) some aspects of international trade; (f) public finance which not only incorporates all the aspects of meeting financial requirements of the state but also focuses on 'newer' aspects of collective decision making.

The present course, Fundamentals of Economics (EEC-11), aims at exposing the learner to each of the above aspects. The course is divided into 9 blocks, spanning over 21 units. Block 1 is concerned with introducing the subject matter of economics along with nature of basic economic concepts and the methodology of this discipline. Block 2 analyses the behaviour of the consumer while Block 3 is concerned with technical specifications of production and cost functions. Block 4 uses information and knowledge gained in previous two blocks and analyses behaviour of the producers under different forms of market organisation. The theories of factor pricing, that is, determination of wages, rent, interest and profits in the society is our concern in Block 5. These five blocks constitute core of micro economic analysis.

Next three Blocks deal with what is popularly known as macro-economic analysis. Block 6 explains the idea of circular flows of money (and goods and services) in the society, and measurement of national income. In Block 7, we present various aspects of determination of income, employment and interest in the society. This block is essentially based on J.M. Keynes' contributions- though, at relevant points, we have also compared Keynesian ideas with 'classical' thinking about aggregative functioning of the society. In Block 8 , we are introducing relationship between quantity of money and price level on the one hand and those between rate of change of prices and levels of unemployment on the other. In this context we discuss Classical, Keynesian and Modern versions of quantity theory of money and Philips curve.

Finally, Block 9 introduces you to the basic aspects of public finance, public goods, externalities and market failure, public revenue and expenditure and various concepts of deficit in the government budget. The other unit in this block examines comparative cost theory of international trade, gains from trade, terms of trade and the structure of balance of payments accounts.

UNIT 4 DEMAND FUNCTION AND ELASTICITY

Structure

4.0 OBJECTIVES

After going through this unit you will be able to:

- analyse and explain the factors influencing the demand for a good;
- specify the consumer's demand function for a good;
- explain the difference between movement along a demand curve (as caused by variation in the own price of the good) and movement of demand curve itself (as caused by variation in other factors, which influence demand for good);
- explain the concept of elasticity of demand and how it is represented graphically, intuitively and algebraically;
- distinguish between point method and arc method of calculating elasticity;
- understand the concept of cross elasticity of demand;
- understand the concept of income elasticity of demand;
- explain the factors on which elasticity of demand depends; and
- understand the importance of the concept of elasticity of demand to a business firm or a policy maker.

4.1 INTRODUCTION

In this unit we will be focussing on the consumer's demand function for a good. The objective ultimately is to relate the individual demand functions for a good to arrive at the corresponding market demand function. Remember that we will be using a demand-supply framework to analyse the working of a market. We will start with the specification of demand function and various influences that work on consumer's demand for a commodity which could be, *own price of the good, the price of a related good, consumer's income and host of other factors*. In particular, we will be concerned with the relationship between the quantity demanded for a good and the

own price of the good. This relationship based on a *ceteris paribus* clause will generate the consumer's demand function for the good. Why a consumer's demand function is downward sloping will be answered in the next unit. In this unit, however, we will be mainly focussing on the concept of elasticity of demand, which is one of the most fundamental properties of a demand function. We will introduce you to various concepts of elasticity, in particular price elasticity of demand and income elasticity of demand. We will be concerned with the meaning, measure, usefulness and significance of various elasticity concepts. In the process we will also highlight some of the major influences affecting elasticity of demand.

4.2 EFFECTIVE DEMAND

We have seen the meaning of *effective demand* in Unit-2. Moreover, we have also understood that the demand means the individual's desire for the good backed by capacity to pay for it. In other words, an individual's desire for a good to satisfy a particular want backed up by her willingness and ability to pay gives rise to demand for a good. If and only if individuals have means to pay, demand becomes effective. An individual's income measures her capacity to pay, purchasing power or means to pay for the goods desired.

Example: Suppose a beggar without money desires milk, but has no purchasing power. Hence the beggar's desire for milk does not constitute an effective demand for milk. As a result the beggar cannot participate in market activities. However, suppose this beggar becomes successful in getting a job, becomes a helper in a shop and for his work as helper gets paid in for money. The beggar who is now a helper earns an income, with which she can buy milk (depending on the price of milk). The beggar's demand for milk, which earlier constituted only an absolute demand, has now become an effective demand. This particular beggar's demand for milk now adds to the market demand for milk, which a potential supplier of milk has to take into account in deciding how much milk must be produced and supplied. Hence, for demand (for a good like milk) to exist two conditions must be fulfilled:

i) individuals must have a desire for that good, and

ii) their desire must be supported by income or purchasing power or means to pay, given of course the individuals willingness to pay or spend

Now, let us look into the factors that affect demand for a good. The demand for a good in a market depends on many factors. Some of the important factors are:

a) **Price of the good under consideration:** For instance, how much quantity of milk one will buy would depend on the price of milk. In other words, the own price of a good is the single most important influence on its demand. The normal behaviour that one would observe is an inverse relationship between the quantity demanded of a good and its own price. As the price of a good rises the quantity demanded falls. And as the price falls the quantity demanded would tend to increase. To appreciate the price-quantity relationship consider the following example: When the price of milk is Rs.20 per litre an individual may buy only half a litre, while if the price is Rs.10 per litre this individual may buy one litre or more. This is a commonly observed aspect of households' behaviour in the market for a good.

b) **Substitutes and Complements:** The second important consideration on which the demand for a good would depend is the prices of related goods. The related good could be a *substitute good* or a *complementary good*. In case of fresh milk, the substitute good could be either powdered milk or

condensed milk. On the other hand, a complementary good could be either corn flakes or oats. The quantity demanded of a good would vary positively with the price of the substitute good, while it will vary negatively with the price of the complementary good. For instance, when the price of powdered milk rises, consumers will move away from powdered milk towards fresh milk. This will increase the demand for fresh milk. On the other hand when the price of corn flakes goes up, the demand for corn flakes will go down and as a result the demand for fresh milk will also go down.

c) **Income level:** Another influence on the demand for a good is the individual consumer's income level. Normally as an individual becomes richer, she would tend to increase the consumption of each and every good, and in particular the consumption of the good being considered. This implies that as our individual consumer's income rises, the demand for milk (the good under consideration) would tend to go up. Hence, we will observe a positive correlation between demand for a good and the household's income.

At this point it is worthwhile to note that if the increase in income results in a decrease in the demand (consumption) for a good, the good will be classified as an *inferior good*. This will, of course, vary from individual to individual. If a consumer consumes less of milk as he/she becomes richer, for such individual milk is an inferior good.

d) **Consumer's tastes and preferences:** The demand for a good depends upon the individual consumer's tastes or preferences. A consumer will demand or desire a good if and only if she has a taste for the good (or, has a definite preference for that good).

4.3 DEMAND SCHEDULE, DEMAND FUNCTION AND DEMAND CURVE

In Section 4.1 we have considered various influences on the demand for a good. Here, it must be pointed out that there could be many other influences working on the demand that we have not considered. However, the above four factors are the major influences on the demand for a particular good. In any particular situation if we keep factors other than own price as constant, we can then derive a *demand schedule*, a *demand function*, and a *demand curve*. A demand schedule lists the various quantities of a good that a potential consumer buys from the market at different prices of the good, observed at a given moment of time. Its tabular representation would be as follows:

Table- 4.1: Demand Schedule for Milk

Price of Milk: per litre (Rs.)	20	15	10	5
Quantity Demanded (Litre)	1	1.5	3	6

The demand function for a good expresses a functional relationship between quantity demanded of the good and its own price. If the good is X (milk in our case), Q_x is quantity demanded and P_x is the own price of good X then *the general form of the demand function will be*

$$Q_x = f(P_x).$$

What it says is that quantity demanded depends on price. Price is the cause variable and quantity demanded is the effect variable. Stated alternatively, *price is the*

independent variable while quantity demanded is the dependent variable. In technical terms, independent variables are also called *exogenous variables,* while dependent variables are called *endogenous variables*. The demand function which expresses the functional relationship between the quantity demanded of a good and its own price is based on *ceteris paribus* assumptions, that is, we only allow the own price to vary with everything else held constant at their pre-assigned values. In other words, when we try to capture the relationship between quantity demanded of a good and its own price, we ignore all other influences on the demand for that good (like prices of substitute goods, complementary goods, the household's income level, tastes and so on). The demand curve is a graphical representation of the demand schedule. The graph of the demand function is plotted on a two dimensional Euclidean space with horizontal axis (abscissa) measuring the quantity demanded of the good and the vertical axis (ordinate) measuring the price. Since the normal behaviour is one of an inverse relationship between quantity demanded of a good and its own price, *the demand curve for a good will be downward sloping*. For the sake of convenience, we can assume the demand curve to be a downward sloping straight line. This is illustrated in Fig. 4.1.

The quantity demanded of good X (milk) is plotted on the *horizontal axis* (x-axis) and the price of good X on the *vertical axis* (y-axis). The quantity demanded is measured in physical units of the good. In case of milk physical unit is litres. Price of the good is, however, expressed in monetary units (in rupees or paise). In case of milk, its price would read as Rs. 8 per litre or Rs. 12 per litre etc. A downward sloping demand curve for good X would imply that when price is lowered, the quantity demanded would increase. And when the price is raised, the quantity demanded would decrease. In Fig. 4.1, when price is OP1 the consumers of good X operate at point N_1, demanding (purchasing or buying) OQ1 quantity of good X. When price falls to OP_2, the consumers operate at point N_2 on the demand curve d, demanding OQ_2 quantity of X. A lowering of price induces a larger quantity of it being demanded. This is supposed to be a commonly observed aspect of consumer behaviour. As to the question why demand curve is downward sloping, we defer the explanation to Unit-5. A downward sloping demand curve reflects the law of demand. This law says that, other things remaining the same, a consumer (or in general, consumers) would buy more when price falls and buy less when price rises.

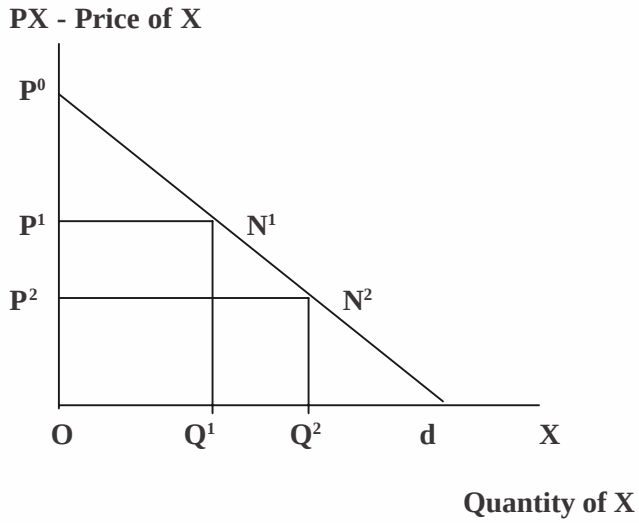


Fig. 4.1 Simple demand relationship or, 'law of demand' is shown with help of this figure. At OP1 price the consumer buys OQ1 of the commodity X. As price falls to OP2 the quantity demanded by the consumer rises to OQ2. This inverse relationship, that is, as price falls, quantity demanded raises, is called law of demand.

Here it can be pointed out that when the demand curve for a good is a straight line, the corresponding demand function will have a linear equation of the type:

$$Q_x = a - b P_x$$

Where 'a' is the intercept and 'b' is the slope. The coefficient 'b' expresses the rate at which quantity demanded changes (increases or decreases) when price changes (falls or rises). That is:

$$b = \partial Q / \partial P$$ (subscript 'x' is dropped for convenience)

Generally, it is seen that the independent variable is plotted on the x-axis and dependent variable on the y-axis. But in case of the demand curve it is the other way round. In fact, in economics, we, as a rule, show price and other price type variables, such as costs, interest rate etc., on vertical axis while the quantity type variables are shown on the horizontal axis. Truly speaking, we plot the inverse demand curve

$$P_x = \alpha - \beta Q_x$$

where $\alpha = a/b$ is the price intercept and $\beta = 1/b$ is the slope of the inverse demand curve and equals $\partial P/\partial Q$. In its normal form $Q_x = a - b P_x$, the demand curve would indicate the *maximum quantity demanded* at any given price of the good. In its inverse from $P_x = \alpha - \beta Q_x$, the demand-curve would indicate for each given quantity demanded the maximum price a consumer (or consumers) would be willing to pay rather than doing without that quantity. Hence, the demand curve (or its inverse form) always indicates the maximum boundary to consumers. No consumer will be willing to pay for OQ_1 quantity a price higher than OP_1 ($=N_1Q_1$) in Fig. 4.1. Also since price paid by buyers reflects the average revenue (AR) earned by sellers, the inverse demand curve is also known by the name *average revenue curve* (this concept will be discussed in Unit-8). Also note that the point where demand curve touches the price axis, gives us the price at which quantity demanded falls to zero. Such a price is called the prohibitive price (price OP_0 in Fig. 4.1).

Shifts in the Demand Curve

We have already seen that demand is influenced by price of the good. But there are several other factors that influence the quantity demanded. For example, when the income of a consumer increases without any change in prices of goods and services, it is generally seen that she will consume more. Similarly, there can be a change in the tastes and preferences of a consumer. For example, I was taking non-vegetarian foods earlier. Now the doctor advises me not to eat non-vegetarian foods on health ground. Suddenly there is a change in my demand schedule for mutton. Even if the price falls, my demand for mutton is zero.

Note that in the above cases the demand curve changes throughout, that is, in all ranges of price. Such changes are expressed by *shifts in the demand curve*. The upward shift is also called *increase in demand* while the downward shift is called *decrease in demand.* In Fig. 4.2 we have indicated two types of shifts in the original demand curve dd. The upward shift is indicated by d_1d_1 and downward shift by d_2d_2.

It may be noted that when the price of a commodity changes the consumer moves on the same demand curve, whereas changes in other factors result in shift in the demand curve.

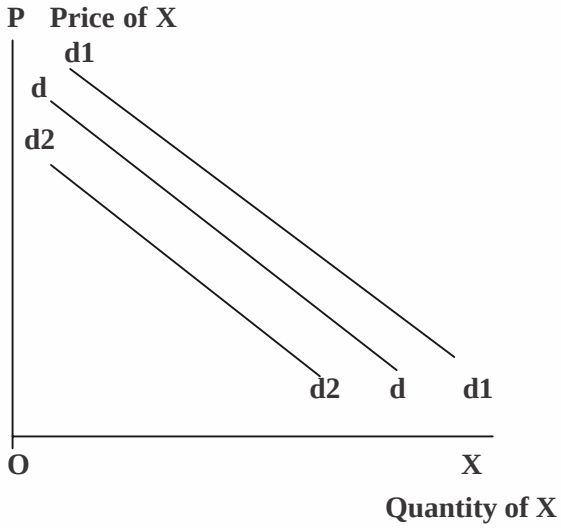


Fig. 4.2 What happens, when 'other things' like income or preferences of the consumer change? A rise in income etc., can prompt consumer to buy a larger quantity at two prevailing prices. Entire demand curve shifts to from dd to d_1d_1. If income falls, the curve shifts to the left, from dd to d_2d_2.

Check Your Progress 1

1. When a consumer's income increases, what happens to the demand curve?

   .....

   .....

   .....

   .....

2. For the following schedule, what happens when price falls from Rs.9 to Rs.8 and when price rises from Rs.5 to Rs.6?

Price (in Rs.)	Quantity demanded of a good (in kg.)
11	100
10	150
9	180
8	205
7	215
6	225
5	230
4	235

   .....

   .....

   .....

   .....

4.4 ELASTICITY OF DEMAND

We have seen that the demand function of a commodity shows the relationship between the quantity demanded of a commodity and its own price, with everything else held

constant. It is of the form $X_1 = f(P_1)$, ceteris *paribus*. Such a relationship indicates the maximum quantity demanded at any given price of the commodity. In its inverse form $P_1 = f(X_1)$ it would indicate the maximum price a consumer would pay for each quantity demanded. This can be stated as the maximum willingness price on the part of a consumer rather than doing without that quantity of the commodity. The slope of such a demand function would indicate the rate of change of the quantity demanded with respect to a change in its own price i.e., $\partial X_1/\partial P_1$ expressed in the units in which the commodity is measured. To get rid of the units of measurement (like kg. litre, ton, bales, etc.), we use the concept of *elasticity.* It captures the extent to which the demand for a commodity would respond when the price of the commodity changes. Formally, the concept of elasticity of demand is defined as the degree of responsiveness of the quantity demanded of a commodity with respect to a change in the variable on which the demand for a commodity depends, like own price, price of substitute or complementary commodity and income. Accordingly, we have (i) own price elasticity of demand (ii) cross price elasticity of demand, and (iii) income elasticity of demand for a commodity. Here, it can be pointed out that the concept of elasticity is borrowed from physics, wherein it refers to the intensity with which a dependent variable is affected when a cause variable changes.

Price elasticity of demand: Take the demand function for a good $Q_1 = f(P_1)$. Price elasticity of demand is defined as the degree of responsiveness of the quantity demanded of good Q_1 when the price of the good, P_1, changes. It measures the extent to which demand for a good would increase or decrease as the own price falls or rises. Thus, elasticity indicates how demand responds when the price changes. Algebraically, elasticity is expressed as a ratio of two terms, the relative change in demand for Q_1 and the relative change in the price P_1. Thus,

$$\varepsilon_{11} = (\Delta Q_1 / Q_1) / (\Delta P_1 / P_1)$$

where ε_{11} = elasticity of good one with respect to its own price;

ΔQ_1 = change in demand for X_1;

Q_1 = original quantity demanded for X_1;

ΔP_1 = change in own price of good one;

P_1 = original price of good one.

Hence, ε_{11} = relative change in the quantity demanded of X_1 over relative change in the price of good X_1 (i.e., P_1). Alternatively, it could be interpreted as a ratio of the proportionate change in the quantity demanded of good X_1 to the proportionate change in P_1. Note that if we multiply the numerator and the denominator by one hundred we get ε_{11} as a ratio of the percentage change in the quantity demanded of X_1 to the percentage change in the price of good one.

Normally, the sign of ε_{11} is negative, since the law of demand implies that quantity demanded and price would be inversely related. In other words, with $\Delta P_1 > 0 \Rightarrow \Delta Q_1 < 0$ and $\Delta P_1 < 0 \Rightarrow \Delta Q_1 > 0$, we find that $\Delta P_1/ P_1$ and $\Delta Q_1/ Q_1$ move in opposite directions. Though price elasticity of demand mathematically has a negative sign, it has to be qualitatively interpreted without the sign. That is, we take $|\varepsilon_{11}|$ (read *modulus* of ε_{11}).

Interpretation of e_{11}

(i) Suppose $\varepsilon_{11} > -1$

or, $|\varepsilon_{11}| > 1$

This would imply that one percent decrease (increase) would cause demand to increase (decrease) by more than one percent. In other words, demand is responsive to price change. The higher the value of ε_{11} above unity, the more significant is the demand response. Whenever ε11 exceeds unity (sign ignored) demand is said to be elastic or demand elasticity is said to be greater than one. We find $\Delta Q_1 / Q_1$ exceeding $\Delta P_1 / P_1$ (in the opposite direction).

(ii) Suppose $\varepsilon_{11} < -1$

or $|\varepsilon_{11}| < 1$.

This would imply that a one percent decrease (increase) in price of the good would cause demand to increase (decrease) by less than one percent. In other words, demand is not very responsive to price change. Even though demand would increase (decrease) as price falls (rises), the (increase or decrease) is not very significant. Hence, whenever ε11 is less than unity (sign ignored) demand is said to be inelastic.

(iii) Suppose $\varepsilon_{11} = -1$

or $|\varepsilon_{11}| = 1$.

In this case a one percent increase (decrease) in price would cause demand to increase (decrease) by exactly one percent.

Demand in such a situation is neither elastic nor less elastic, and is said to be *unitary elastic*. In other words, $\Delta X_1 / X_1 = \Delta P_1 / P_1$ in the opposite direction. If price rises by 10 percent, demand would fall by 10 percent. The demand response is 'Middle-of-the Road, type and provides the 'Great Divide' between more elastic and less elastic demand.

Check Your Progress 2

1. When price is Rs.10 the demand for the good is 100 unit, and the price elasticity of demand is 1.5, what will happen to demand when price fall by ten percentage points?

   .....

   .....

   .....

   .....

2. For the following table determine price elasticity when price rises from Rs. 7 to Rs. 8.

Price (in Rs.)	Quantity demanded (kg.)
10	50
9	60
8	65
7	68
6	69
5	70

4.5 DISTINCTION BETWEEN ARC ELASTICITY AND POINT ELASTICITY

Consider the demand curve dd in Fig. 4.3. We want to capture the demand response over the arc N_1N_2 on the non-linear demand curve dd allowing the price to vary between OP_2 and OP_1. The demand response must be the same whether the consumer moves from N_1 to N_2 or from N_2 to N_1. However, when we use the definition of price elasticity we will get one type of answer (response) if the initial price is OP_1 and another distinct answer (response) when the initial price is OP_2. When price falls from OP_1 to OP_2

$$\varepsilon_{11} = (Q_1Q_2 / OQ_1) / (P_1P_2 / OP_1) \quad \text{or,}$$

$$\varepsilon_{11} = (Q_1Q_2 / P_1P_2) / (OP_1 / OQ_1) \quad ...(i)$$

On the other hand, when price rises from OP_2 to OP_1

$$\varepsilon_{11} = (Q_1Q_2 / OQ_2) / (P_1P_2 / OP_2) \quad \text{or,}$$

$$\varepsilon_{11} = (Q_1Q_2 / P_1P_2) / (OP_2 / OQ_2) \quad \ldots(ii)$$

If we compare (i) and (ii) we find that the value of ε_{11} differs since OP_1/OQ_1 is not equal to OP_2/OQ_2 even though Q_1Q_2/P_1P_2 is common to both. To overcome the above problem created by the choice of the initial reference point (whether N_1 or N_2) we use the concept of *arc elasticity*.

If we use the method of arc elasticity, then price elasticity of demand for good X_1 is defined as

$$\varepsilon_{11}^{A} = \{(Q_1^1 - Q_1^0) / (Q_1^0 + Q_1^1)\} / \{(P_1^1 - P_1^0) / (P_1^1 + P_1^0)\}$$

$$= \{\Delta Q / (Q_1^1 + Q_1^0)\} / \{\Delta P / (P_1^1 + P10)\}$$

where ΔQ_1 = change in the quantity demanded of Q1

ΔP_1 = change in the own price of Q_1

Q_1^0 = the quantity demanded at the price OP_1 (i.e., P_1^0)

Q_1^1 = the quantity demanded at the price OP_2 (i.e., P_1^1)

P_1^0 = the price OP_1 in Fig.4.3

P_1^1 = the price OP_2 in Fig.4.3

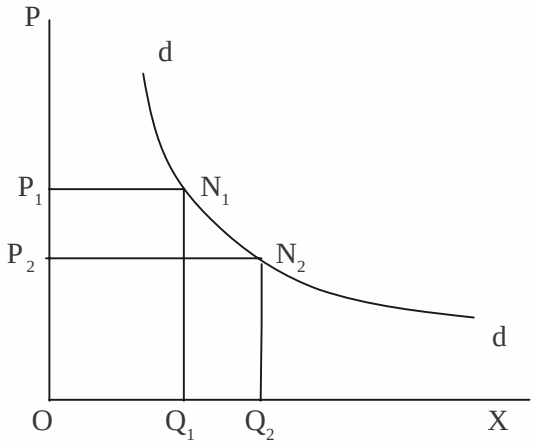


Fig. 4.3 A curvy-linear demand curve is shown. As price changes from OP_1 to OP_2 quantity demanded changes from OQ_1 to OQ_2 Thus, change in price, $\Delta P = P_1P_2$ and change in quantity, $\Delta Q = Q_1Q_2$.

If arc elasticity is used to measure the demand response over the arc N_1N_2 of the demand curve **dd**, then we will get the same value for price elasticity whether we move up or move down the range N_1N_2 of the curve **dd**.

In the arc elasticity formula if we let ΔP_1 approach zero then we get the formula for point elasticity. In other words, the limit of arc elasticity as ΔP_1 tends to zero is point elasticity. That is,

$$\lim_{\Delta P_1 \to 0} \varepsilon_{11} = (\partial Q_1/\partial P_1)(P_1/Q_1) = \varepsilon_{11}^P$$

Conceptually, point elasticity measures demand response for an infinitesimal change in price, while arc elasticity measures demand response for a finite (discrete) change in price. Those of you who are familiar with logarithmic differentiation may note that:

$$\partial Q_1/Q_1 = \partial \log Q_1 \text{ and } \partial P_1/P_1 = \partial \log P_1$$

$$\text{Thus, } \varepsilon_{11} = \partial \log Q_1 / \partial \log P_1$$

Examples:

1. *Consider the following demand schedule where Q_1 is quantity demanded of a good and P_1 is price of the good. The quantity is in kilogram, while the price is in Rupees.*

Q_1	100	150	220	250	280
P_1	5	4	3	2	1

What is the price elasticity when price falls from Rs.4 to Rs.3?

Here, since the price has fallen by one rupee, which is a finite change, we use the concept of arc elasticity. Hence,

$$\varepsilon_{11}^A = -\{\Delta Q_1 / (Q_1^0 + Q_1^1)\} / \{\Delta P_1 / (P_1^1 + P_1^0)\} = \{70/(150+220)\}/\{-1/(3+4)\}$$
$$= -490/370 = -1.3$$

2. *For the demand function $Q_1 = 50 - 0.8\ P_1$ find price elasticity of demand at $P_1 = 10$.*

Here, since we have to find elasticity at a point ($P_1 = 10$) we use point method.

$$\varepsilon_{11}^P = (\partial Q_1 / \partial P_1)\ (P_1 / Q_1)$$

For the demand function $Q_1 = 50 - 0.8\ P_1$, $\partial Q_1 / \partial P_1 = -0.8$

At P=10, we get, $Q = 50 - 0.8\ (10) = 42$.

By substituting the above values we get $\varepsilon_{11}^P = -0.8\ (10 / 42) = -0.2$

As mentioned earlier, value of elasticity is interpreted without considering the minus sign. It is the absolute value of ε_{11} that matters. If demand is elastic, $\varepsilon_{11} > 1$ and if demand is less elastic, $\varepsilon_{11} < 1$. Thus in Example 1 above demand is elastic while in Example 2 it is less elastic.

Note that elasticity measure, unlike the slope of a demand curve is a *pure number*,

i.e., unitless, free from the units of measurement of Q_1 as well as P_1. One advantage of such a measure is that it can be used (for a comparative evaluation) across all goods and commodities expressed in different physical units. For instance, we can say that 1.8 is greater than 1.7 while the same cannot be said about 1.8 kg. and 1.7 litres.

Check Your Progress 3

(i) In example 2 above find out the price elasticity when price falls Rs.9 to Rs.8, using

(a) arc method

(b) point method

4.6 GRAPHICAL REPRESENTATION OF ELASTICITY

Consider the linear inverse demand curve AB in Fig. 4.4. Find out the elasticity at a point N on this demand curve.

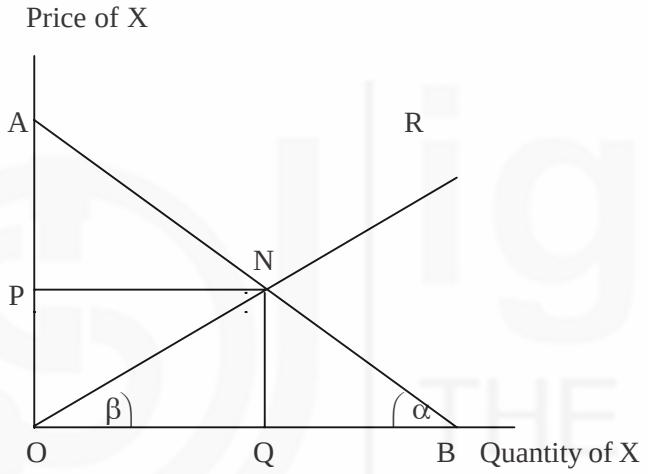


Fig. 4.4 AB is demand curve. It makes an angle a with quantity axis. ON joins the point N with origin making angle b with the quantity axis. tanb = NQ/OQ = P_1/Q_1. Also, tan α = NQ/QB = ΔQ/ΔP if price falls to zero. Thus, elasticity at N = ε_N = tanβ/tanα .

By point method elasticity is defined as

$$\varepsilon_{11}{}^{P} = (\partial Q_1 / \partial P_1)\ (P_1 / Q_1) = (P_1 / Q_1) / (\partial P_1 / \partial Q_1)$$

$$= \text{Average Function/ Marginal Function}$$

In Fig. 4.4 at point N of the demand curve AB, P_1/Q_1 is given by tanβ while $\partial P_1/\partial Q_1$ is given by tanα. Hence elasticity at N = ε_N = tan β / tan α. When tan β is positive and tan α is negative, elasticity is negative. We can vary point N on AB to see how elasticity changes from point to point. To understand this point, look at the following cases:

Case I: N is the midpoint of AB. This would imply β = α or tan β = tan α and hence $\varepsilon_N = 1$.

Case II: N lying above the midpoint of AB, implying tan β > tan α so that $\varepsilon_N > 1$.

Case III: Point N lying below the midpoint of AB. This would imply tan β < tan α and hence $\varepsilon_N < 1$.

Case IV: N coincides with the corner point A implying $\beta = 90^o$. In this case tan β = infinity. Hence ε_N = (infinity).

Case V: N coincides with the corner point B implying tan β =0, hence ε_N = 0 (zero).

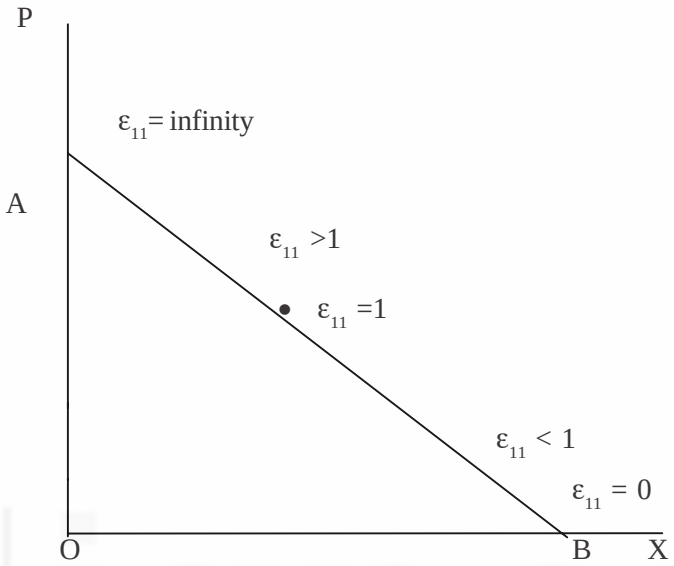


Fig. 4.5 AB is demand curve. Elasticity at any point is equal to lower segment divided by upper segment. Thus, it is equal to unity at the mid-point, as we go down, the elasticity declines. So, higher the price, higher the elasticity and lower the price, lower the elasticity.

From the above we learn that on a linear inverse demand curve the price elasticity falls in value as we move down the demand curve from infinity at A to zero at point B. Fig. 4.5 summarises the elasticity values on a demand curve.

The same result can be derived in a different way by using the property of similar triangles. Elasticity is equal to $(P_1/ Q_1) / (\partial P_1/ \partial Q_1)$. At point N in Fig.4.6 below (P_1/ Q_1) = NQ/OQ while $(\partial P_1/ \partial Q_1)$ = NQ /QB. Hence ε_N = (NQ /OQ)/(NQ/QB) = (NQ/OQ)/(QB/NQ) = QB/OQ. NQ gets cancelled out. The negative sign is ignored. The triangles AGN and NQB in Fig. 4.6 are similar triangles.

Hence, AN/GN = NB/QB.

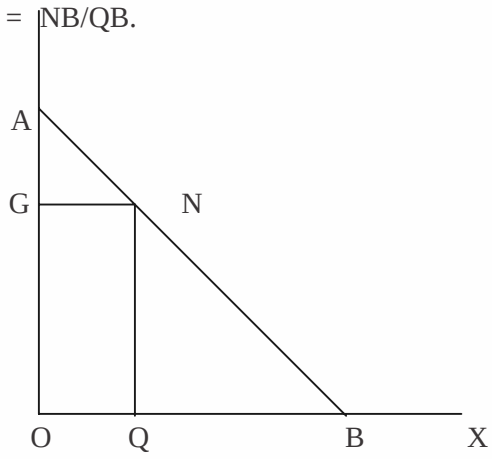


Fig. 4.6 Explains derivation of elasticity of demand using properties of similar triangles. eN turns out to be equal to BN / AN and also, QB / OQ as well as OG / AG.

Since GN = OQ, AN/OQ = NB/QB, by rearrangement of the terms gives

QB/OQ = NB/AN.

Since elasticity at N equals QB / OQ, it also equals NB / AN, a ratio of the lower

segment of the demand curve to the upper segment. Exactly in an identical way we can show, by using the properties of similar triangles (ANG and NBQ) that elasticity at N equals OG / GA.

Hence, elasticity at N is given by QB/OQ = NB/NA = OG/GA. By expressing elasticity as NB/NA, we can derive the following results:

Case I: when N is the midpoint of AB, NB = NA, hence elasticity equals unity $\varepsilon_N = 1$.

Case II: when N lying above the midpoint on AB, NB will exceed NA, as a result ε_N will exceed one. That is, Demand is elastic.

Case III: when N coincides with the corner point A. NA is zero. Hence, ε_N = NB/ 0 = $\propto$ (infinity) and demand is infinitely elastic,

Case IV: When the Point N lying below the midpoint of AB, NB is less than NA. Hence ε_N will be less than one and demand is less elastic.

Case V: When N coincides with the corner point B, NB equals zero.

As a result, ε_N = 0 /NB = 0.

Once again we get the same result obtained earlier that on a linear inverse demand curve elasticity falls in value as we move downward along such a curve.

Let us now take two parallel demand curves and compare elasticities at a (i) given price, and (ii) given quantity. [see Fig. 4.7(a) and Fig. 4.7 (b)]

In Fig 4.7 (a) we consider elasticity at the same price level on two demand curves.

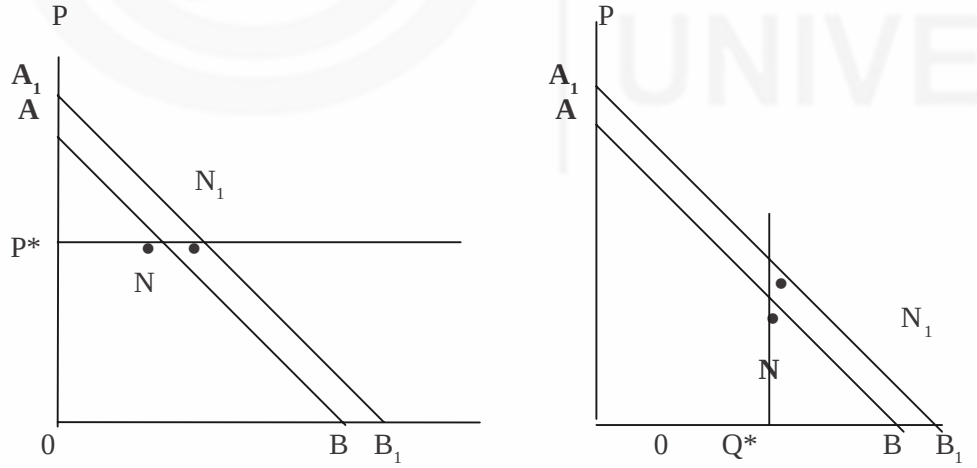


Fig. 4.7(a) shows that at a given price, two parallel demand curves will not show equal elasticities as lower segment is same but not the upper one.

Fig. 4.7 (b) we find that at same quantity also the parallel demand curves are not equally elastic. Now upper segments are equal but not the lower ones.

If we compare ε_N and $\varepsilon_N{}^1$, we find numerators are same (i.e., OP*) while P*A is less than P*A_1 in the denominator. Hence OP*/P*A is greater than OP*/P*A_1 implying ε_N is greater than $\varepsilon_N{}^1$. In other words, at a given price elasticity is higher on a lower demand curve AB than on a higher demand curve A_1B_1. In the second case, in Fig. 4.7 (b), at a given quantity OQ*

ε_N = Q*B/OQ* = NB/NA while $\varepsilon_N{}^1$ = Q*B_1/ OQ* = N_1B_1/N_1A_1.

This time denominators are same, OQ^*, while Q^*B_1 is greater than Q^*B in the numerator. Hence $\varepsilon_N = Q^*B_1/OQ^*$ is greater than Q^*B/OQ^*. As a result ε_N is greater than $\varepsilon_N{}^1$. In other words, at a given quantity, price elasticity is higher on an upper demand curve A_1B_1 than on a lower demand curve AB.

At this point let us look at Fig. 4.8, which shows two non-parallel demand curves AB and A_1B_1. At N, is demand more elastic on AB or A_1B_1? Elasticity is OG/GA on AB; while it is OG / GA_1 on A_1B_1. Since GA is greater than GA_1, elasticity is higher on A_1B_1 than on AB.

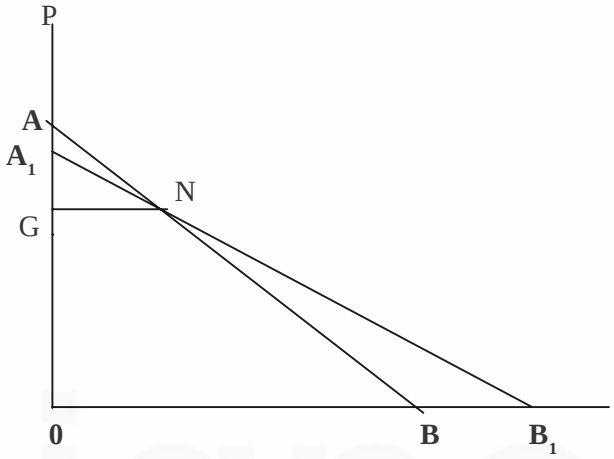


Fig. 4.8 If two demand curves intersect at point N, then we can compare their elasticities at N, using the relationship developed in Fig. 4.6; ε = OG/AG.

Suppose next an inverse demand curve that is non-linear and we want to find price elasticity at a point. How do we go about? N is the point on the non-linear demand curve D_1 in Fig. 4.9 (a). We draw a tangent line through point N. In other words, we apply a linear approximation at N. The line AB is such a tangent line.

Elasticity at N can then be expressed as NB/NA. As we move N along the demand curve D_1, the tangent line changes as is shown in Fig. 4.9 (b).

At point N_1 elasticity is N_1B_1/N_1A_1

At point N_2 elasticity is N_2B_2 /N_2A_2

At point N_3 elasticity is N_3B_3/N_3A_3

Elasticity at any other point on D_1 can be derived in a similar manner.

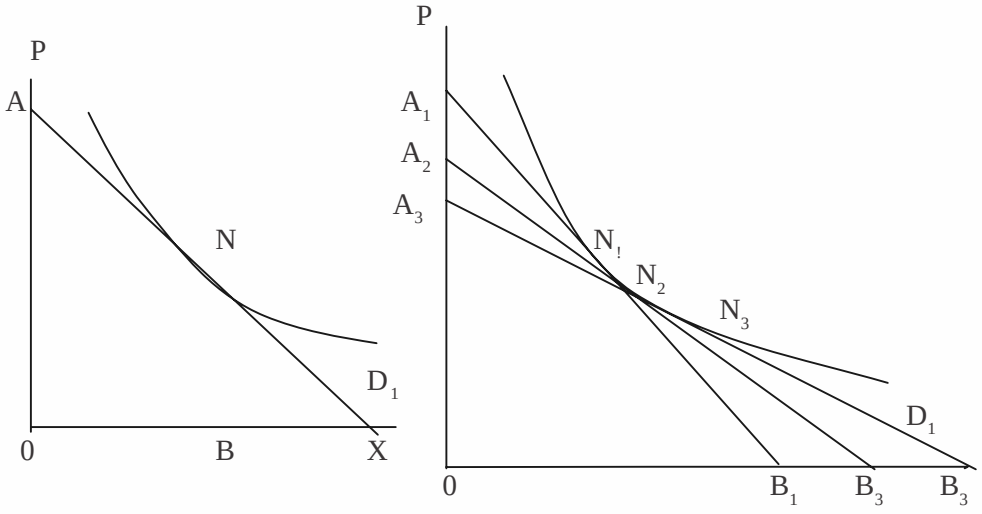
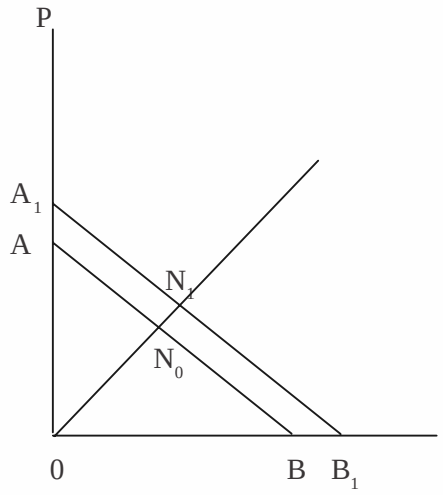


Fig. 4.9 Highlights the point that for a curvilinear demand relationship we take tangent to the curve at the relevant point N. Estimation of elasticity for that tangent is taken as approximate value of the curve at the point of tangency.

Check Your Progress 4

1. For the following diagram what will be elasticity at point N_0 and N_1?



4.7 PRICE ELASTICITY: THE OUTLAY METHOD

We have postulated that demand is a function of price. Therefore, as price changes along a demand curve, the quantity demanded will change. With such changes, an individual consumer's expenditure on the commodity would also change. By observing how this expenditure changes in response to price change, we can predict whether demand is elastic, less elastic or unitary elastic.

Remember that t*otal outlay, expenditure or revenue is price multiplied by quantity, i.e.*,

$R_1 = P_1X_1$

Take the case when price is reduced (i.e., P1 falls). Then this method says

(i) the total outlay or expenditure on the good could increase and demand for the good would be price elastic, i.e., $|\varepsilon_{11}| > 1$.

(ii) the total outlay or expenditure on the good could fall and demand for the good would be price inelastic, i.e., $|\varepsilon_{11}| < 1$.

(iii) the total outlay or expenditure on the good could remain the same and demand would be of unitary elastic, i.e., $|\varepsilon_{11}| = 1$.

When the price of the good rises this method would say that

(i) the total outlay or expenditure on the good falls demand is elastic, i.e., $|\varepsilon_{11}| > 1$.

(ii) the total outlay or expenditure on the good also increases demand is inelastic, i.e., $|\varepsilon_{11}| < 1$.

(iii) the total outlay or expenditure on the good remains the same demand is unitary elasticity, i.e., $|\varepsilon_{11}| = 1$.

Note that with the help of the outlay method, we cannot derive the exact value of elasticity. The calculated value can be used to indicate only whether ε11 is greater than or less than or equal to 1.

Check Your Progress 5

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

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

4.8 CROSS ELASTICITY OF DEMAND

In the demand function for good X_1 we hold the own price (P_1) constant and vary the price P_2 of a related (either substitute or complement) good X_2. Under *ceteris paribus* assumption, when P_2 varies the demand for good X_1 will normally change. The extent of demand change will depend on the concept of cross price elasticity of demand for good X_1. We use the notation ε_{12} for elasticity of demand for good X_1 with respect to a change in the price P_2 of related goods. It is defined *as the degree of responsiveness of the quantity demanded of X_1 with respect to a change in the price P_2*. Remember, this concept measures the extent to which demand for X_1 changes in response to a change in the price P_2.

$$\varepsilon_{12} = (\Delta Q_1 / Q_1) / (\Delta P_2 / P_2)$$

In ε_{12} numerator gives the relative change (proportionate change) in demand for X_1 while the denominator provides the relative change (proportionate change) in the price P_2. Multiplying both the numerator and the denominator by 100 we get the percentage changes.

If we use *arc method* then

$$\varepsilon_{12} = (\Delta Q_1 / Q_1^0 + Q_1^1) / (\Delta P_2 / P_2^0 + P_2^1)$$

If we use *point method* then $\varepsilon_{12} = (\partial Q_1 / Q_1) / (\partial P_2 / P_2) = (\partial Q_1 / \partial P_2) (P_2 / Q_1)$

The sign of cross-elasticity will indicate the nature of the relationship between the commodities, X_1 and X_2. If ε_{12} is positive, it would imply that X_1 is a substitute of X_2. On the other hand when ε_{12} is negative X_1 would be a complement of X_2. If ε_{12} is zero then X_1 and X_2 are independent. (There is no direct relationship though indirect one exists). *Note that price elasticity of demand is always negative in sign. But cross elasticity of demand can take both positive and negative values. It takes positive values for substitutes and negative values for complements*. The numerical magnitude of ε_{12} would indicate how strong is the inter-relationship between X_1 and X_2 and its sign will show whether it is one of substitutability or complementarily.

Check Your Progress 6

1. If cross elasticity of demand for X is 2.5 in terms of price of Y, is commodity X a substitute or complement of commodity Y?

   .....

   .....

   .....

2. For the following demand schedule, find cross elasticity of demand for good X when price of Y falls from Rs.5 to Rs.4. Comment on the nature of good.

Price of Y (Rs.)	8	7	6	5	4	3
Demand for Y	40	56	62	67	70	72

4.9 INCOME ELASTICITY OF DEMAND

So far we have considered two cases, viz., (i) price elasticity of demand where own price varies while prices of related goods and income of the consumers remains constant and (ii) cross elasticity of demand where price of related good varies with other variables remaining constant. Now let us analyse the effect of income changes on the quantity demanded of a good. Consider the demand function,

$X_1 = f(P_1, P_2 \; M)$

where P_1 is price of good X_1, P_2 is price of good X_2 and M is the consumer's money income.

When we hold P_1 and P_2 constant and allow M to vary, we will get a relationship between consumption of good X_1 and money income M. Such a relationship defines an *Engel's curve* for good X_1. It is named after a German statistician Ernst Engel (1821-1896) who studied the budgets of a large number of families in 1857 to derive the famous Engel's Law. *The concept of income-elasticity of demand for a good X_1 (to be denoted by ε_{1M}) is defined as the degree of responsiveness of the quantity demanded of X_1 with respect to a change in consumer's income*. It measures the extent to which demand for a good X_1 responds when the consumer's income changes. Formally, it is defined as the degree of responsiveness of the quantity demanded of X_1 with respect to a change in consumer's income. Thus, this elasticity of demand defines the percentage change in quantity demanded in response to percentage change in income of the consumer. Symbolically,

$$\varepsilon_{1M} = (\Delta Q_1 / Q_1) / (\Delta M / M)$$

where $\Delta Q_1 / Q_1$ is the relative (proportionate) change in demand for X_1 and $\Delta M / M$ is the relative (proportionate) change in money income.

Interpretation of income elasticity

(i) When 1 percent change in income leads to more than 1 percent change in demand for X_1, ε_{1M} exceeds one. The demand for X_1 is said to be income elastic.

(ii) When 1 percent change in income leads to less then 1 percent change in demand for X_1 then ε_{1M} is less than unity. In this case demand for X_1 is said to be income-inelastic. An implication of such an outcome is that demand is not very responsive to a change in income.

(iii) When 1 percent change in income leads to just 1 percent change in demand for X_1, ε_{1M} equals unity. It provides the dividing line between more elastic and less elastic demand.

For a finite change in income, the method of *arc elasticity* must be used where

$\varepsilon 1MA = (\Delta Q1 / Q_1^0 + Q_1^1) / (\Delta M / M_1^0 + M_1^1)$

where M_1^0 is income in one situation and M11 income in another. Q_1^0 is demand when income is MM_1^0.

When change in income is infinitesimal, *point method* must be used to get

$$\varepsilon_{1M}^{P} = (\partial X_1 / \partial M_1)\ (M_1 / X_1)$$

Again, those of you who are familiar with logarithmic differentiation may see that

$\varepsilon_{1M} = \partial \log X_1 / \partial \log M$

4.9.1 Classification of Goods on the basis of Income Elasticity of Demand

Normally, the income elasticity sign is positive. This is because when income increases a utility maximising individual would increase the consumption of good. But there may be cases when income increases results in decline of consumption of a commodity.

(i) A good is said to be a *normal good* when the income-elasticity is positive but less than unity.

(ii) A good is said to be a *superior* (luxury) *good* when the income-elasticity is positive and exceeds unity.

(iii) A good is said to be an *inferior good* when the income- elasticity is negative, i.e., when the consumer's income increases demand for the good falls.

The Engel's Curves for these three types of goods are given in Fig. 4.10.

Engel's curve for good X1 shows the relationship between income and quantity demanded of good X1. On the Y-axis we plot money income (M) and on the X-axis the quantity demanded

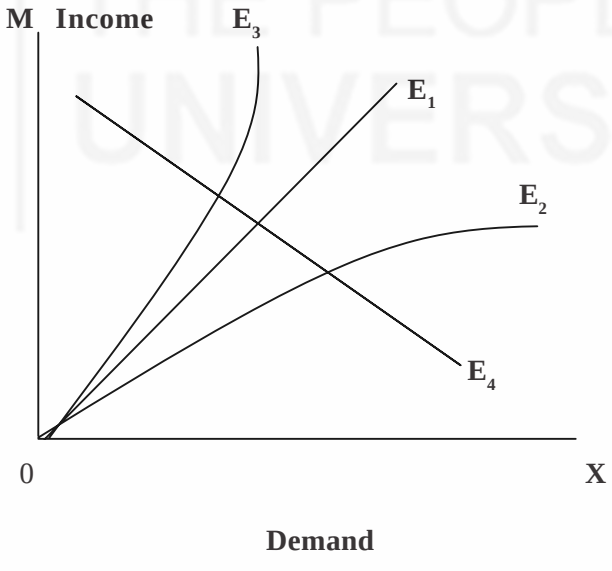


Fig. 4.10 Shows four curves describing the relationship between income (M) and demand for commodity X. We can call these curves income-demand curves so as to differentiate them from price demand curves shown in Figs. 4.1 to 4.9. These curves are also called Engel Curves. E_4 shows an Engel Curve for an inferior good. which has negative income elasticity. E_1 has elasticity equal to unity, E_2 show elasticity greater than utility and E_3 is the one with elasticity less than unity.

For normal and superior goods Engel's curve will be upward sloping. In Fig. 4.10 Engel's curve E_1 shows income elasticity of demand for X_1 to be equal to unity. Engel's curve E_2 shows income elasticity of demand to be greater than unity (luxury

or superior good). Engel's curve E_3 shows income elasticity to be less than unity (normal good). Engel's curve E_4 represent the case of an inferior good with income elasticity of demand negative. When income elasticity is positive, by observing the behaviour of the proportion of income spent on a good, one can predict whether income elasticity exceed, falls short of, or equals one.

If the proportion of income spent on good X (= P_1X_1/M) is constant, ε_{1M} must be unity since X_1 and M would have increased by the same proportion. In such a case, Engel's curve is a straight line. Remember that the slope of the straight line can be +1, greater than +1, less then +1 depending upon the share of expenditure on good X1.

In Fig. 4.11, Engel's curve E has a slope of +1, E_1 has a slope of greater than +1, this would imply P_1X_1/M less than 1. On all such curves the proportion of income spent on X_1 is constant. On E_2 the slope is less than +1

When the proportion of income spent increases as income increases ε_{1M} must exceed unity for X_1 increasing faster than income, as shown by the Engel's curve E* below (see Fig. 4.12).

Fig. 4.11 Engel Curve E shows income elasticity. E_2 shows a superior good with income elasticity exceeding unity. For the Engel curve E_1, lying above 450 line E, income elasticity is less than unity.

When the proportion of income spent decreases as income increases ε_{1M} must be less than one. This would imply demand for X_1 increasing less faster than income as shown by the Engel's curve E_1 below (see Fig. 4.13).

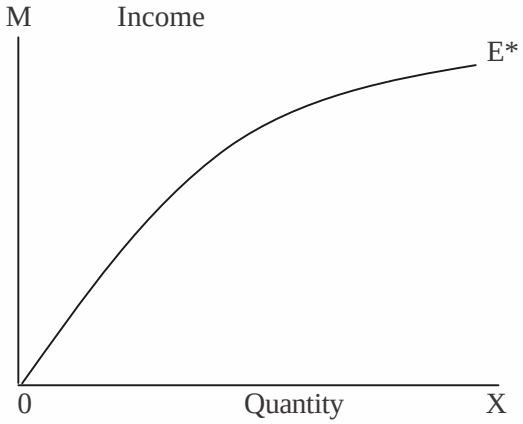


Fig. 4.12 As the proportion of income spent on a commodity rises, its income elasticity tends to become greater than unity.

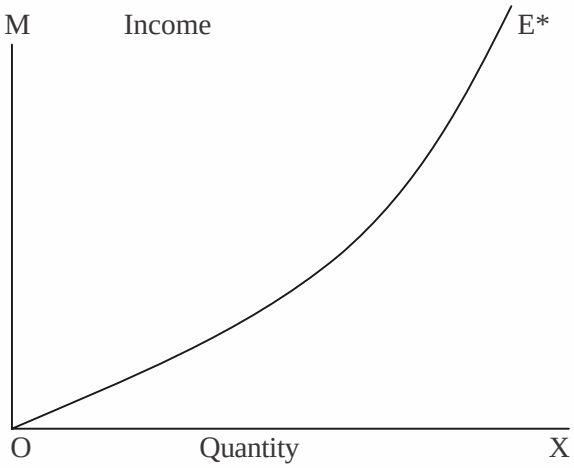


Fig. 4.13 As the proportion of income spent on X rises, its income elasticity falls.

Check Your Progress 7

1. For the following demand schedule calculate income elasticity when income rises from Rs. 500 to Rs. 600.

Income (Rs.)	400	500	600	700	800
Demand (kg.)	10	25	45	55	60

Hint. use arc method.

4.10 FACTORS ON WHICH ELASTICITY DEPEND

Now let us identify factors that influence the elasticity of a good. They are:

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

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

(iii) **Proportion of income spent on a good i.e.,** importance of the commodity in consumer's budget. Higher the budget proportion more strongly will the income-effect be felt. As a result demand for such goods will be highly responsive to price change. Hence, demand will be elastic. For example, all consumer durables like refrigerator, television, cooking range, geyser, motor bikes and washing machine tend to be more demand elastic. On the other hand, smaller the budget proportion, the more weak will be the income effect. Demand for such goods will be insensitive to price change. Demand will be less elastic. For this feature, take the example of salt, sugar, match sticks, certain vegetables, public transport etc. This factor - the

proportion of income spent on a commodity - will not only make demand more or less price elastic but will also determine the magnitude of income-elasticity.

(iv) **Level of the price of a good**: If the price of a good is very high (almost prohibitive) then a small change in it would generate a substantial income effect and would thus cause demand to be highly responsive to price change. As an example, consider the electronic household gadgets. The opposite happens when the price is very low - the income effect will be very weak. Hence, demand would be less elastic. For example, take cases like salt, match sticks, spinach, a cup of tea at *Dhabas*, etc.

(v) **Time period for adjustment**: Another factor is the time period over which demand adjustment is supposed to take place. Longer the adjustment period the higher the elasticity of demand. Tastes and preferences can be changed only in the long run, and through search, tastes for new goods can be acquired. As a result demand is more responsive in the long run than in the short run. Of course, it will also depend on how informative the consumers are, as well as on their mental make up.

4.11 IMPORTANCE OF THE CONCEPT OF ELASTICITY OF DEMAND

The concept of elasticity of demand (either price or cross or income) is important to any decision maker, be it a business firm, government policy-maker, an economic planner, or an international economic institution.

A business firm cannot fix its profit maximising price unless it has knowledge of *direct price elasticity* as well as *cross price elasticity* of demand for the good it produces. Since in oligopolistic market situation the firms are strategically interdependent in decision making, the concept of cross price elasticity of demand becomes more relevant in determining the price structure. Ignoring cross elasticities might prove to be disastrous in oligopoly situation.

Similarly, a government policy-maker, say, in the fiscal division, cannot determine the tax structure without the knowledge of price elasticity as well as income elasticity. For maximising tax revenue, taxes must be levied on goods with low price and income elasticities of demand. The same is true of government policy-makers responsible for fixing rates, charges, tariffs, fees, prices like electricity tariffs, railway fares and freight, dairy products, grains, taxi fares, public transport charges etc.On the other hand *economic planners* would not be able to fix the output targets of various goods during a plan period unless they are able to estimate the income elasticities of demand for various goods. Similarly, international economic institutions dealing with *world trade and currency exchange rates* must know the various demand and supply elasticities for policy interventions in trade and balance of payments situations of countries.

4.12 LET US SUM UP

In this unit we have investigated in great detail an important property of demand function for a good. That is, how demand responds when some of the variables on which it depends changes. Such demand response could be: (i) due to variation in price of the good under consideration (own price), (ii) price of a related good, either a substitute or a complement, (iii) consumers' disposable income (or just income if no taxes exist). Accordingly, we have derived own price elasticity, cross price elasticity and the income elasticity of demand for a good. We have also discussed two ways of viewing elasticity - *arc method and point method*.

For a finite or discrete change in price we use *arc method* (for instance when price

falls from Rs.10 to Rs.9, the change is one rupee, a finite change). When the change in price is very small (infinitesimal) we use point method (for instance when we are asked to find elasticity at price equal to Rs.6, the price at a particular point on the demand curve). The implicit assumption is that the price change is infinitesimal, the two points on the demand curve are very close to one another so that your bare eyes cannot visualise the difference. We have shown how elasticity can be measured geometrically. These apart we have also shown how cross elasticity can be used to determine whether goods are substitutes or complements. For substitutes the sign will be positive while it is negative for complements. We have concluded this unit by describing the importance of the concept for policymakers, business firms and economic planners.

4.13 KEY WORDS

Arc elasticity : Demand response measured on a finite range of the demand curve, that is, for a discrete price change.

Contraction in demand : A movement along a given demand curve in the upward direction. When quantity demanded decreases due to an increase in own price of the good.

Decrease in demand : A shift of the demand curve to the left. Quantity demanded decreasing due to a change in the other influences on demand (non-price change).

Demand function : Expresses a causal relationship between quantity demanded of a good and its own price.

Elastic demand : When demand response outweighs the price change in the opposite direction.

Expansion in demand : A movement along a given demand curve in the downward direction. When quantity demanded increases due to a fall in own price.

Income elasticity : Degree of responsiveness of quantity demanded of a good when consumer's income changes.

Increase in demand : A shift of the demand curve to the right. Quantity demanded increasing due to a change in the other influences on demand (non-price change).

Inelastic demand : When demand response is weak. The change in demand is not very significant induced by a price change.

Point elasticity : Demand response measured for a very very small(epsilon) changes in the price of a good.

Price elasticity : Degree of responsiveness of quantity demanded when own price of the commodity changes.

Unitary elastic demand : When demand response equals price change in the opposite direction. A situation in which a one percent change (fall or rise) in price leads to a one percent change (rise or fall) in quantity demanded.

4.14 SOME USEFUL BOOKS

Same as given in Unit-1.

4.15 ANSWERS OR HINT TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1 Shifts to the right

2. Expansion in demand, contraction in demand

Check Your Progress 2

1. Price falls by ten per cent, demand increases by 15 percent. Demand is said to be elastic, elasticity greater than unity.

2. $\varepsilon_{11} = 0.338$
   Price elasticity less than unity, demand is inelastic.

Check Your Progress 3

1. By arc method demand is inelastic.

2. By point method elasticity is less than 1.

Check Your Progress 4

1. Elasticity is the same at N_0 or N_1

Check Your Progress 5

1.

Price elasticity	Quantity	Total Outlay	Direction
10	60	600	—-
9	75	675	$\varepsilon>1$
8	80	640	$\varepsilon<1$
7	84	582	$\varepsilon<1$
6	97	582	$\varepsilon=1$

Check Your Progress 6

1. Since the sign is positive it is a substitute commodity.

2. Cross elasticity = { 3 / (67 + 70) } / {-1/ (5+4)} = -27 / 137. = 0.197
   The good is a complement.

Check Your Progress 7

1. Income elasticity = {20 / (25 + 45)} / { 100 / (500 + 600)} = 3.14
   Hence income elasticity is greater than unity, demand is said to be income elastic.

UNIT 5 CONSUMER EQUILIBRIUM: CARDINAL AND ORDINAL APPROACHES

Structure

5.0 OBJECTIVES

This unit will enable you to:

- understand and analyse how a consumer attains equilibrium;
- use cardinal utility theory to explain consumer behaviour;
- describe the law of diminishing marginal utility;
- explain consumer's equilibrium in terms of the Marshallian law of equi-marginal utility. Also use this law to explain the law of demand;
- explain the concept of consumer's surplus;
- explain consumer's behaviour in terms of ordinal utility theory, the Hicks-Allen approach
- describe consumer's equilibrium condition in terms of ordinal utility theory;
- decompose price effect into substitution effect and income effect;
- graphically derive price consumption curve and income consumption curve, and demand curve for a good;
- understand the difference between normal, inferior, and Giffen goods;
- provide a comparative evaluation of the two competing theories.

5.1 INTRODUCTION

In the previous unit we have introduced the concept of demand function, various determinants of demand and its elasticity. In this unit, we continue the discussion on

demand and focus our attention on consumer's behaviour in order to explain the law of demand. The law of demand says that when price of a commodity is lowered a larger quantity is demanded, and when price rises a smaller quantity is demanded, other things remaining the same. In other words, the law states that price and quantity demanded are inversely related. In this unit we will introduce you two contending theories - Alfred Marshall's *cardinal utility theory of demand*, and J.R. Hick's and R.G.D. Allen's *preference approach* (or the indifference curve theory, or the ordinal utility theory) of consumer behaviour.

In Hicks-Allen approach some of the restrictive assumptions of the *Marshallian* approach are dropped. Particularly, that utility is a cardinal concept and is measurable on a numerical scale with an absolute zero and that marginal utility of money is constant are relaxed. *Marshallian* theory is also based on the law of diminishing marginal utility as well as on inter-personal comparisons of utility. In the preference approach these limitations are overcome with the help of Hicks-Allen formulation, which is based on the indifference curve technique. We will first develop the properties of indifference curves. Using the indifference curves and in conjunction with prices of goods and the consumer's money income (or budget) we will be showing how a rational consumer attains equilibrium.

Since consumer's choice depends on prices and money income, and as prices change or money income changes, the consumer's equilibrium choice will also change. We explain how to derive the price-consumption curve and the income-consumption curve. We will then show how the demand curve for a good can be derived from the price consumption curve. This part of the discussion ends by pointing out the difference among normal good, inferior good and Giffen good. It is only in the case of Giffen good that the law of demand is violated and the demand curve for a good is upward sloping rather than downward sloping. The law of demand need not be violated in case of inferior good. As we will be showing it, all depends on the working of two opposing forces -the substitution effect and the income effect.

5.2 THE CARDINAL UTILITY APPROACH

Alfred Marshall (1842-1924), an important member of the neo- classical school of economics, gave us the cardinal *utility theory* of consumer behaviour in his book ***Principles of Economics*** (1890). According to him, a consumer derives utility from consuming a commodity. Following ***Jeremy Bentham*** (1748-1832) the founder of the Utilitarian School of Ethics, utility is defined as the subjective sensation - pleasure, satisfaction, wish fulfilment, cessation of need - which are derived from consuming a commodity and the experience of which is the object of consumption. Marshall assumed that utility (which is the want satisfying power of a commodity) could be measured quantitatively in the same way as one can measure weights and heights. In other words, utility is cardinally measurable - numerical or quantitative scale exists for measuring it. See that this is a very highly restrictive assumption. For instance, it is possible to say that a person, say, Mili gets 2 units of utility from a cup of tea. If utility is a cardinal concept, then it requires a complementary assumption specifying the unit of measurement. Bentham used a psychological unit of measurement called Utils. However, it cannot be taken as a standard unit for measurement due to its variation from individual to individual. Hence, Marshall took money as the unit of measurement. It has the advantage of uniformity for all individuals in the economy. In the illustration above Mili would receive 2 rupees worth of utility from a cup of tea. Besides adopting money as a measuring rod for utility, Marshall made another complementary assumption. He assumed the marginal utility of money to remain constant for each consumer. That implies the measuring rod must remain constant.

Cardinal measurability of utility also implies that utilities derived from the consumption of different quantities of a commodity can be added and also can be compared across

various individual consumers. Thus, one can speak of total utility and marginal utility derived from consuming a commodity. Marginal utility (MU) is defined as the addition to total utility when an additional unit of a commodity is consumed. Thus, MU is the ratio of extra utility to an extra unit of the commodity consumed.

I. Illustration of MU

Quantity	Total Utility	Marginal Utility
commodity X	(TU)	(MU)
1	2.00	-
2	5.00	3.00
3	9.00	4.00
4	14.00	5.00
5	17.00	3.00
6	9.00	.00
7	20.00	1.00

For the first unit the marginal utility cannot be calculated. The dash sign in the illustration indicates that. For the second unit of commodity X consumed the total utility are 5. Hence marginal utility is (5-2)/(2-1)=3/1=3, since the change in quantity is only one unit ($\Delta X=1$). In other words,

$$MU = \frac{(TU_2 - TU_1)}{(\Delta X = 1)}$$

Where TU_1 is utility derived from consuming one unit of X and TU_2 is total utility derived from consuming two units of commodity X. In general, then

$$MU = \frac{TU_n - TU_{n-1}}{X_n - X_{n-1}}$$

where n and n - 1 are the number of units of the commodity consumed.

Another important assumption, which Marshall made, is independence of utility. What it means is that utility derived from, say, consuming a *Samosa* is independent of utility derived from consuming, say, *Sandwiches*.

Together, all these assumptions would imply that if our consumer's taste can be represented by means of a utility function of the form

$$U = f(X_1; X_2; Xn),$$

then such a function will have the property of additivity and separability. This would mean that $U = U(X_1) + U(X_2) + ... + U(X_n)$

where $U(X_1) = f_1(X_1)$

$U(X_2) = f_2(X_2)$

.....

.....

$$U(X_n) = fn(X_n).$$

Utility derived from a good depends on the quantity consumed of that good alone. And total utility or total satisfaction derived from consumption will depend on the sum of utilities derived from consuming all the commodities. The object of consumption is to make this total utility as high as possible.

It is important to remember that when a utility function is used to represent a consumer's taste, like $U = f(X_1)$, marginal utility derived from consuming, say, commodity 1 is given by the first partial derivative of U with respect to X_1, that is $\delta U/\delta X_1$, marginal analysis is based on calculus technique. The use of calculus method requires the assumption that each and every commodity must be perfectly divisible or as finely divisible as possible. So, consumption of any commodity can be varied in as small an amount as possible. This makes the utility function continuous and twice differentiable.

The ***Marshallian*** theory of consumer behaviour is also based on the non-satiation assumption. In other words, consumers are never satiated with any good. Satiation would imply that the marginal utility of a good becomes zero. Non-satiation also implies that more of a good is preferred to less of the same good.

Check Your Progress 1

1. The following table shows the relationship between total utility derived from consuming various quantities of milk. Calculate the marginal utility.

Qty. of milk (good X) in litres	1	3	5	8	12
Total Utility (in Rupees)	10	15	28	40	56

5.3 THE LAW OF EVENTUAL DIMINISHING MARGINAL UTILITY

The law of eventual diminishing marginal utility forms the basis of the *Marshallian* theory of demand. This law says that after sufficient quantity of a good is consumed, consumer experiences diminishing marginal utility from additional units consumed. To put it differently, the law states that after sufficient quantity of a good has been consumed each additional unit of consumption yields less and less additional utility. This law is based on introspection and has the following rationale: when a fewer units of a good are available, a utility maximising consumer would be using them to satisfying the most pressing (urgent) needs. However, as more and more units of a good become available, the needs to which they are used or put become less and less important and hence yield less and less additional utility.

Suppose that there is water shortage in your residential area, and you get only one tumbler a day, how will you use it? Surely, you will use it for drinking purpose only, and may be for cooking. But, suppose, you get one or two additional tumblers a day. You may then use it for bathing and washing. As more and more water becomes available you may start using it to satisfy less and less urgent needs, like cleaning your car, watering your garden, and if enough water still remains available, you may even involve in water fights - "activities that are far removed from the notion of water as an absolute essential for human survival". (J.QUIRK).

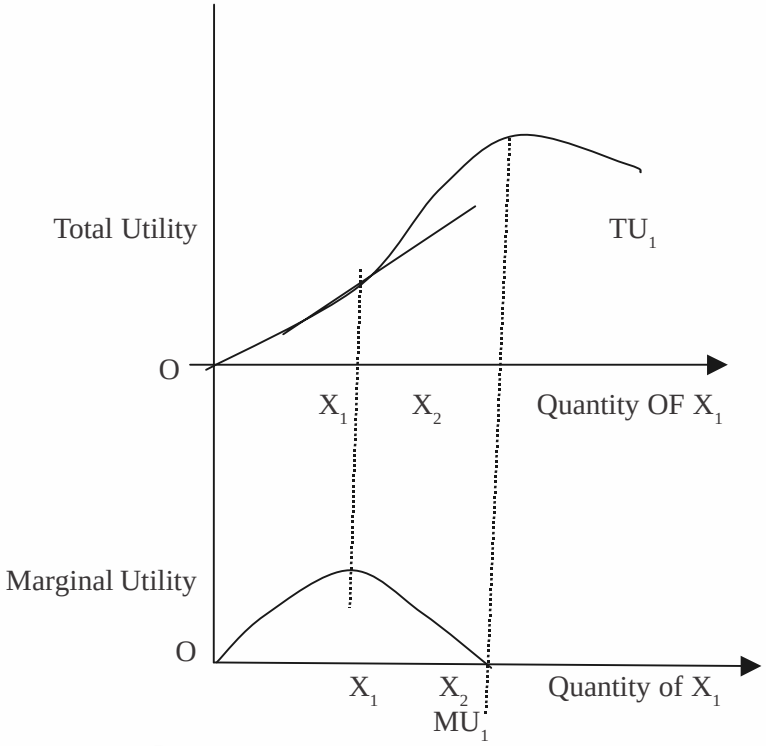


Fig. 5.1 The upper segment shows total utility curve TU_1 and the corresponding MU_1 is in the lower segment. Note that at OX_1, MU_1 is at its maximum and TU_1 shows the highest rate of rise. Beyond this point, TU_1 keeps rising, but at smaller and smaller rates. It reaches its maximum at OX_2 when MU_1 falls to zero.

As consumption increases from an initial low level (say, from zero), total utility increases at an increasing rate. This feature implies that the marginal utility is increasing (upto X_1). Beyond OX_1 for all successive consumption of X_1 total utility starts increasing at a diminishing rate. As a result marginal utility tends to fall. Hence, the law of eventual diminishing utility starts operating from OX_1.

Let us examine some other aspects of the law of diminishing marginal utility through an example.

When you are very thirsty, the first glass of water will give you a very high level of utility. The second glass might give you even a higher utility. But as you go on taking glass after glass of water, a point will be reached when you will not wish to have it any more. At that point you are completely satiated with it. In the diagram, at the quantity OX_2 total utility reaches a maximum and marginal utility becomes zero. Beyond OX_2 total utility decreases implying that marginal utility becomes negative. So, a utility maximiser will not go beyond this point. In fact, we will be showing below that a rational consumer will be attaining equilibrium in the range OX_1 and OX_2. That is the range where the law of diminishing marginal utility holds as well as the assumption of non-satiation.

5.4 CONSUMER'S EQUILIBRIUM

Let us assume that a consumer is consuming only two goods X_1 and X_2. The utility which she receives from consuming X_1 and X_2 is given by the utility function $U=f(X_1, X_2)$ and satisfies the property of eventual diminishing marginal utility. The consumer has a given money income to be spent on these two goods during the period we are analysing her behaviour. She cannot influence the prices, P_1 and P_2, of the goods, through her own action. Prices are given as parameters in decision-making (consumption) as this consumer is one of the numerous consumers demanding X_1 and X_2. Thus, she has no market power. Since she is required to spend her entire income

on X_1 and X_2, the budget equation is given as,

$$M = P_1X_1 + P_2X_2$$

Where M is her nominal income. Since the consumer is a utility maximiser, her consumption problem can be formulated as follows:

Maximise

$$U = f(X_1, X2)$$

subject to the budget constraint

$$M = P_1X_1 + P_2X_2.$$

By using the ***Marshallian Equi-Marginal Principle*** (which is based on the Lagrange ***Multiplier Technique)***, we get the equilibrium condition,

$$MU_1/P_1 = MU_2/P_2 = \lambda$$

This is the first order condition (necessary) for achieving equilibrium. The second order condition (sufficiency) of equilibrium is given by the law of eventual diminishing utility. The second order condition will be automatically fulfilled so long as the marginal utility schedules for each good MU_1 and MU_2 are both downward sloping. It must be noted that whenever a consumer maximises utility, equilibrium is said to be attained. What the equilibrium condition says is that to maximise total utility $[U = f(X_1, X_2)]$ an individual consumer must equalise the ratio of marginal utility to price for each and every good and which in turn must be equal to constant marginal utility of money. In other words, to obtain maximum total utility a rational consumer must equalise the marginal utility per rupee of expenditure on each and every line of expenditure (that is, on each and every good).

This relationship represents the consumer's equilibrium condition. A consumer attains equilibrium when she maximises total utility from consuming X_1 and X_2. The equilibrium condition can also be stated in an alternative form:

$$MU_1 / MU_2 = P_1/P_2 \quad ...(1)$$

The ratio of marginal utilities of goods X_1 and X_2 must equal the price ratio of the same two goods. This in turn, must equal marginal utility of money, which is constant by assumption.

Condition (1) above is the famous ***Marshallian law of the equi-marginal utility***. It can be shown that if the ratio of marginal utility of the two goods is not equal to the price ratio, then without spending more in the aggregate, just by re-allocating the given amount of money income as between the two goods X_1 and X_2, the consumer can increase her total utility from consumption.

An illustration*:*

The following table gives an individual's marginal utility schedules for goods X_1 and X_2. If the prices of X_1 and X_2 are Rs. 2.00 each and that the individual has Rs. 20.00 of Income, which she spends on X_1 and X_2, what is the individual's equilibrium purchase of X_1 and X_2?

Q	1	2	3	4	5	6	7	8	9	10	11
MU_1	16	14	11	10	9	8	7	6	5	3	1
MU_2	15	13	12	8	6	5	4	3	2	1	0

The solution: The individual's equilibrium purchase is given by the conditions

$$MU_1/P_1 = MU_2/P_2$$

and the budget constraint must be fully satisfied. From the above table we derive the following.

Q	1	2	3	4	5	6	7	8	9	10	11
MU_1/P_1	8	7	5.5	5	4.5	4	3.5	3	2.5	1.5	0.5
MU_2/P_2	7.5	6.5	6	4	3	2.5	2	1.5	1	0.5	0

At $X_1 = 6$

MU_1/P_1 is 4

At $X_2 = 4$ units

MU_2/P_2 is 4

Hence $MU_1/P_1 = MU_2/P_2 = 4$.

The amount spent is $P_1X_1 + P_2X_2$ which is Rs.20.00 (2*6+2*4 = 12 + 8). Money income is also Rs. 20.00. Hence, the budget constraint is satisfied. The equilibrium purchase is $X_1 = 6$ units and X_2=4 units.

Since MU_1 falls from 16 to 1 as X_1 increases from 1 to 11 and MU_2 declines from 15 to zero as X_2 increases from 1 to 11, the second order condition is also fulfilled.

When there are more than one combination of two goods (X_1, X_2) at which the equi-marginal principle holds, one has to take recourse to the budget constraint to obtain the equilibrium combination and all other combinations violating the budget constraint have been rejected.

It should be noted that when the consumer consumes n goods, the law of equi- marginal utility would then read as:

$MU_1/P_1 = MU_2/P_2 = MU_3/P_3 = ... = ..MU_n/P_n = \lambda$ (the marginal utility of money)

with the second-order conditions (the law of eventual diminishing marginal utility must hold for each of the n goods).

Further illustration: What is the maximum total utility, which the consumer derives from consuming 6 units of X_1 and 4 units of X_2?

Remember total utility equals sum of marginal utilities.

From goods X_1 the total utility derived is 16+14+11+10+9+8, which equals 68 units of utility. From good X_2 it is 15+13+12+8, which equals 48. Hence, the maximum total utility derived from consuming 6 units of X_1 and 4 units of X_2 is (68 + 48), which equals 116 units of utility. At the given prices of X_1 and X_2, any other combination of these goods would generate less than 116 units of total utility.

Check Your Progress 2

1. The following table shows total utility (TU) and marginal utility (MU) schedules of ***Roomali roti*** and ***Chicken curry*** for an individual consumer.

Qty:	Roomali roti		Qty:	Chicken curry	
	TU	MU		(TU)	'MU
0	0	-	0	0	-
1	25	-	1	10	-
2	40	-	2	-	6
3	50	-	3	21	-
4	-	-	4	24	-
5	59	3	5	-	-
6	-	1	6	27	1

(a) Fill in the blanks on the table.

(b) Suppose that consumer's income is Rs.12 and prices of *Roomali roti* and *Chicken curry* are Rs.2 per 100 gm. each. What is the utility maximising combination of *Roomali roti* and Chicken *curry*?

5.5 THE BASIS OF THE LAW OF DEMAND IN THE CARDINAL APPROACH

The demand function for a good is not to be confused with the utility function. The utility function of a good expresses the relationship between the consumer's intake (or consumption) of the good and the resultant psychic satisfaction, happiness and utility derived. Remember that utility is defined in terms of the subjective sensation which one experiences in her mind and the experience of it is the object of consumption. In the utility function quantities of goods appears as arguments. In a way it reflects consumer's taste or preference scale. Consumer's money income and prices of goods do not enter the utility function.

The demand function of a good on the other hand, expresses a relationship between the quantity demanded and its own price, *ceteris paribus.* For each price, it would indicate the maximum quantity demanded. Moreover, for each quantity demanded, it would indicate the maximum price the consumer would be willing to pay. This implies that the demand function is based on some kind of maximising behaviour of the consumer. The demand curve shows the graph of the demand function in the price-quantity axis. Such a curve indicates the consumer's intentions and reflects the maximum boundary for the consumer. Thus, the demand curve indicates what the consumer plans or intends to purchase and consume at alternative prices of a good. Hence, it is the consumer's planning curve. At each price it records the consumer's utility maximising choice. The demand function (and its graph the demand curve) is derived from the utility function by using an optimisation (maximisation) process. Such a demand curve would be negatively inclined implying that the quantity demanded of a good and its own price will be inversely related. A fall (rise) in price leads to an increase (fall) in the quantity demanded. It reflects the law of demand, which states that other things remaining the same (*ceteris paribus),* a reduction in the price of a good leads to a larger quantity of the good being demanded, while an increase in the price of a good leads to a smaller quantity of the good being demanded.

The basis of Law of Demand in the Marshallian analysis is the Law (based on introspection) of Eventual Diminishing Marginal Utility. Take two commodities X_1 and

X_2, whose prices are known and given to the consumer. The consumer's money income is also given over the period we are analysing her behaviour. The consumer's utility function is

$$U = U(X_1, X_2)$$

with $U_1 = U_1(X_1)$

$U_2 = U_2(X_2)$

and $U = U_1 + U_2$ using U_1 and U_2, marginal utilities of X_1 and X_2 are derived as

$$MU_1 = \partial/\partial X_1 [U(X_1, X_2)] = \partial U/\partial X_1$$

$$MU_2 = \partial/\partial X_2 [U(X_1, X_2)] = \partial U/\partial X_2$$

Let the prices of X_1 and X_2 be P^0_1 and P^0_2 and the money income of the consumer is M^0. Total income is spent on X_1 and X_2. So

$$M^0 = P^0_1 X^0_1 + P^0_2 X^0_2 \quad \text{...(i)}$$

The consumer's equilibrium condition is

$$MU_1/P_1 = MU_2/P_2 = \lambda, \text{ where}$$

λ= marginal utility of money. That is

$$MU^0_1/P^0_1 = MU^0_2/P^0_2 \quad \text{...(ii)}$$

Solving (i) and (ii) we get the equilibrium consumption of X_1 and X_2 to be X_{01} and X_{02}. This generates the point N_0 in the diagram below (for goods X_1). Next consider a situation where the price of good X_1 falls from P^0_1 to P^1_1 with the price of good X_2 and the consumer's nominal income remaining unchanged at P^0_2 and M^0. The consumer's original equilibrium is disturbed.

To restore equilibrium, MU_{01} must be reduced. MU^0_1 will be reduced if and only if quantity consumed of X_1 is increased. This follows from the law of diminishing marginal utility. At the new price of X_1, P^1_1, the consumption of X_1 is X^1_1; with

$MU_1/P^1_1 = MU_0/P^0_2$ $(=\lambda)$. This gives us the point N^1_0 on the demand function (curve) of good X_1. We repeat the exercise for all values of P_1. The locus of the consumer's equilibrium consumption points defines the demand curve (Fig.5.2).

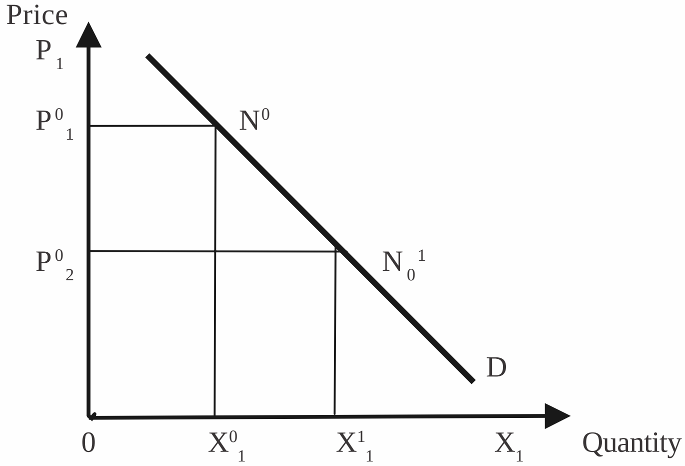


Fig. 5.2 shows the demand curve for the commodity X_1.It has been derived as locus of equilibrium consumption points of the consumer at different levels of price P_1.

The demand curve is downward sloping since the marginal utility schedule is downward sloping. However, the demand curve is not identical with the marginal utility schedule (unless λ is equal to unity). Consider the marginal utility schedule for good $X_{1,}$ as shown in Fig. 5.2.

Check Your Progress 3

(i) Suppose the marginal utility schedule is given as $MU_x = 40 - 0.5 Qx$. If marginal utility of money is equal to one and is constant, find out the demand schedule?

(Hint: set $P_x = MU_x$)

5.6 CONSUMER'S SURPLUS

Though Marshall was not the originator of the concept of consumer's surplus, he played a very significant role in providing it a theoretical structure, which could be used to derive many welfare propositions in economics. In particular it was used to measure the benefit derived by consumers from consuming a commodity (expressed in monetary terms). Marshall defined consumer's surplus as the difference between the price a consumer is willing to pay for a given quantity of a good rather than doing without it and the price that the consumer actually pays to acquire that quantity. It is the difference between the total willingness price and the actual price that one pays to acquire the good. Graphically it is illustrated as follows:

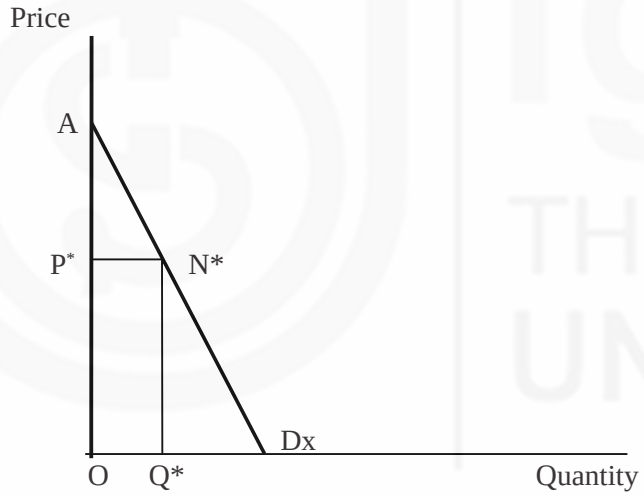


Fig. 5.3 shows consumer's surplus. As demand curve D_x shows maximum price per unit of X that the consumer is willing to pay, the area under D_x until point Q^*(= OAN^*Q^*) is her willing to pay. But market price is OP^*. So she actually pays only $OP^*N^*Q^*$. Thus the difference, that is, the area AP^*N^* is the consumer's surplus.

Let D_x be the demand curve to a particular consumer. At the price OP^* she consumes OQ^* quantity. For OQ* the total price that the consumer is willing to pay is given by the area under the demand curve upto OQ^* quantity. This area is OAN^*Q^*. The actual amount spent by the consumer is the area $OP^*N^*Q^*$. Consumer's surplus is the difference between the maximum (total) amount that the consumer is willing to pay and the actual amount that she pays. Hence, it is given by OAN^*Q^*- $OP^*N^*Q^*$. This equals the area AP^*N^* which represents the net gain to the consumer from consuming OQ^* quantity of good X. It must be noted that the Marshallian measure of consumer surplus is based on all the assumptions on which the Marshallian theory of demand is based. In particular, the assumption that utility of money is constant implies that Marshall ignored the income effect of a price change. Thus, the Marshallian demand curve only incorporates the substitution effect of a price change. Hence, consumer surplus must be measured on an income-compensated demand curve to reflect correctly

the gain to consumer from consuming a particular good. If one uses an observed demand curve to measure consumer surplus, it will be valid only in an uninteresting case of one per cent increase in income increasing consumption of each and every good by one per cent. The concept of consumer's surplus was used by Marshall and his disciple A.C. Pigou and many other neo-classical economists in deriving many welfare propositions and policy prescriptions in economics.

Check Your Progress 4

1 (a) When MU is zero, what happens to TU?

(b) When total utility is increasing what happens to marginal utility?

(c) When marginal utility is increasing what happens to total utility?

(d) When total utility is stationary, neither increasing nor decreasing, what happens to marginal utility?

2. Given the following MU schedule, when the price of good X is Rs.10 and marginal utility of money equals one, what will be the consumer's demand?

Qty. of X: (in kg.)	10	12	14	16	18	20	22
MU_x (in Rs.)	25	20	18	10	6	5	4

5.7 ORDINAL UTILITY APPROACH TO CONSUMER BEHAVIOUR: THE INDIFFERENCE CURVE

In the 1930s, an alternative theory of consumer behaviour was developed by J.R. Hicks of the Oxford University and R.G.D. Allen of the London School of Economics. It came to be known as *the preference approach to consumer behaviour*. It was called the preference - indifference curve approach and was developed as a response to increasing dissatisfaction with the Marshallian approach. In particular, economists of the neo-classical tradition started questioning the scientific validity of the assumptions of cardinal measurement of utility and constant marginal utility of money. Such assumptions, it was pointed out, reduced the explanatory as well as the predictive power of the model. It was also realised that those assumptions were not required to derive the law of demand and other related propositions of demand theory.

The preference theory starts with the premise that a consumer is able to express her preference for various commodity bundles. To give it a proper structure the preference function must satisfy certain properties. Before doing that some preliminaries must be discussed.

We use the symbol $\geq$ to mean preference indifference relations. Let us take two bundles (one can use the term market basket of goods), X_0 and X_1, consisting of quantities of two goods X_1 and X_2 (may be wheat and milk). $X_1 > X_0$ implies that the bundle X_1 is preferred to the bundle X_0.

$X_1 = X_0$ implies that the bundle X_1 is indifferent to the bundle X_0.

Hence $X_1 \; X_0$ implies that the bundle X_1 is preferred to or indifferent to X_0 bundle.

If $X_1 \geq X_0$ then X_0 cannot be preferred to X_1 at the same time.

An indifference curve is defined as the set of combination of two goods (X_1 and X_2) which give the consumer same level of utility or satisfaction or which are equally preferred.

Properties of Indifference Curves

We now describe the properties of indifference curves.

(i) Indifference curves are downward *sloping*. A consumer would be indifferent between bundles if these contain more of one good and less of the other. Since total utility is constant on a given indifference curve, if something of one good is taken away, the consumer must be compensated by other good so that total utility remains constant.

Let an important point to be noted: Indifference curves must not touch the axes. Touching the axes implies consumption of one good becomes zero, and the essence or the main thrust of the indifference curve approach is lost. The formulation that an individual gets satisfaction or utility from a combination of two goods consumed is lost. Bundles, which have mixture of goods, are always preferred to bundles having only one of them. If we impose this condition implying $U(0; X_2) = U(X_1, 0) = 0$ then indifference curves will not touch the axis.

(ii) Indifference curves *must not intersect*, otherwise higher indifference curves will not reflect higher utility. That is, bundles lying on higher indifference curves cannot be preferred to bundles lying on lower indifference curves. The property of transitivity, non-satiation as well as reflexivity, and completeness would ensure that indifference curves do not intersect. In such a situation, the preference direction would be North and East.

(iii) Indifference curves are strictly convex to the origin. This has something to do with the curvature of the indifference curves. The law of diminishing marginal rate of substitution (MRS), to be discussed below, ensures strict convexity of the indifference curves.

A consumer's taste or preference scale is represented by not one indifference curve but by a set of indifference curves, which we call the indifference curve map. The graphical representation of which would be as follows:

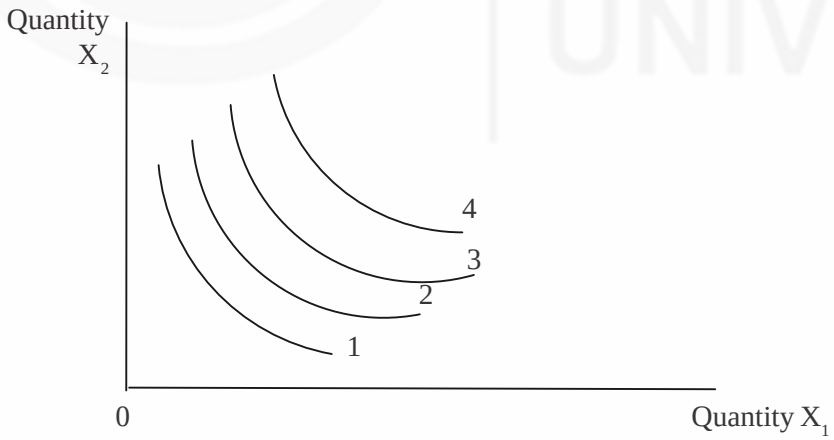


Fig. 5.4 shows the preference set of an individual. Each curve shows one level of satisfaction. A curve that lies to the North-East of any given curve, will show higher levels of satisfaction. These curves touch neither each other nor the axes.

We have shown only four indifference curves. In fact, there will be an indifference curve passing through each point in the commodity space. This is made possible by the assumption that the goods consumed are perfectly divisible (as finely divisible as possible). Hence, the commodity space can be jammed with indifference curves with their normal properties.

Marginal rate of substitution can be defined, as the rate at which commodity X_2 can be substitutional for X_1 in such a manner that consumer's total satisfaction remains constant. This rate keeps falling as we move along an indifference curve towards right. The idea is not so difficult to appreciate. We know that as the consumer gets larger quantity of X_1 he has to give up some of X_2. Why does it happen? As quantity of X_2 decreases, consumers "intensity" of liking for this increases, while the same for X_1 tends to decline. Hence for every increase in X_1 (of same magnitude) consumer will be willing to give up progressively smaller equilibrium of X_2 only. This diminishing MRS gives indifference curve the property of being strictly convex to the origin.

5.8 CONSUMER'S BUDGET CONSTRAINT

Since any individual consumer is one of the many consumers in the market for a good, she will not have any market power, implying that the market price of the good cannot be influenced through her own action. Also given is the consumer's money income (the budget) for the period we are analysing her consumption behaviour. The given money income in conjunction with the given market prices of goods would define the consumer's budget set, the feasible consumption choice set, or simply the consumption possibilities curve (line).

Let M_0 be the consumer's money income, X_1 and X_2 are the two goods consumed, P_1 and P_2 are the prices of X_1 and X_2. Then the budget equation is,

$$M_0 = P_1X_1 + P_2X_2$$

This implies that money income must be spent on the two goods, X_1 and X_2, and income equals expenditure. The budget must be completely exhausted in buying (purchasing). The activity of savings/dissavings does not give any utility/disutility to the individual consumer. The theory of consumer choice (behaviour) is a static partial equilibrium theory. We are not yet concerned with inter-temporal decisions. The graph of the budget equation would look something like the following:

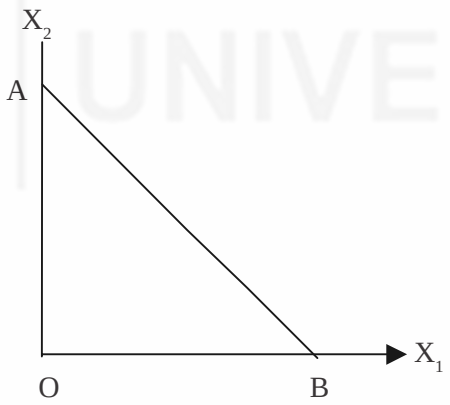


Fig. 5.5 shows the budget line. If all the money income M_0 is spent on two goods X_1 and X_2 at given prices P_1 and P_2 respectively, the consumer can buy one combination out of those given by line AB.

AB is the graph of the budget equation. The co-ordinate of 'A' is M_0/P_2 and that of 'B' is M_0/P_1. They indicate that if the entire income is spent on X_1 then the maximum amount of X_1 that the consumer can consume is OB. The consumption of X_2 would be zero. Similarly, if the consumer spends only on X_2, the maximum attainable consumption is OA, with consumption of X_1 zero. Any allocation of income between X_1 and X_2 would lie on the linear line AB. This line is called the *budget line* or real income line or expenditure line or the price line. It is linear as prices are constant.

Query: What happens to the budget line when money income and prices change in the same proportion? Say, money income doubles and prices double?

Let $M_1 = 2\ M_0$; $P_{11} = 2P^0_1$; $P^1_2 = 2\ P^0_2$.

The co-ordinate of 'A' will be M_1/P_{12}. This equals $2M_0/2P^o_2$, hence equals

M_0/P^0_2.

The co-ordinate of 'B' will be M_1/P^1_1. This equals $2M_0/2P^0_1$, which equals

M_o/P^0_1.

Hence, the co-ordinates of 'A' and 'B' remain unchanged. The line AB remains unchanged.

Check Your Progress 5

1. If the consumer's money income is Rs.100, price of good one is Rs.5, price of good two is Rs.10, draw the budget line.

2. When prices both the commodities double and money income also doubles, what happens to the budget line?

5.9 CONSUMER'S EQUILIBRIUM IN THE ORDINAL UTILITY APPROACH

A consumer attains equilibrium whenever she is maximising utility. This implies that in the preference approach, or, indifference curve analysis, the consumer chooses the most preferred commodity bundle, which she can buy with her budget. We have to combine the consumer's tastes or preference with the budget constraint and the consumer's indifference curve map is defined in the same commodity space. The indifference curve mapping is super imposed on the budget get the consumer's equilibrium or utility maximising choice. This is given by the highest indifference curve consistent with the budget.

In the figure below the indifference curves touch or intersect the budget line at five different points N_1, N_2, N^*, N_3 and N_4, giving us five possible equilibrium points.

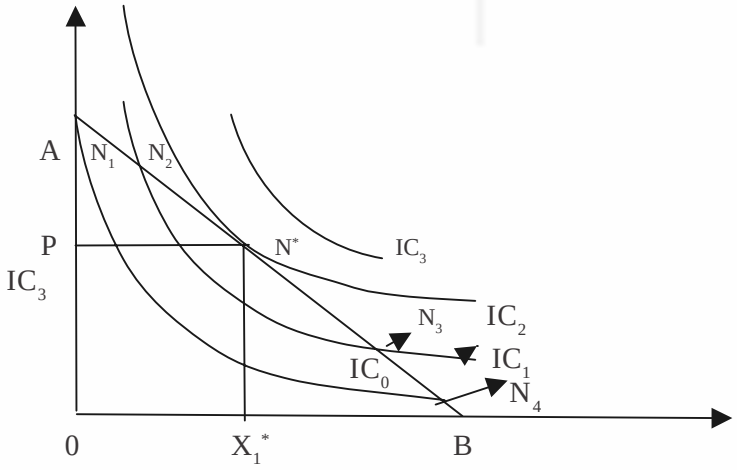


Fig. 5.6 shows indifference map and budget set of a consumer. The budget line comes in contact with three indicated indifference curves at N_1, N_2 , N_3 , N_4 and N^* points. But at N^* it merely touches IC_2 – making it the highest attainable IC.

At N_1 the budget is exhausted and the level of welfare (utility) attained is shown by indifference curve IC_0.

At N_2 the budget is exhausted and the level of welfare is shown by indifference curve IC_1.

At N_3 the level of utility is once again shown by indifference curve IC_1.

At N_4 it is shown by IC_0.

At N^* the budget is exhausted and the level of welfare (utility) attained is shown by IC_2.

None of the points N_1, N_2, N_3, N_4 and N^* violate the budget constraints. In terms of the level of welfare or utility attained the ranking of the points would be:

$$N^* > N_2 > N_1$$

or $$N^* > N_3 > N_4$$

Hence, the consumer's utility is maximised at N^*.

IC_2 is the highest indifference curve consistent with the budget constraint. With given money income and prices the consumer cannot attain the level of welfare indicated by the indifference curve IC_3. N^* is the most preferred commodity bundle since it has the highest ranking among the bundles in the attainable set. N^* represents consumer's equilibrium point. What is the characteristic of the equilibrium point? The budget line (price line) AB is tangential to the indifference curve, IC_2, at the point N^*. This implies that the slope of the budget line equals the slope of the indifference curve IC_2. Slope of the budget line is the price ratio (P_1/ P_2) while that of the indifference curve is dX_2/ dX_1. Both the slopes are negative. At the point N^*,

$$P_1/ P_2 = dX_2/ dX_1, \text{ with } dX_2/ dX_1 \text{ negative.}$$

Multiplying both sides by -1 we get $P_1/ P_2 = - dX_2/ dX_1$.

(-) dX_2/ dX_1 is the consumer's MRS_{X1X2}. Hence $P_1/ P_2 = MRS_{X1X2}$ at point N^*

This is the first order condition of equilibrium. The law of diminishing MRS gives the second-order condition of equilibrium. This implies that strict convexity of the indifference curves is sufficient to ensure the fulfilment of the second-order condition of equilibrium.

This will also guaranty that equilibrium is also unique. As

$$MRS_{X1X2} = MU_1/ MU_2,$$

The equilibrium condition can also be restated as

$$P_1/ P_2 = MRS_{X1\,X2} = MU_1/ MU_2,$$

or, $$MU_1/ P_1 = MU_2/ P_2.$$

Thus, we are back to the Marshallian equi-marginal principle of equilibrium condition without, however, the restrictive assumptions of measurability of utility and constant marginal utility of money.

5.10 SPECIAL CASES

Let us relax some of the assumptions of indifference curve analysis and suppose that the indifference curves are just convex, or, are straight lines with MRS constant. There are three possibilities, which will follow:

Case I

$P_1/ P_2 > MRS_{X1X2}$ at corner point A. The implication is

$P_1/ P_2 > MRS_{X1X2} = MU_1/ MU_2$, or, $MU_2/ P_2 > MU_1/ P_1$

Hence, per rupee of expenditure, the consumer gets higher marginal utility from good 2 than good 1. As a result the consumer would go on spending more on good X_2 till the entire income is exhausted on it and will consume OA of X_2 and zero amount of X_1.

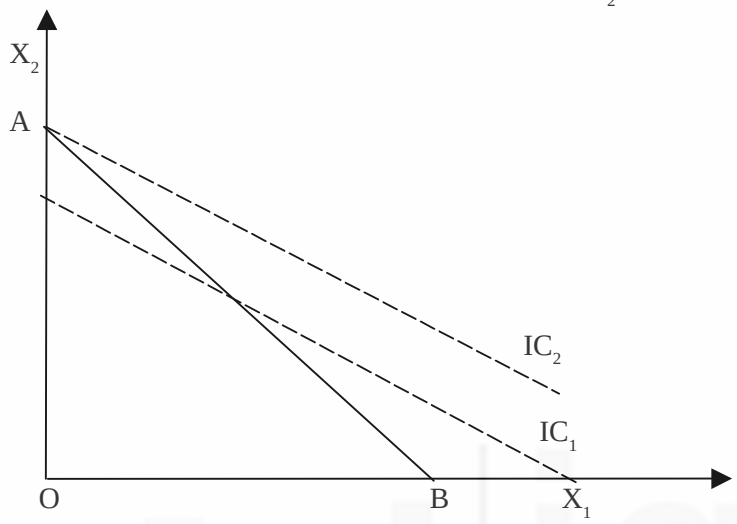


Fig. 5.7 (a)

However, though IC_2 is the highest indifference curve touching the budget line at point A, the equilibrium condition $P_1/ P_2 = MRS_{X1X2}$ is not satisfied and the second-order conditions are also not met.

Case II

Once again we will get a corner solution. This time $P_1/ P_2 < MRS_{X1X2}$.

Since $MU_1/ P_1 > MU_2/ P_2$, the individual gets higher utility from good X_1 than from good X_2. The entire money income will be spent on good X_1. We have a corner solution at point B; IC_1 being the highest indifference curve touching the budget line AB. Again equilibrium condition is not satisfied. The consumer consumes only X_1 equal to OB and nothing of X_2. We will have unique equilibrium.

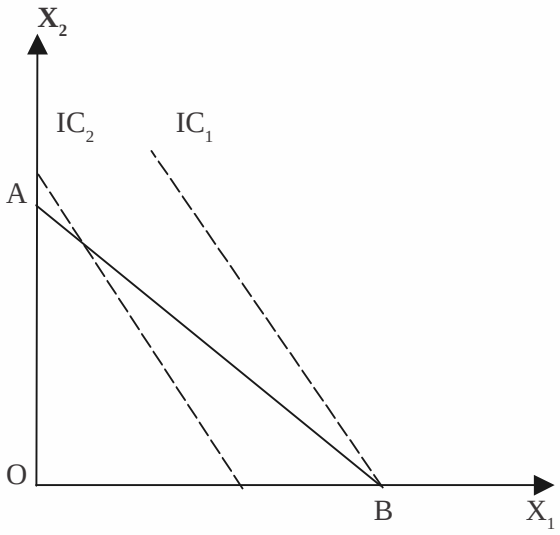


Fig. 5.7 (b)

Case III

In the third case the slopes of indifference curve and the budget line would be identical. One indifference curve, IC_1, overlaps the budget line AB. Hence, every point on the budget line AB from A to B is an equilibrium point. In other words, we get multiple equilibrium solution in this case. Note that at every point on AB, $P_1 / P_2 = MRS_{X1\,X2}$ and equilibrium condition is satisfied.

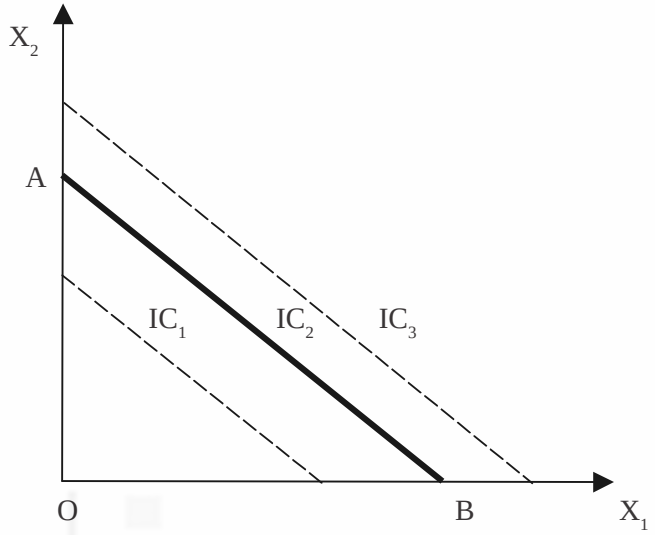


Fig. 5.7(c)

Next, Suppose indifference curves are concave rather than convex to the origin

Concavity of indifference curves implies that MRS_{X1X2} is increasing as we move down an indifference curve. As a consequence, at the point an indifference curve is tangential to the budget line, instead of maximising utility the consumer is in fact minimising utility since the second-order condition is not met. To maximise utility the consumer has to move to a corner point either A [Fig. 5.8(a)] or B [Fig. 5.8(b)]. Once again note that the consumer's equilibrium condition is not satisfied both at A and B.

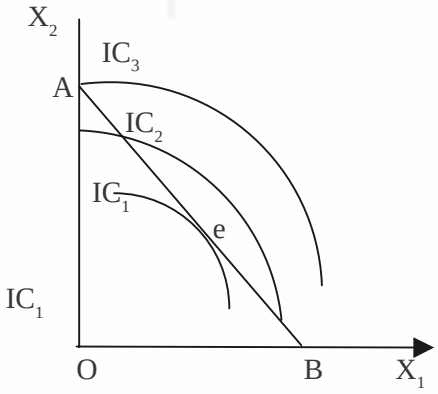


Fig.5.8 (a)

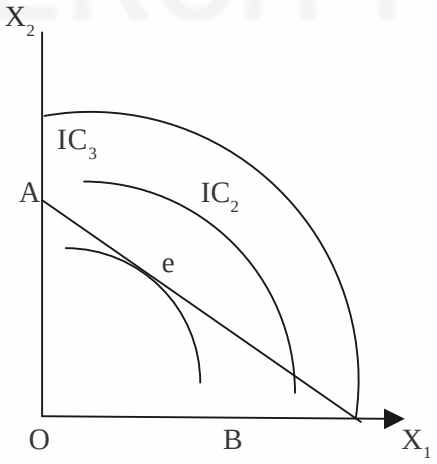


Fig. 5.8(b)

Note: Corner point solutions imply the consumer consuming only one commodity. The consumer is said to be a monomaniac.

Warning: It is quite possible that even with strict convexity assumption the consumer might attain a corner solution unless we invoke the restriction $U(O, X_2) = U(X_1, O) = 0$.

If the consumer gets utility only from a combination of two goods consumed then corner solutions are ruled out.

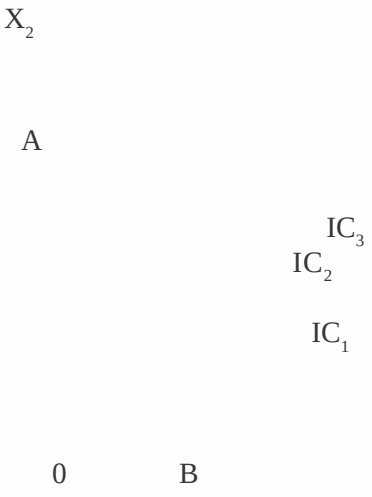


Fig.5.9(a)

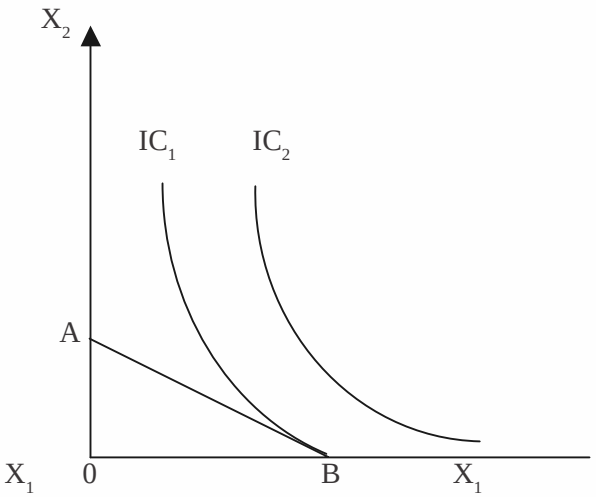


Fig.5.9(b)

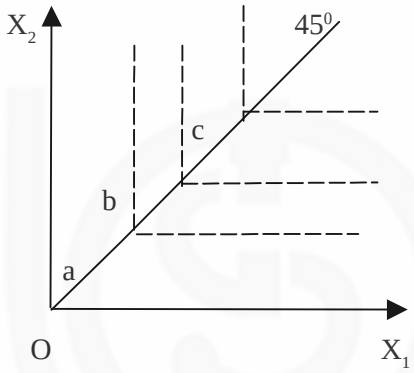


Fig.5.10(a)

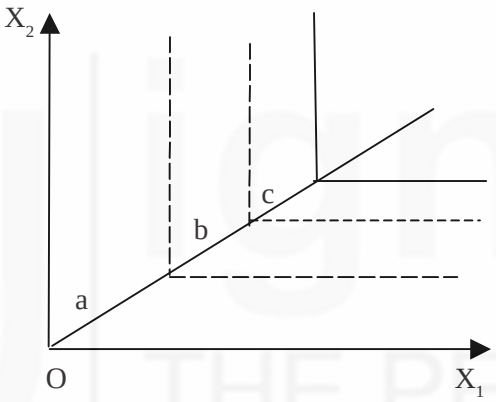


Fig.5.10(b)

Example (i)*:* The right and the left of a pair of shoes are perfect complements. If you have one of right shoes and two of the left then the extra left shoe will be totally useless. It will not add to your utility unless you are able to get an extra right shoe to match that of extra left. Let X_1 be right and X_2 be left shoes. Then the consumer will always operate at the corner points a, b, c of the indifference curves I_0, I_1, I_2 of Fig. 5.10.

On the horizontal and vertical segments, the marginal utilities of X_1 and X_2 will be zero. Utility increases when moving from I_0 to I_1 to I_2 only when the consumer is able to increase consumption of both X_1 and X_2 in a fixed proportion.

Example (ii)*:* If X_1 is hydrogen and X_2 is oxygen, to produce water we need to combine hydrogen and water in the ratio - two units of hydrogen and one unit of oxygen. Once again, the indifference curves are L- shaped with the consumer always operating at the corner points, a, b, c on indifference curves I_0, I_1, I_2. This time the corner points lie on a linear ray from the origin with a slope less than 45^o. If X_1 is oxygen and X_2 is hydrogen then the slope of the ray would be greater than 45^o.

Consumer's Equilibrium with L-shaped Indifference Curves

IC_2 is the highest (farthest from the origin) indifference curve touching the budget line AB, at point N*.Hence N* is the most preferred commodity bundle, the utility-maximising choice. However, the equilibrium condition cannot be applied here since at the corner point of an L-shaped indifference curve slope is not defined, as a result MRS_{X1X2} is not defined. Hence, we cannot equate price ratio P_1/P_2 to the consumer's MRS.

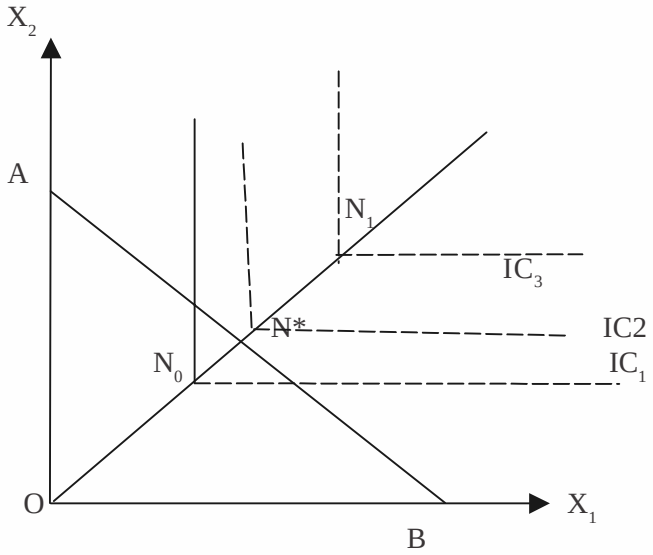


Fig.5.10(c)

Check Your Progress 6

1. The following table shows combinations of two goods X and Y along an indifference curve.

X:	0	1	2	3	4	5	6	7	8	9	10
Y:	30	23	17	12	8	5	3	2	1.2	0.5	0.0

(a) Calculate the marginal rate of substitution at each point on the indifference curve.

(b) When price of X is Rs.3.0 and the price of Y is Re.1.0, utility is maximised on this curve. What is the utility maximising combination? Explain.

5.11 THE PRICE-CONSUMPTION CURVE

Let us now consider a situation where the price of a commodity changes (falls or rises) with everything else (the price of the other goods and the consumer's money (nominal) income, as well as tastes) unchanged. We assume further that the consumer was initially in equilibrium. We show resultant outcome with help of Fig. 5.11.

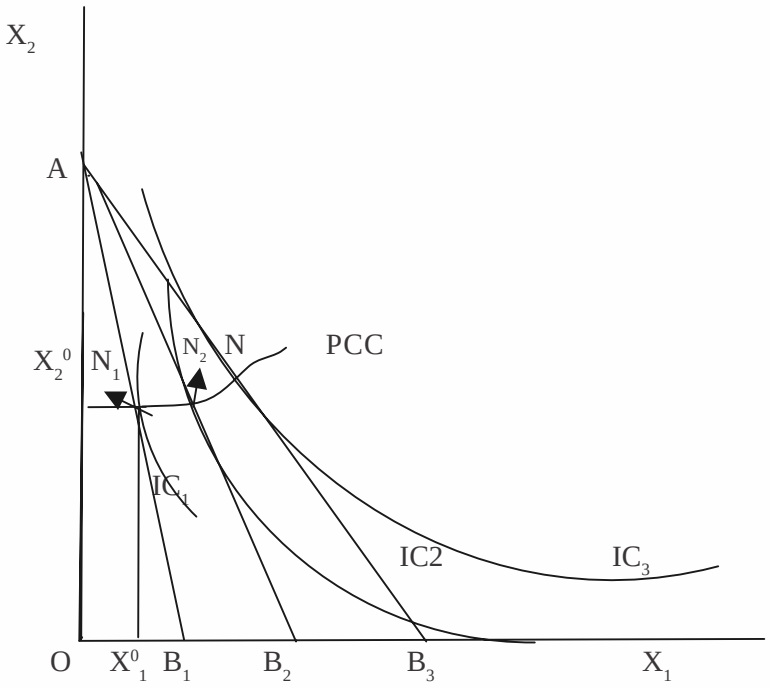


Fig. 5.11

With the original set of prices of the two goods P_{01} and P_{02}, and money income M_0, the consumer's equilibrium is at point N_1. It satisfies both the first-and-second order conditions of equilibrium. The utility maximising quantities of the two goods consumed are given by OX_{01} and OX^0_2. Next, let the price of X_1 fall continuously from P^0_1 to P^1_1 with P_2 constant at P^0_2 and money income is also constant at M_0. The point A remains unchanged. However, point B moves to the right to B_1, B_2, B_3 with $OB_1 = M_o/P^1_{1,}$ $OB_2 = M_o/P_1^2$ and $OB_3 = M_o/P^3_{1.}$

The purchasing power in terms of X_2 remains unchanged, and is equal to OA. When price of good X_1 falls the budget line becomes flatter and flatter (P_1/P_2 falls). Since the budget line shifts from AB to AB_1 to AB_2, the consumer's equilibrium shifts from N_1 to N_2 to N_3. The level of welfare (utility) increases from IC_1 to IC_2 to IC_3. With the shift in equilibrium points the equilibrium quantity of X_1 bought and consumed expands. If we take the locus of the consumer's equilibrium with changing prices of good X_1 we get the price-consumption curve (PCC) of good X_1 whose shape can be anything from downward sloping to horizontal to upward sloping. The shape depends on the amount spent on good X_1 as its price falls. With amount spent on X_1 increasing, PCC will be downward sloping. With amount-spent constant, PCC will be horizontal. Lastly, with amount spent decreasing, PCC will be upward sloping. But how do we discern the amount spent on good X_1? At the prices P^0_1 and $P_2{}^0$ the individual is in equilibrium at N_1. The two goods were bought in $X_1{}^0$ and $X_2{}^0$ quantities. Therefore, the amount spent on X_1 was $P_1{}^0X_1{}^0$. As price falls to $P_1{}^1$, the consumer buys $X_1{}^1$ quantity. Hence the amount spent is $P_1{}^1X_1{}^1$.

If $P_1{}^1X_1 > P_1{}^0X_1{}^o$, then we can infer from the total expenditure method that the elasticity of demand is greater than one. The PCC slopes downward to the right. If $P^1_1X^1_1 = P_1{}^0X_1{}^0$, elasticity will be unity. Finally, when $P_1{}^1X_1{}^1 < P_1{}^0X_1{}^0$, the demand is inelastic and the PCC slopes upwards.

Check Your Progress 7

Draw a price-consumption curve (PPC) for good X_1 when P_1 rises with price of X_2 and money -income being unchanged.

5.12 THE INCOME -CONSUMPTION CURVE

Starting from an initial equilibrium situation, we now allow consumer's money income to vary with prices of the two goods X_1 and X_2 remaining constant. With money income changing and prices of goods remaining constant, the consumer's budget line shifts outward (for a rise in income) and inward (for a fall in income) in a parallel manner.

In the following Fig. 5.12 when income raises the budget line moves from AB to A_1B_1 to A_2B_2 with price ratio P_1/P_2 given by the slope of the budget lines remaining constant.

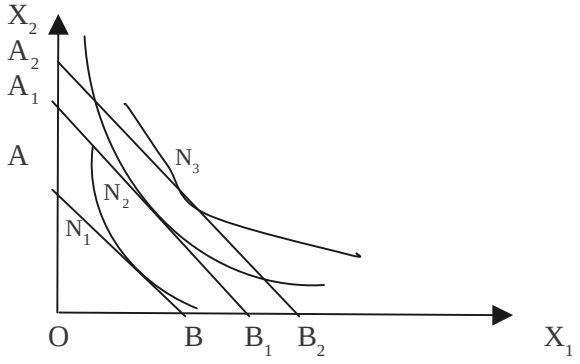


Fig. 5.12

When income increases and we record the consumer's consumption behaviour by taking the locus of consumer's equilibrium points we trace out the income-consumption curve for goods X_1 and X_2.

An income-consumption curve (ICC) for a good (say, X_1) shows the effect on consumption of good X_1 when money income varies with prices remaining unchanged. Such a curve can be linear or non- linear. At every point on such a curve (points like N_1, N_2, N_3), (P_1/P_2) = MRS $_{X1X2'}$ and since (P_1/ P_2) is constant MRS $_{X1X2}$ will also be constant. Moreover, indifference curves are also parallel.

5.13 THE PRICE EFFECT, SUBSTITUTION EFFECT, INCOME EFFECT

Starting once again from an initial position with money income M_0 and prices P^0_1 and P^0_2, the consumer attains equilibrium at N_1 consuming OX_{01} of good X_1 and OX^0_2 of good X_2 (see Fig. 5.11). Suppose that price of good X_2 falls, with money income and price of good X_2 remaining unchanged. What happens? We explain it with the help of Fig. 5.13.

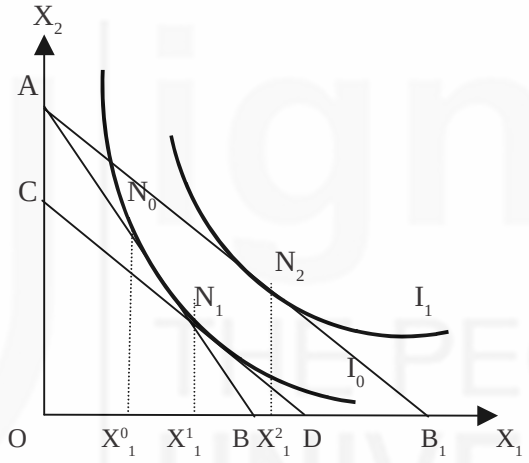


Fig. 5.13

The consumer's initial position is depicted by budget line AB and equilibrium is at N_0, the point where the indifference curve I_0 is tangential to the budget line AB. At the initial situation the quantity demanded of X_1 is OX^0_1. We now let the price of X_1 fall from P^0_1 to P^1_1 with everything else unchanged. *The budget line rotates about point* A in an anti-clockwise direction and becomes flatter. Its position moves from AB to AB_1. With P_2 remaining constant, the purchasing power of money in terms of good X_2 remains unchanged in OA. With P_1 falling the purchasing power in terms of goods X_1 increases from OB to OB_1. Originally at N_0, $P_1/P_2 = MRS_{X1X2}$. With P_1 falling, this equilibrium is disturbed and P_1/P_2 becomes less than MRS_{X1X2}. When the consumer is thrown out of equilibrium she undertakes consumption adjustment. With P_1 falling, she can attain N_2 on indifference curve I_1. At N_2 once again the price ratio (new P_1/P_2) becomes equal to MRS, which has also changed between N_0 and N_2. Both the first-order and second-order conditions are fulfilled at N_2. Note that *equilibrium point is always characterised by the tangency between budget line and indifference curve*. At the new equilibrium point quantity demanded increases from $OX_1{}^0$ to OX^2_1. When price of good X_1 falls the quantity demanded of X_1 increases by $X^0_1X^2_1$. This we call the price effect. It gives the change in demand induced by a change in own price.

When price of a good changes, the relative price ratio changes. From the consumer's point of view, good X_1 becomes relatively cheaper and even though the price of good

X_2 has not changed it has become relatively more expensive. Any change in relative price brings into play two forces: *the substitution effect and the income effect*. The substitution effect results from the fact that a rational consumer would always tend to substitute the relatively cheaper good for the relatively more expensive good. The income-effect results from the fact that when price of good changes the consumer's real income changes. Induced by this change consumer's consumption will change in general and consumption of the good whose price has changed in particular.

How do we separate these two effects? There are two approaches to the problem. One is the ***Hicksian*** method of decomposing the total price effect into substitution effect and income effect. The second method is due to a Russian economist, ***E. Slutsky***. In order to ascertain the magnitude of the two effects of a price change, we have to eliminate, first, one of the two effects. Normally the income effect is eliminated first in order to provide us with the magnitude of the substitution effect. Since income effect arises due to a change in real income brought about by a price change with money income held constant, to eliminate income effect we have to hold real income constant. This becomes necessary due to the fact that a change in real income causes consumption to vary. The real income can be held constant if money income can be adjusted to price changes. You know that

Real Income = Money Income/Price of a Good. That is, $RI = MI / P = M_0 / P^0_1$.

(i) When P_1 falls from P^0_1 to P^1_1 with M_0 constant real income increases. Hence money income needs to be reduced at the same time when price of a good falls. So, $RI = M_1/P^1_1$, where

$P^1_1 < P^0_1$ and $M_1 < M_0$.

(ii) When P_1 rises from P^0_1 to P^1_1 (with $P^1_1 > P^0_1$) and M_0 constant, real income falls. Hence to hold real income constant money income needs to be adjusted upward. New money income is M_1, with $M_1 > M_0$. In the process,

$RI = M_0/P^0_1$ becomes M_1/P^1_1.

An upshoot of the whole discussion is that to eliminate income effect consumer's money income needs to be adjusted to hold real income constant. With real income held constant, if the quantity demanded increases (decreases) with price of the good falling (rising), the change in demand (increase or decrease) must be due to operation of pure substitution effect. Once we have got the substitution effect, we restore the consumer's original level of money income, which will cause a further change in the quantity demanded of the good whose price has changed due to variation in real income. This latter variation in the quantity demanded (increase or decrease) is known as the income-effect of a price change.

The difference between methods of *John Hicks* and *E. Slustky* lies in their interpretation of holding real income constant. In Hicks, holding real income constant implies holding utility constant at the original level. While in Slutsky, holding real income constant means holding purchasing power of money income constant. In other words, when price changes the consumer should be given that much money income, which would make the original bundle of goods affordable. In the following, we will concentrate on the Hicksian method only.

Hicksian Method of Isolating Substitution and Income Effects

With money income at M_0 and prices of two goods at P^0_1 and P^0_2, the consumer's budget line is AB with the slope given as P^0_1/P^0_2. The consumer's equilibrium point will be at $P^0_1/P^0_2 = MRS_{X1X2}$ (see Fig. 5.13).

The quantity demanded of X_1 is OX^0_1. With price of X_1 falling to P^1_1, price of X_2

remaining at P^0_2 and money income at M_0 the budget line swings to $AB_{1.}$ The consumer's equilibrium shifts to N_2. The quantity demanded increases to OX^2_1.

The total increase in quantity demanded is $X^0_1X^2_1$, which is the price effect. The price effect (PE) equals substitution effect (SE) plus income effect (IE), that is,

PE = SE + IE.

In order to eliminate income-effect, Hicks would like the consumer's original utility to be restored since a price fall would increase utility from I_0 to I_1. To hold utility constant at I_0 we have come to reduce consumer's money income by AC in terms of good X_2 or DB_1 in terms of good X_1. The intermediate budget line CD shows the money income required at the new price ratio P^1_1/P^0_2 as indicated by the slope of the price line AB_1 to allow the consumer to attain the level of utility achieved before the price fall. The budget line CD is tangential to the indifference curve IC_0 at $N_{1,}$ the point where first and second order conditions for equilibrium are satisfied. The movement from N_0 to N_1 is Hicksian substitution effect. Consequently *the increase in quantity demanded from* OX^0_1 *to* OX^1_1 *is due to the Hicksian substitution effect.* This increase is caused by a fall in the relative price of good X_1, with real income held constant. We now restore the consumer's original money income (M_0). At the lower price of good X_1 the consumer is back on the budget line AB_1. On this budget line the consumer attains equilibrium at point N_2. The movement from N_1 to N_2 is the income-effect. The quantity demanded increases farther from OX^1_1 to OX^2_1. This increase is due to a gain in real income when price of a good X_1 falls.

Hence PE = $X^0_1X^2_1$

SE = $X^0_1X^1_1$

IE = $X^1_1X^2_1$

PE = SE + IE.

$X^0_1X^2_1 = X^0_1X^1_1 + X^1_1X^2_1$

Mathematically seen, the sign of substitution effect is always negative. This is because of quantity demanded and price moving in opposite direction. The sign of income effect is also negative for normal goods. This implies when price of a good falls (rises) as a consequence of real income increases (decreases), the demand for normal good will increase (decrease). Hence via the change in real income, quantity demanded and price move in opposite direction when we consider normal goods. For inferior goods, however, the mathematical sign of income-effect will be positive, implying quantity demanded and own price will move in the same direction. For a normal good the mathematical sign of price-effect will also be negative, because price-effect is the sum of substitution effect and income- effect. Since for a normal good the quantity demanded and own price are inversely related, the demand curve for the good will be downward sloping. The price effect has a negative sign.

What happens when the good (X_1) is an inferior good? By *definition an inferior good is a good whose consumption varies inversely with real income*. That is, if price remains constant money income and real income will move in the same direction. Let us analyse the sign of PE, SE and IE.

We have seen that PE = SE + IE

PE = - ve and + ve

as PE >, =, < 0

In case of normal goods, substitution and income effects reinforce each other by working in the same direction. However, in case of inferior goods they work in the

opposite direction. Thus, substitution *effect has a negative sign while income-effect has a positive sign. The sign of price-effect will depend on which of the two effects dominate (SE or IE).* So we have the following situations:

If SE dominates and outweighs IE then the price effect will still have a negative sign. The demand curve for the inferior good is still downward sloping. The law of demand continues to be valid

If IE dominates and outweighs the negative SE, PE will have a positive sign. The law of demand is invalidated. The demand curve for the inferior good is upward sloping. This is a perverse demand curve and we get the case of Giffen *Paradox*. The goods for which demand curves are upward sloping are called Giffen Goods.

SE and IE have equal weights. The price effect is zero. The quantity demanded does not respond when price changes. The demand curve for the good is a vertical straight line parallel to the price axis.

5.14 DERIVATION OF THE DEMAND CURVE FOR A GOOD

Normal Good Case

Now we turn to derivation of demand curve for a good. We shall first consider the case of normal goods. The upper panel of Fig. 5.14 is simply a reproduction of Fig. 5.13.

When X_1 is a normal good and its price falls, demand for X_1 is increased by $X^0_1X^2_1$ in the diagram. In the lower panel diagram the demand curve for X_1 is derived from the upper panel. In the lower panel, the horizontal axis measures the quantity demanded of X_1 as in the upper panel. However, the vertical axis in the lower panel measures the price of good X_1 given by the slopes of the budget lines. The initial price of good X_1, is given by P^0_1 = OA/OB; when price falls, it is given by P^1_1 = OA/OB$_1$ with $P^1_1 < P^0_1$.

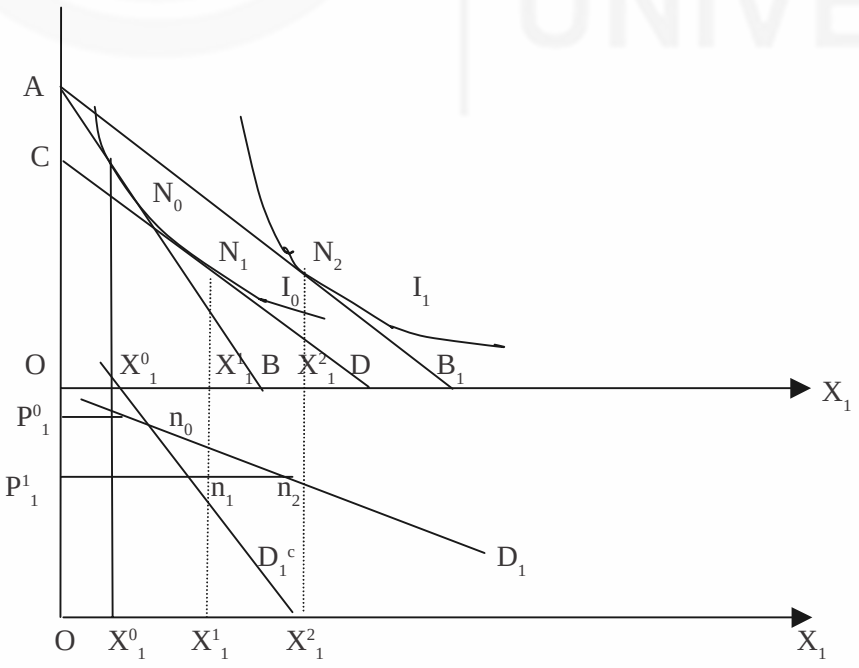


Fig. 5.14

At these prices of good X_1 the quantity demanded are OX^0_1 and OX^2_1. These are given by points n_0 and n_2 in the lower segment of Fig. 5.14. Joining points like n_0 and

n_2 we get the demand curve for X_1. We can generate more points like n_0 and n_2 by taking the other values of P_1. If X_1 is a normal good the demand curve for X_1 is downward sloping. The demand curve D_1 incorporates both the substitution effect and the income effect of a change in price of good X_1. In the lower panel there is another demand curve labelled $D^c_{1,}$ which is called the compensated demand curve for X_1. The consumer is compensated for the price change through an adjustment in her money income to hold real income constant (that is, utility constant). This demand curve incorporates only the substitution effect of a price change, income-effect being eliminated via the adjustment in money income. We record the quantity demanded at points N_0 and N_1 that is, OX^0_1 and OX^1_1. This information is translated into the lower panel diagram to give us point's n_0 and n_1 when the price of good X_1 is P^0_1 and P^1_1. Joining points like n_0 and n_1 we derive the compensated demand curve for X_1. The compensated demand curve will always be downward sloping and the mathematical sign of substitution effect is always negative. Its shape does not depend on the nature on the good, whether it is normal or inferior.

5.15 INFERIOR GOOD & GIFFEN PARADOX

In case of an inferior good whose consumption falls as the consumer's real income increases, and consumption rises when real income falls. On the other hand, the *Giffen goods are goods whose consumption falls when price of the good falls, and rises when price of the good rises.* These are illustrated below in Fig. 5.15.

Inferior Good Case

AB is the initial budget line with M_0 income and P^0_1, P^0_2 as prices of X_1 and X_2. On AB the individual is in equilibrium at N_0 with quantity demanded of X_1 being $OX_1{}^0$. Let P_1 fall from P^01 with income remaining at M_0. The individual attains equilibrium at N_2 with OX^2_1 as the quantity demanded of X_1. The Hicksian substitution effect is the movement from N_0 to N_1 on indifference curve IC_0, increases the quantity demanded from OX^0_1 to OX^2_1.

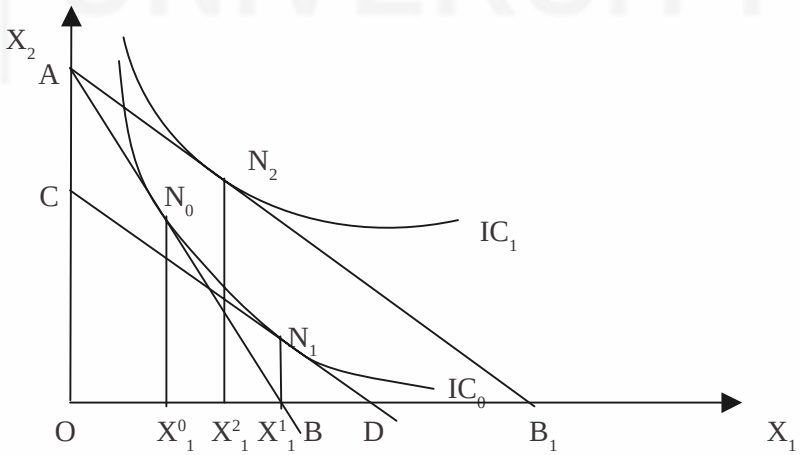


Fig. 5.15

However, in the case depicted here substitution effect dominates over income effect. Hence there is net increase in demand equal to $X^0_1\ X^2_{1.}$ The demand curve for X_1 will still be downward sloping being steeper than the demand curve for a normal good.

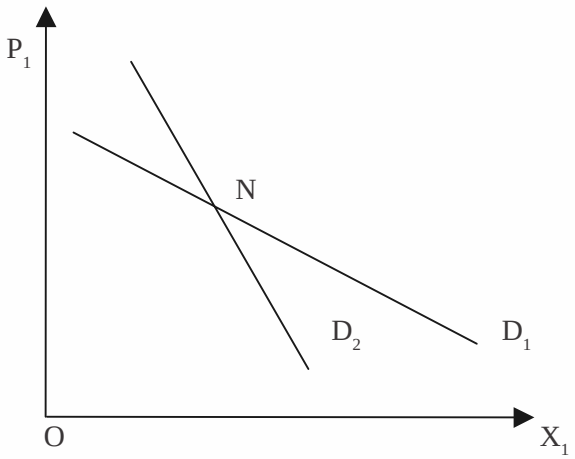


Fig. 5.15 a

D_1 is the demand curve for a normal good.

D_2 is the demand curve when X_1 is an inferior good.

The Giffen Good Case

This case is depicted in Fig. 15.16 below

In the diagram, X_1 is still an inferior good. When P_1 has fallen, the substitution effect leads to an increase in consumption from $OX^o{}_1$ to $OX^1{}_1$, by $X^0{}_1X^1{}_1$.

The income effect leads to a decrease in consumption from $OX^1{}_1$ to $OX^2{}_1$ by $X^1{}_1\ X^2{}_1$.

As you can make out from the diagram the income effect dominates over the substitution effect. As a result the quantity demanded falls as price falls. On balance demand has fallen by $X^0{}_1\ X^2{}_1$. The law of demand is violated as the demand curve for X_1 is upward sloping. This is the **Giffen** case or **Giffen paradox**. Mr. Robert Giffen, a British statistician of the late nineteenth century observed from empirical studies of household expenditures that consumers buy more of some goods (like brown meat), when price rises. Such goods are called **Giffen goods** after the name of its originator. For Giffen **Paradox** to hold, the following conditions must be met:

The income-effect must have a positive sign. Consumption of the good (X_1 in our case) falls as income rises, and rises as income falls. In other words, the good must be an **inferior good**.

(ii) The positive income-effect must outweigh the negative substitution effect to make the price effect positive. This implies price and quantity demanded having a positive correlation.

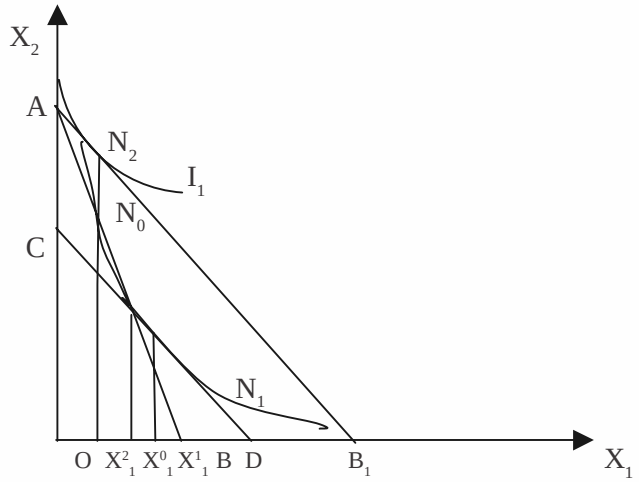


Fig. 5.16

Check Your Progress 8

1. Can complementary goods be inferior goods?

5.16 LET US SUM UP

In this unit we have covered a wide range of theme, starting from the Marshallian cardinal utility approach to demand theory. We end up with a discussion on inferior goods and Giffen paradox using the indifference curve approach. We saw how the law of eventual diminishing marginal utility formed the basis of the Marshallian explanation to the law of demand. We have also seen how the law of equi-marginal utility principle was used to derive the consumer's equilibrium condition. Though Marshall was aware of the Giffen paradox, he could not explain it, as he was unable to make a distinction between substitution effect and the income effect of a price change. The reason behind such a lapse on the part of Marshal was his assumption of constant marginal utility of money.

We have discussed the indifference curve approach in great detail. Starting from the properties of indifference curve, which reflects consumer's preferences we have tried to explain consumer's equilibrium. An important concept in this approach is the marginal rate of substitution (MRS) between any pair of commodities, as determined subjectively by consumer's preferences. It is supposed to be a psychological rate of exchange. The concept of MRS together with consumer's budget and the market prices of goods would determine the equilibrium choice for each consumer. In the Marshallian theory the second order condition of equilibrium is given by the law of eventual diminishing utility. On the other hand, in the indifference curve approach, or, Hicks-Allen approach the second order condition is given by the law of diminishing MRS. It must be noted that conceptually there is no connection between the law of diminishing marginal utility and the law of diminishing MRS.

We have also seen how in the indifference curve approach consumption of goods responds to a change in either prices or consumer's budget. From the first type of response we get the price-consumption curve while from that of the second we get the income-consumption curve. We then derived the demand curve for a good from the price-consumption curve. The shape of price consumption curve as seen would be depending on whether the good is normal, inferior or Giffen. In the same way the shape of the income consumption curve would also depend on whether the good is normal or inferior. In the last section, we have made clear the technical differences between an inferior good and a Giffen good.

5.17 KEY WORDS

Cardinal : The quantitative numbers used for measurements, like 1,2,3, and so on.

Consumer Equilibrium : A consumer attains equilibrium whenever marginal utility per rupee of expenditure is equalised on each and every good. In other words, the ratio of marginal utility to price must be equalised for each and every good, and must equal the constant marginal utility of money. This is also known as the Law of Equi-marginal Utility.

Consumer's Budget Constraint : A consumer given money income in conjunction with the given prices of goods defines the budget constraint. In other words, what combinations of goods are affordable is indicated by the budget constraint. It indicates what money income will buy.

Consumer's Surplus : Defined as the difference between the total prices that a consumer is willing to pay for a given quantity of a good rather than doing without it and the actual price that the consumer pays to acquire the given quantity. In other words, it is the difference between the total value of the good to the consumer and the total amount that the consumer pays on it. The concept was used by Alfred Marshall and his disciple A.C. Pigou to derive welfare propositions in economics.

Giffen Good : When consumption of a good and its own price move in the same direction, it is called a Giffen good. For such a good the demand curve is upward sloping. When price falls consumers buy less and when price rises they buy more.

Income Consumption Curve : It shows how consumption varies when a consumer's money income changes with prices of goods and preferences remaining unchanged.

Income Effect : The effect on consumption of a good when real income changes with relative price of the good unchanged. For normal goods the mathematical sign of income effect is negative whereas for inferior goods it is positive.

Inferior Good : When consumption of good changes in the opposite direction to a change in income, it is called an inferior good. For such a good, when income rises consumption falls, and when income falls consumption rises.

Indifference Curve : A curve showing the combinations of two goods, which would be equally preferred by the consumer.

Marginal Utility : Defined as additional utility per additional unit of the commodity consumed. It measures the change in total utility resulting from an extra unit of consumption of a commodity.

Marginal Rate of Substitution (MRS) : It is the psychological rate of substitution between any pair of goods defined on a given indifference curve. In other words, how much of one good the consumer must give up per unit of the other good acquired so that the consumer remains on an indifference curve. It is something, which the consumer works in her mind.

Normal Good : When consumption of good changes in the same direction as income changes it is called a normal good.

Ordinal : The numbers used to represent an ordering like 1st, 2nd and 3rd.

Price Effect : The change in consumption of a good when price of good changes with money income and price of the other good held constant.

Price-Consumption Curve : The curve that shows how consumption changes when price of a good changes with everything else unchanged including money income, prices of all other goods and

the consumer's taste.

Substitution Effect : The effect on consumption of a good when relative price of good changes with money income adjusted to hold real income constant. Mathematically, the sign of substitution effect is always negative.

Total Utility : The total satisfaction derived from consumption, which will be the sum of marginal utilities.

Utility : is defined as a want satisfying power of a commodity. It is the subjective sensation, which an individual derives from consuming a commodity. If such a sensation can be measured quantitatively on a numerical scale we call it cardinal utility.

5.18 SOME USEFUL BOOKS

For this unit, the books referred to in Block -1 (units 1,2,3) are useful.

5.19 ANSWERS OR HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1.

X	1	3	5	8	12
MU_x	-	5/2	13/2	12/3	16/4

Check Your Progress 2

In equilibrium $MU_x/P_x = MU_y/P_y$

Hence, 4 units of X and 3 units or Y will be consumed where X is *Roomali roti* and Y is chicken curry.

Check Your Progress 3

(i) $MU_x = P_x$ when $l = 1$ ($MU_m = 1$)

Since $MU_x = 40 - 0.5\ Q_x$,

$MU_x = P_x = 40 - 0.5\ Q_x$

Hence, the demand equation is

$P_x = 40 - 0.5\ Q_x$

The demand schedule coincides with the MU schedule.

Check Your Progress 4

(i) (a) TU is contant

(b) MU is positive

(c) Total utility is increasing at an increasing rate

(d) MU is zero.

(ii) In equilibrium $P_x = MU_x$

Hence, 16 units of X will be demanded.

Check Your Progress 5

(i) Read section 5.8 and answer.

(ii) Budget line remains unchanged.

Check Your Progress 6

(i) MRS = - dy/ dx

7/1, 6/1, 5/1, 4/1, 3/1, 2/1, 1/1, 0.8/1, 0.7/1, 0.5/1

(ii) $P_x/P_y = 3.0/1.0 = 3.0$

In equilibrium P_x/P_y = MRS

and MRS = 3 when X=5, and Y=5 units.

Check Your Progress 7

(i) Read 5.11 and answer.

Check Your Progress 8

(i) No, since complementary goods are consumed together. When income increases their consumption must go up, when income falls the consumption also falls together.

ignou
THE PEOPLE'S
UNIVERSITY

Indira Gandhi
National Open University
School of Social Sciences

EEC-11
Fundamentals of Economics

Block

2

Theory of Consumer Behaviour

Expert Committee

Mr. Kalyanjit Roy Chaudhary
St. Stephens College,
University of Delhi,
Delhi.

Mr. B.S. Bagla,
PGDAV College,

University of Delhi,
Delhi

Mr. R.S. Malhan,

St. Stephens College,
University of Delhi,
Delhi.

Mr. R.S. Bharadwaj
Shivaji Collage
University of Delhi,
Delhi

Dr. Gopinath Pradha
IGNOU

Dr. Narayan Prasad
IGNOU

Dr. Madhu Bala
IGNOU

Dr. Kaustuva Barik
IGNOU

Mr. Saugato Sen
IGNOU

Prof. S.K. Singh (retd.)
IGNOU

Course Editor	Course Coordinator	Block Coordinator
Mr Bhawani Sankar Bagla	Dr. Gopinath Pradhan	Dr. Gopinath Pradhani

Block Preparation Team

Unit No.	Unit Writer	Unit Editor (IGNOU Faculty)
4	Mr Kalyanjit Roy Chaudhary	Dr. G. Pradhan & Dr. K. Barik
5	Mr Kalyanjit Roy Chaudhary	Dr. G. Pradhan & Dr. S. Sen

Production	Cover Design	Word Processing
Mr. Arvind Kumar	Ms. Arvinder Chawla	Mrs. Rekha Mishra
Mr. Manjit Singh		Ms. Daisy Lal

March, 2002

ISBN-81-266-0449-2

Further information on the Indira Gandhi National Open University courses may be obtained from the university's office at Maidan Garhi, New Delhi-110 068.

Printed and published on behalf of the Indira Gandhi National Open University, New Delhi, by Professor Kapil Kumar, Director, School of Social Sciences.

Laser typeset by : HD Computer Craft, EA1/75, Main Market Inderpuri, New Delhi-110012. Ph: 5811437

Printed at :

BLOCK 1 THEORY OF CONSUMER BEHAVIOUR

This block deals with consumer behaviour in course of satisfaction of demand. Unit 4 discusses the demand theories, derivation of individual consumer's demand curve, market demand as a summation of individual demands and the concept of elasticity. The other unit (i.e., Unit 5) explores the process of consumer attaining equilibrium employing cardinal and ordinal approaches offered by economic theory. In general you will be exposed to utility maximising choice exercised by a consumer and the process of attaining equilibrium when there are constraints like limited income to be spent on a number of contending commodities. The condition of equality between marginal utility and price of a commodity as an equilibrium point will be shown to you.

INTRODUCTION TO ELECTIVE COURSE IN ECONOMICS EEC-11

Economics is no longer the preserve of those who practiced statecraft and regarded it to be confined to the ways and means of raising finances to meet the "requirements" of the ruling elite. The discipline has moved from such confines to the domain of the common man. It is now concerned with our day-to-day decisions such as: Which commodities to produce? How to produce? Which techniques to use? Which factors or resources to use in which combinations to produce what quantity of a commodity? Not only this, it shows which consumer may gain access to what specific amounts of different goods ? How to increase/decrease production of which good(s) in future? In other words, economics has moved away from financing the activities of state to helping the common man in the street to make many a crucial decisions impinging on their day-to-day life.

It must be remembered, however, that we have not moved from one extreme to another - from the state to the street. We, today incorporate a rather wide spectrum of activities in the domain of economics. These activities are (a) consumers' behaviour or choice process; (b) producers' behaviour or how is the production organised and carried on, what is the special role of cost functions therein and also the different forms of market organisations; (c) different individuals co-operate in the process of production to contribute factors owned by them. How do we determine their 'rewards'? Or, how do we distribute aggregate output among the members of society? (d) estimation of national (social) product and various aggregates, determination of level of income, employment and interest and also the relationship between money supply and prices; (e) some aspects of international trade; (f) public finance which not only incorporates all the aspects of meeting financial requirements of the state but also focuses on `newer' aspects of collective decision making.

The present course, Fundamentals of Economics (EEC 11), aims at exposing the learner to each of the above aspects. The course is divided into 9 blocks, spanning over 21 units. Block-1 is concerned with introducing the subject matter of economics along with nature of basic economic concepts and the methodology of this discipline. Block-2 analyses the behaviour of the consumer while Block-3 is concerned with technical specifications of production and cost functions. Block-4 uses information and knowledge gained in previous two blocks and analyses behaviour of the producers under different forms of market organisation. The theories of factor pricing, that is, determination of wages, rent, interest and profits in the society is our concern in Block 5. These five blocks constitute core of micro economic analysis.

Next three Blocks deal with what is popularly known as macro-economic analysis. Block 6 explains the idea of circular flows of money (and goods and services) in the society, and measurement of national income. In Block 7, we present various aspects of determination of income, employment and interest in the society. This block is essentially based on J.M. Keynes' contributions- though, at relevant points, we have also compared Keynesian ideas with 'classical' thinking about aggregative functioning of the society. In Block 8 , we are introducing relationship between quantity of money and price level on the one hand and those between rate of change of prices and levels of unemployment on the other. In this context we discuss Classical, Keynesian and Modern versions of quantity theory of money and Philips curve.

Finally, Block 9 introduces you to the basic aspects of public finance, public goods, externalities and market failure, public revenue and expenditure and various concepts of deficit in the government budget. The other unit in this block examines comparative cost theory of international trade, gains from trade, terms of trade and the structure of balance of payments accounts.

UNIT 6 THEORY OF PRODUCTION

Structure

6.0 OBJECTIVES

After going through this unit you should be able to:

- understand the concept of production function as the relation between output and inputs;
- deduce the Scale Line when production is subject to different returns to scale;
- identify the basis of the choice of technique; and
- analyse the operation of the law of variable proportion under variable returns to scale.

6.1 INTRODUCTION

Traditionally production is defined as "the creation of utility." This means we produce goods and services consuming which our wants are satisfied. Seen in such a context, therefore, production includes a wide variety of activities such as making of fabricated material goods, writing of a book for B.A. (Economics) students, making of a movie, or providing services as a beautician.

While "production" is referred to creation of any good or service, its concept is understood better when we speak only of goods. For example, you can say that

you have produced 100 bicycles in the month of January 1997. To produce this number of bicycles, you have spent some money to buy machinery, steel and labour. All these activities can now be grouped in two categories to define clearly what is involved in production. One category is put under output (bicycle in the above example). Production of output requires putting together the necessary materials. These constitute the other category and are called 'inputs'. Thus, you put together the inputs in a certain way to produce a specified quantity of output. A theory of production analyses how a producer combines various inputs to produce a specified quantity of output when the technology of producing the commodity is given. The relationship between quantities of inputs and output produced is termed **production function**.

6.1.1 What is an Input?

Input is any good or service that contributes to the production of an output. A producer normally, employs different inputs to produce a unit of output. Some of these inputs could be the outputs of other firms. For example, a producer of car uses steel as input. But steel itself is an output for the steel producer. There are inputs such as labour, which are not produced.

Inputs used in production may also be classified as fixed or variable. A fixed input is one that is used for production but its quantity does not change with the quantity of output produced. For instance, take the case of manager of a firm. There is one manager when one unit of output is produced. Even when 100 units of output are produced, the services of the single manager (not 100) only are required.

In contrast to the fixed input, variable input is dependent on the quantity of output produced. For example, the quantity of steel required to produce one car will be different from that of the production of more than one car.

It must be remembered that the distinction between fixed and variable inputs is related to a span of time. Inputs, which are fixed for one period of time, are variable for a longer period. Remember that a machine employed for production may be considered as a fixed input in the short-period (i.e., when the scale of output production cannot be changed). But given a sufficiently long period, the number of machines will vary in accordance with the requirement of the scale of production.

6.2 PRODUCTION FUNCTION: THE CONCEPT

You might have seen machines of different kinds used for the production of commodities. If you look at the functioning of a machine closely, you will notice that number of workers operating the machine is fixed. Take for example, a simple saw machine. Two people are engaged to operate it. When the machine employed for cutting a tree happens to be an axe, only one person is required to handle it. These examples indicate that once the machine is chosen, the number of people required for its operation is automatically known. Extending this simple observation to more complex technologies, you would like to infer certain general features. Perhaps, you would like to say, given a technology, the requirements of inputs (like, labour, capital and raw materials) are known. The quantity of output forthcoming from the application of the technology is also known.

What is being attempted to tell you is the relation between inputs and output, when technology is given. Thus, take the case of a production process, which uses two

variable inputs (x_1 and x_2) and one or more fixed inputs for the production of a single output (Q). The production function would state the quantity of output (Q) as a function of quantities of variable inputs (x_1 and x_2). This relation gives the idea of maximum output that can be produced from different combination of inputs (x_1 and x_2). Thus **production function depicts the technical relation between inputs and output** once the technology is given. This concept is used by economists to analyse laws of production relating inputs and output.

6.2.1 Product Curves

We have said above that a production function depicts the relationship between inputs and output. When other inputs are kept fixed and only one is allowed to vary, the total output produced exhibits some interesting features. To see these, take the help of a hypothetical example.

Suppose that you are experimenting with production of paddy and have land as a fixed and labour as a variable input. Take 8 blocks of land each measuring one-acre. Assume that all these blocks are equally fertile. Select a group of equally efficient workers and employ one worker in the first, two in the second and keep on increasing this input upto the last block of land to have 8 workers. Record the output produced in each block of land. These pieces of information put together will give you Table 6.1

You can see from the table that upto 6th tract of land, output increases as a result of greater number of employment of labour. With one labourer working in the first block, output yield is 10 units (say, bags of paddy), while it increases to 66 units in the 6th block as the labour employment was increased to 6. The 7th block of land, with 7 labourers, has produced 66 units of output, which is equal to the yield of 6th block. Thus, despite an attempt to employ one more worker production has remained at the level of 66 units. In the last block, with the highest labour to work at has yielded only 64 units of output. The production level is lower than 6th and 7th tracts of land.

Table 6.1: Employment of Labour and Production of Paddy

Tract of land	Labour employed	Output (Paddy)
1	1	10
2	2	24
3	3	39
4	4	52
5	5	61
6	6	66
7	7	66
8	8	64

The data in Table 6.1 can be represented in the form of a graph. This is done in Figure 6.1. In order to draw the graph, it may be useful to remember that output produced depends on inputs used. It is customary to plot the independent variable (labour in our example) on x-axis. The dependent variable (production of paddy) is

plotted on y-axis. Joining successive points of labour-output combinations from Table 6.1, total product curve is obtained. It is necessary to take note of an important feature of this curve. It first rises slowly. Then it increases an increasing rate and slows down again reaching a maximum. After wards there appears a decreasing phase.

Fig. 6.1

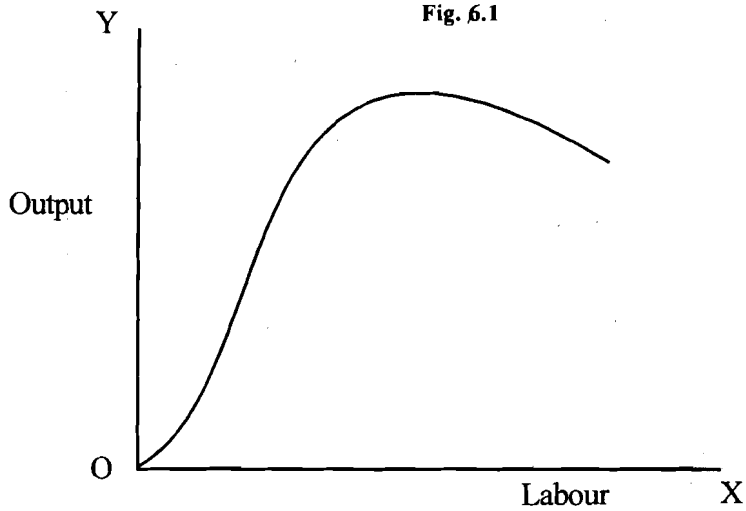


Fig.6.1: Depicts the relation between variable input, labour, and output. Output tends to rise as the use of labour increases. After a while, however, it starts declining.

Average and marginal products of labour can be defined on the basis of above formulation. The average product (AP) of labour is the total product divided by the number of workers employed (X). That is,

$$AP = \frac{Q}{X}$$

Similarly, the marginal product (MP) of X is the rate of change of the total product with respect to variations in quantity of X. To see the calculation of AP and MP from total output see Table 6.2 in the following. Features of AP and MP can be seen from their curves. These are presented in Figure 6.2. See that AP for a point on a total product curve equals the slope of a line segment connecting that point with the origin. It increases for movements along the total product curve from the origin upto a point and decreases thereafter. Similarly, MP increases initially and declines after a point. MP and AP are equal at the maximum AP.

Fig. 6.2

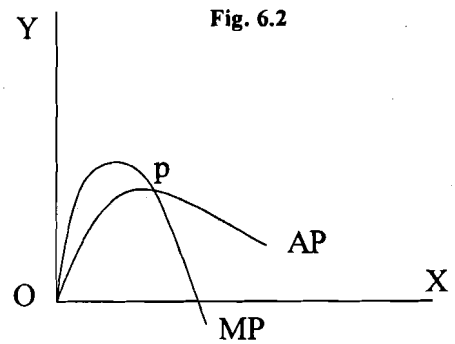


Fig. 6.2 shows the shapes of AP and MP curves. Notes that MP rises faster than AP initially, attains the peak earlier and starts declining faster as well. MP cuts AP at p, which is the maximum height attained by AP. As it falls below AP, goes to the negative level of output. In contrast, AP remains at a positive output level.

6.2.2 Law of Diminishing Marginal Product

This law states that the MP of input (say, x_1) will eventually decline as its use is increased keeping other inputs unchanged. It does not rule out the initial phase of an increasing MP. See, for example, a production process in which labour and land are combined for the production of a commodity (say, paddy). Keep the quantity of land constant and increase the labour input to produce paddy. You will notice that initially an increase in the employment of labour will result in increasing MP of labour. This process, however, will not last long. Increased application of labour will result in smaller increases in the output of paddy. This characteristic is called the **law of diminishing marginal product**. It operates when one input relative to other increases and will not be observed if both the inputs are increased. The output elasticity with respect to x1 can be obtained on the basis of above information. It is defined as the proportionate rate of change of q with respect to x_1:

$$\varepsilon_{01} = \frac{\Delta q}{q} \div \frac{\Delta x_1}{x_1} = \frac{x_1 \Delta q}{q \Delta x_1} = \frac{\Delta q}{\Delta x_1} \div \frac{q}{x_1} = \frac{MP}{AP}$$

where ε_{01} = Output elasticity with respect to input x_1

Δq = Change in output and

Δx_1 = Change in input, x_1

6.2.3 Fixed and Variable Proportions in Production

As stated above, different quantities of output can be produced by using a fixed amount of one input and a variable amount of another. As the amount of one is changed, other remaining constant, the ratio of inputs varies. This process is called **production under variable proportions**. In order to understand the underlying idea behind variable proportion, just recall that a factory building and capital equipments (machineries) in it may remain fixed for quite a while. The labour services, on the other hand, are variable in day-to-day production. That means capital-labour ratio varies.

The above type of situation can be contrasted with fixed proportion production. Under it, only one ratio of inputs is available to produce a good. Output expansion or contraction requires variation of all inputs to ensure the condition of fixed ratio.

6.2.4 Three Stages of Production

When the variable input is combined with a fixed one in the production process, three distinct stages appear. Look at Figure 6.3 to identify these stages. In the figure, variable input is measured on the X-axis and total, average and marginal products are plotted on Y-axis.

Stage 1

The first stage corresponds to the use of variable inputs upto point 5 (see Figure 6.3). In this stage the total product increases at an increasing rate initially. MP rises in this range. However, the slope of TP declines after a point (after point 1 in Figure 6.3). As a result, TP increases at a diminishing rate. Consequently the marginal product falls but remains positive. At point 5, average product reaches its maximum. Stage 1 ends there. In stage 1 MP rises first and falls after some level. It remains

above average product. On the other hand, AP continues to rise throughout the stage. This part of production, therefore, is known as **stage of increasing returns**.

Fig. 6.3

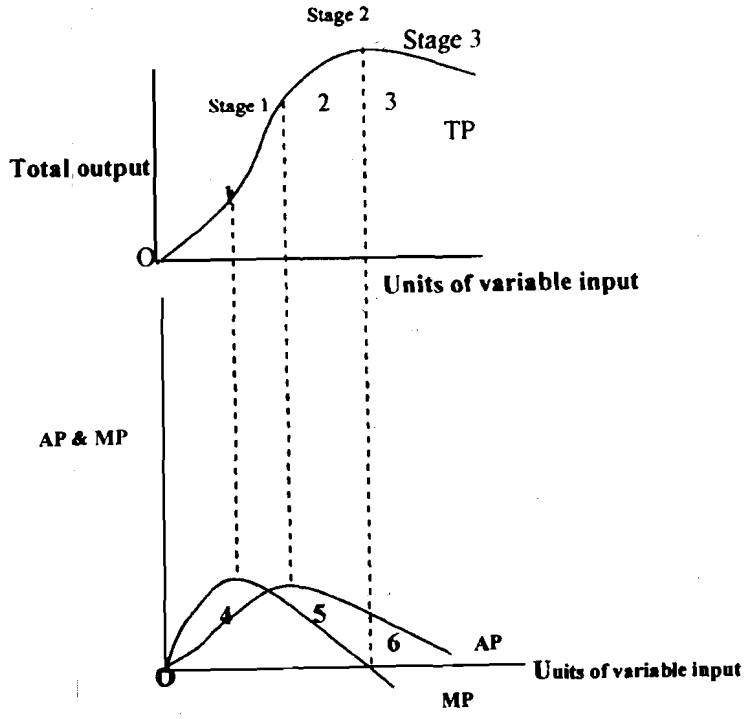


Fig. 6.3 shows the shapes of total, average and marginal product curves. To appreciate the underlying relation between these, think of the figure as consisting of two parts. In the upper part see the total product curve (TP). Its relation with average and marginal product curves (AP and MP) can be seen by coming down to lower part. The three stages of production are evaluated with the help of TP, AP and MP

Stage 2

The second stage corresponds to use of the variable input between points 5 and 6. In this part, TP continues to rise at a diminishing rate. The second stage ends at point 3 when TP reaches the maximum. In this stage both AP and MP are diminishing but are positive. Point 6 gives the condition where MP is equal to zero and indicates the end of stage 2. This is known as the **stage of diminishing returns** as both AP and MP are diminishing.

Stage 3

Stage 3 is indicated by the use of variable input to the right of point 6. TP declines and MP becomes negative at this stage. See that MP curve goes down the X - axis. This stage is called the stage of negative returns. An evaluation of three stages of production discussed above will indicate that a rational producer will never produce at the last stage. Since MP is negative there, the producer will increase the output by reducing the quantity of variable input.

6.2.5 Why Law of Diminishing Returns Operates?

The preceding discussion on stages of production indicated a phase when average and marginal product curves start declining. To understand such a feature see the following discussion on the law of diminishing returns. This principle states that as one variable input is increased, with all others remaining fixed, a point will be reached beyond which the marginal physical product of the variable factor will begin to decrease. To be more specific, take the help of an example. Suppose that paddy is produced by making use of land, labour and a tractor. The law of diminishing returns tells that when the amount of labour is increased holding the use of land and tractor constant, there will eventually be a stage when rate of increase in production of paddy will decrease.

Initially, one worker is employed on the farm of the size 10 acres and is given one tractor to work with. Suppose one more worker is engaged. What type of change has taken place in the relationship between factors of production? Now each worker is getting 5 acres of land and ½ of the tractor to work with. Labour's, output is bound to change. This phenomenon is depicted in the table. When land and capital are kept fixed and labour input is increased and total, average and marginal products can be recorded. With one worker 10 units of output are obtained. When the employment of labour is increased to 8, 64 units of output are produced.

Look at the average product and record the changes taking place. With one worker the average product was 10. When labour was increased to 4, average production went upto 13. This implied an increase in labour employment resulted in higher productivity. But when the number of labourers was increased to 5, the average production came down to 12.2. Further increase in labour inputs pushed down the average production. A similar picture of increasing at the beginning and decreasing afterwards was evident from the column giving marginal product. When 3 units of labour were in employment, the marginal product was the maximum at 15. However, a decline set in after that. When number of workers employed reached 8, the marginal product became negative (-2).

Table 6.2: Diminishing Average and Marginal Products

Units of Labour	Total Output	AP	MP
1	10	10	-
2	24	12	14
3	39	13	15
4	52	13	13
5	61	12.2	9
6	66	11.0	5
7	66	9.4	0
8	64	8	(-)2

6.2.6 Iso-Product Curve and Marginal Rate of Technical Substitution

Take output, Y, to be constant. A functional relation tells us how such a constant level of output can be produced with different combinations of capital and labour. ABCD curve has been drawn from Table 6.3 Each of the A, B, C, D combinations of capital and labour can produce 100 units of output. Between A and B combinations, for producing 100 units of output, 10 units of labour can be substituted for 30 units of capital. Again, the shift from B to C can keep the level of output constant if 10 units of labour replace 20 units of capital. This way 10 units of labour replace 20 units of capital for producing the same level of output for a shift from C to D.

Table 6.3: MRTS of Labour for Capital

Combinations	Output	K	ΔK	L	ΔL	MRTS = $\frac{\Delta K}{\Delta L}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)
A	100	90				
B	100	60	-30	20	10	-3:1
C	100	40	-20	30	10	-2:1
D	100	30	-10	40	10	-1:1

Here it is shown that the replacement of capital with labour can keep the level of output constant only if each additional unit of labour replaces lesser and lesser amount of capital. This is due to the operation of the law of diminishing returns in both the inputs. As capital is reduced and labour is increased two forces operate. First, a reduction of capital increases its returns. Second, since capital is replaced by labour, the increased amount of labour has to work with less capital. Hence, labour's marginal productivity decreases. This is because no two inputs would be perfect substitutes of each other, unless they were absolutely identical. In that case, by definition, they would be the same inputs. Since the marginal productivity of labour decreases and that of capital increases, with the decrease in capital and corresponding increase in labour, the substitution of labour for capital to produce a constant level of output is possible only if every extra unit of labour displaces less and less of capital per unit. The rate, at which labour can replace capital keeping the level of output constant, is the **marginal rate of technical substitution** (MRTS) of labour for capital. MRTS declines as capital decreases and labour increases or labour-capital ratio increases. The diagrammatic presentation of the iso-product line takes the shape of a falling curve, convex to the origin as shown in Figure 6.4. Its properties are the same as that of the usual Indifference curves. This is also called Iso-quant. It is a continuous curve, which assumes that two inputs are continuously divisible and substitutable. The slope of the curve

$= \frac{\Delta K}{\Delta L}$, which is negative.

Since output is constant throughout the curve, the marginal reduction in output with respect to decrease in capital is just equal to the marginal increase in output as a result of increase in labour. The marginal reduction in output due to decrease in capital is $MP_K \times \Delta K$ where MP_K is the marginal productivity of capital.

The marginal increase in output due to increase in labour = $MP_L \times \Delta L$

where MP_L is the marginal productivity of labour. Hence, the change in total output due to substitution of labour for capital is zero. That is,

$$\Delta Y = MP_K \Delta K + MP_L \Delta L = 0$$

$$\text{or } \frac{MP_L}{MP_K} = -\frac{\Delta K}{\Delta L}$$

This means the marginal rate of technical substitution of labour for capital

Fig. 6.4

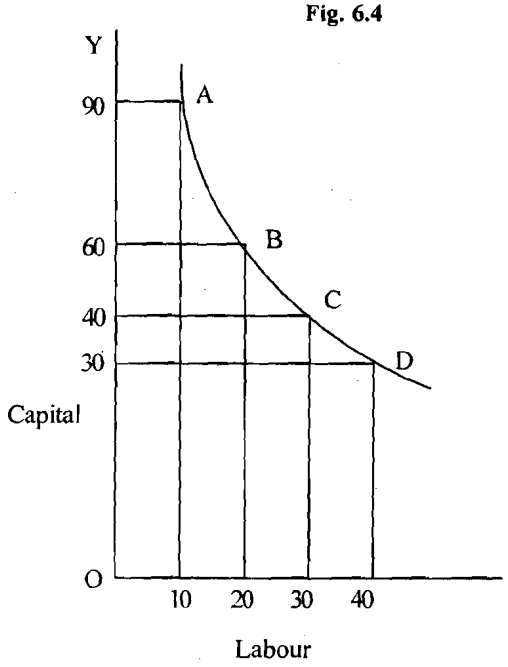


Fig. 6.4 shows that at points A, B, C and D same levels of output are being produced. However, to such a level of output, defferent combinations of capital and labour as given in Table 6.3 are employed

$MRTS_{LK} = (-)\frac{\Delta K}{\Delta L}$

equals the inverse of the ratio of their marginal productivities.

The relation between $MRTS_{LK}$ and the ratio of labour to capital is also deducible from Table 6.3. From this relation the elasticity of substitution can be defined as

$$ES_{LK} = \frac{\text{Percentage change in MRTS}_{LK}}{\text{Percentage change in L/K}}$$

$$= \frac{\Delta MRTS_{LK}}{MRTS_{LK}} \div \frac{\Delta (L/K)}{L/K}$$

$$= \frac{\Delta MRTS}{\Delta (L/K)} \times \frac{L/K}{MRTS}$$

This elasticity is greater than, equal to, or less than one according as a given percentage change of L/K ratio induces greater, equal, or less percentage change in MRTS in the opposite direction.

6.2.7 Economic Region of Production

In production analysis, generally, isoquants are shown as convex to origin. Their slopes are monotonically negative. This shape of isoquants gives them a special feature: these curves represent the input combinations, which are economically efficient. However, a production function need not give us this shape all the times. It may have an upward sloping segment as well. But that segment will not be economically efficient. We can draw two curves, OC and OL, which separate efficient and inefficient segments of the isoquants. These curves enclose what may be called 'economic production region'. Note that OC passes through all the points on the isoquants are horizontal, or their slopes are equal to zero. (Fig. 6.5)

Fig. 6.5

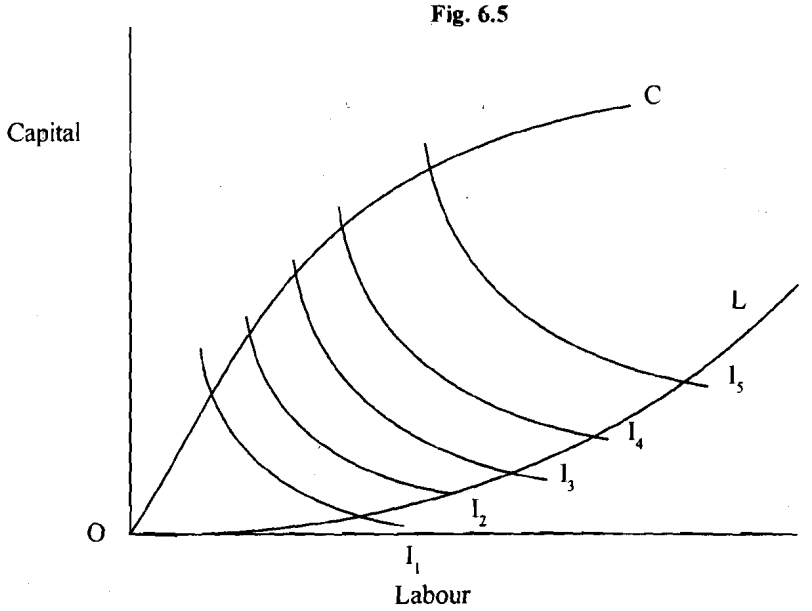


Fig. 6.5 shows that the economic production region is enclosed by Ridge Lines OC and OL. The line OC is the locus of points where Iso-quants are vertical. Similarly, OL is the locus of points on which Iso-quants are horizontal.

6.2.8 Some Specific Forms of Production Function

Some specific forms of production function allowing different degrees of substitutability between inputs are given below:

a) **Fixed Technique of Production and Iso-Product Curve**

Fig. 6.6

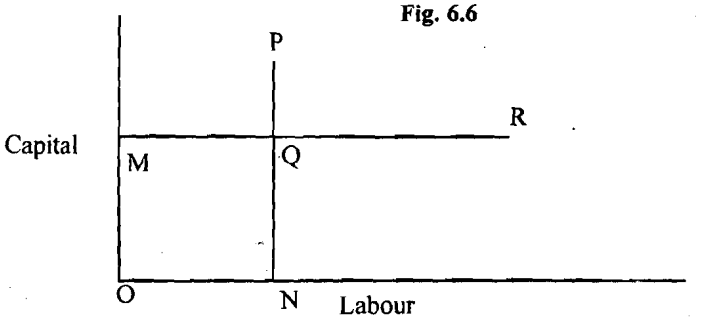


Fig. 6.6 is drawn by assuming that capital and labour are not substitutable and have to be used in fixed proportions. Thus, when the technique of production is fixed, the shape of the Iso-product curve is like PQR. The Iso-product curve is flat running parallel along the capital and labour.

When the technique of production is fixed, the shape of labour and capital axes would show non-substitutability between these two inputs. If capital or labour is used in excess of OM or ON amount respectively, their marginal productivity will be zero.

b) Homogeneous Production Function: Degree One

When continuous substitutability is allowed between these inputs, the production function can be homogeneous or non-homogeneous. In case of the homogeneous production function, the level of production can be increased by using the same technique of production, i.e., by increasing both the inputs in the same proportions. This way of expanding production gives the returns to scale. The scale here refers to change in inputs in the same proportion.

Thus, production function, homogenous of degree one is described by terms such as linear homogenous or constant returns to scale. All these drive home the essential concept: if all inputs are expanded in the same proportion, output is expanded in that proportion. To understand the underlying idea take an example.

Consider a production function:

$$q = f(x, y) = Ax^{\alpha} y^{1-\alpha}$$

where x,y = two inputs (say, labour and capital)

and $0 < \alpha < 1$

Now let x and y be increased by proportion λ. Then we write f(x,y) as

$$f(\lambda x, \lambda y) = A(\lambda x)^{\alpha} (\lambda y)^{1-\alpha} = A\lambda^{\alpha} x^{\alpha} \lambda^{1-\alpha} y^{1-\alpha}$$

$$= \lambda^{\alpha+1-\alpha} A x^{\alpha} y^{1-\alpha}$$

$$= \lambda f(x, y) = \lambda q$$

Therefore, the expansion of inputs by λ has resulted in output expansion by λ. This is what is meant by a constant returns to scale. Such an idea is also presented with the help of a diagramme in the following.

Fig. 6.7

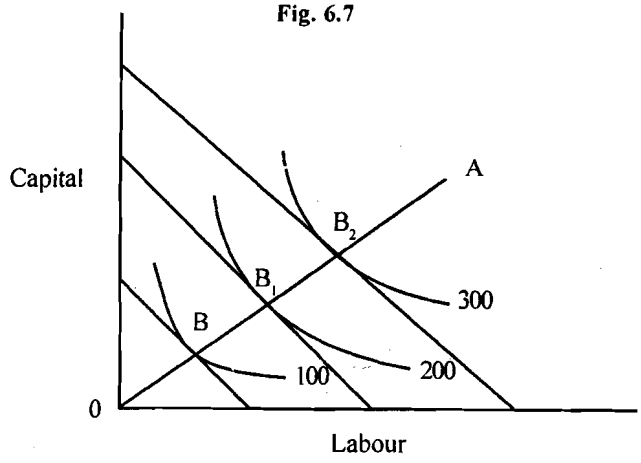


Fig. 6.7: In figure 6.7, parallel Iso-product curves are drawn. A higher Iso-product curve using more of two inputs represents a higher level of output (say, 200 units). A ray through the origin (OA), crossing the Iso-product curves at B,B¹ and B" is known as the scale line or expansion path of output.

The parallel Iso-product curves in the figure indicate that a higher level of output is produced when greater amounts of the inputs, capital and labour are employed. It also shows that a higher point on the scale line corresponds to a higher Iso-product curve and indicates a higher level of output. If Iso-product curves are parallel, and the same marginal rates of substitution between inputs prevails on all the points of intersection of the Iso-product curves and the rays through the origin, the production function related to these Iso-Product curves are homogenous. In the figure slopes at B, B' and B", i.e., points of intersection between OA line and Iso-product curves, are the same. These Iso-product curves are a set of parallel convex curves, like the Indifference Curves. Then, production can be increased by increasing the inputs by the same proportion and the scale line is a straight line having a constant slope. When the production function is linear and homogeneous, as in Figure 6.7, proportionate increase in output is just equal to the proportionate increase in inputs i.e., OB = BB' = B'B". While the scale line is a straight line the proportional increase in output can be higher than that in inputs and can also be less than that in the inputs. In that case, production function is not linear, though homogeneous.

c) Non-Homogeneous Production Function and Scale Line

When production function is non-homogeneous, the increase in production is not proportional to the increase in the inputs, at the given technique of production. In Figure 6.8, OA is an upward sloping curve through the origin showing that its slope increases as output increases. This implies that capital intensity of the technique of production, i.e., the ratio of capital to labour rises as production increases. It is possible to have a reverse situation where the slope of the expansion curve will decrease and capital intensity will fall with increase in output.

Fig. 6.8

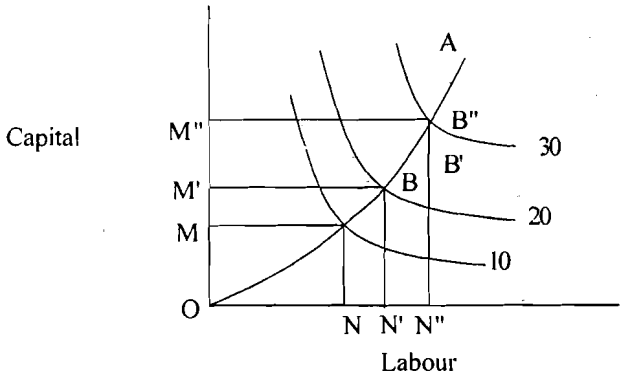


Fig. 6.8 shows that the slopes of the Iso-product curves are the same on the expansion curve OA.

Check Your Progress 1

1) Indicate the following statements as true(T) or false(F)

i) The law of diminishing marginal production states that increased application of one input keeping others fixed will yield smaller output. ( )

ii) In production under variable proportions the ratio of all variable inputs varies. ( )

iii) Marginal rate of technical substitution of labour for capital increases as capital decreases and labour increases. ( )

iv) One unit increase in output due to an extra unit of increment in input is known as marginal product. ( )

v) In case of homogenous production function production can be increased by increasing the inputs in the same proportion. ( )

2) Discuss the three stages of production . Explain why the law of diminishing returns operates?

.....

.....

.....

3) Write a short note on economic region of production.

.....

.....

.....

6.3 INPUT PRICES AND ISO-COST LINE

The technology of production discussed so far offers the possibilities of combining different techniques of production. But the choice of a technique out of the technological options is based on input prices. A firm chooses the technique, which produces maximum output with minimum cost of production. Here, the average cost of production, i.e., total cost divided by total product is minimum. The choice of technique will depend on the highest possible Iso-product curve, which can be reached by a given Iso-cost curve deduced on the basis of input prices and total outlays.

6.3.1 Iso-cost Line

Let the price of capital be r and of labour be w.

Then Total Outlay = rK + wL,

Obviously, $\frac{w}{r}$ is the exchange ratio between labour and capital. Exactly as in the consumer budget line, the intercept on the capital axis (Y) represents the number of units of capital that the total outlay can purchase if it were spent on capital only. Similarly the intercept on the labour (X-axis) represents the number of units of labour that the total outlay can purchase if it were spent entirely on labour. Thus,

$\frac{w}{r}$

i.e., the ratio of prices of the two inputs is the slope of the line, which is negative and constant. The slope will remain the same as long as the relative price ratio of the inputs is constant. Intercepts change as total outlay is changed. In Figure 6.9, 'WT' line is drawn on basis of the above equation to show that the total outlay remains constant over different combinations of two inputs. This is the Iso-cost line, very similar to the budget line of the consumer buying two goods. This, in fact, is the producer's budget, i.e., outlay line for the inputs.

Now, the choice of technique is determined where the Iso-cost line is tangent to the Iso-product curve at point R. This is the maximum output the outlay can produce. If the technique represented by the ray OA is used, capital-labour ratio is OM/ON. At point R, the slope of the Iso-cost curve (w/r) is equal to the slope of the Iso-product curve. Hence, at this point the input-price ratio, the marginal rate of technical substitution and the ratio of marginal productivities of labour and capital inputs are equal. This is similar to consumer equilibrium.

Fig. 6.9

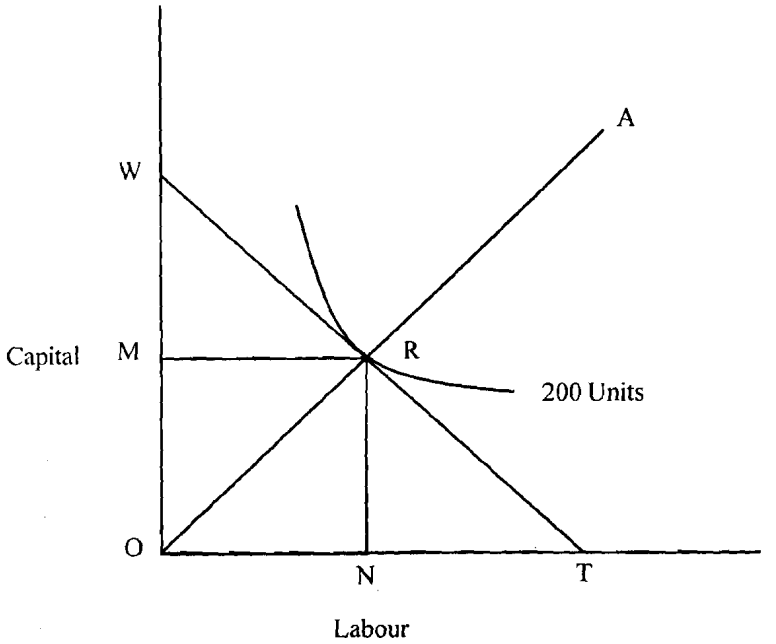


Fig. 6.9 shows that OW amount of capital is hired to spend the entire outlay. Similarly, by spending that outlay, OT units of labour can be hired. The straight line WT indicates various combinations of capital and labour, which a producer can employ for the given outlay. It is the Iso-cost line. The Iso-quant (200 units of output) is tangent to WT at R.

6.3.2 Returns to Scale : A Diagrammatic Presentation

We have already established that returns to the scale need not be constant. It will be interesting to examine how to present increasing, constant and diminishing returns to scale on a single diagram. For that purpose you may look at Fig. 6.10.

If we draw a straight line OA through the origin it will cut all the isoquants. This line OA is called the 'expansion path'. Let the first isoquant shows 100 units of output, second 200, third 300 and so on. The gap between two successive curves as shown along the line OA can give us very important clues about the returns to scale. If the curves are at a constant distance from none another, they show constant returns to scale. If on the other hand, this gap is increasing, diminishing returns would prevail and narrowing down of the gap would point to increasing returns.

Consider Fig. 6.10. It has 6 isoquants. The first one represents 100 units of output and the last denotes 600. Gap between I_1 and I_2 is equal to that between I_1 and I_3. This equal gap points towards equal change in the utilisation of labour and capital. Use of labour has changed from OB to OC to OD. But note that BC is equal to CD. A similar trait is observed along the vertical axis, which shows changes in use of capital. Thus, equal increase in both the inputs leads to equal rise in the output

Fig. 6.10

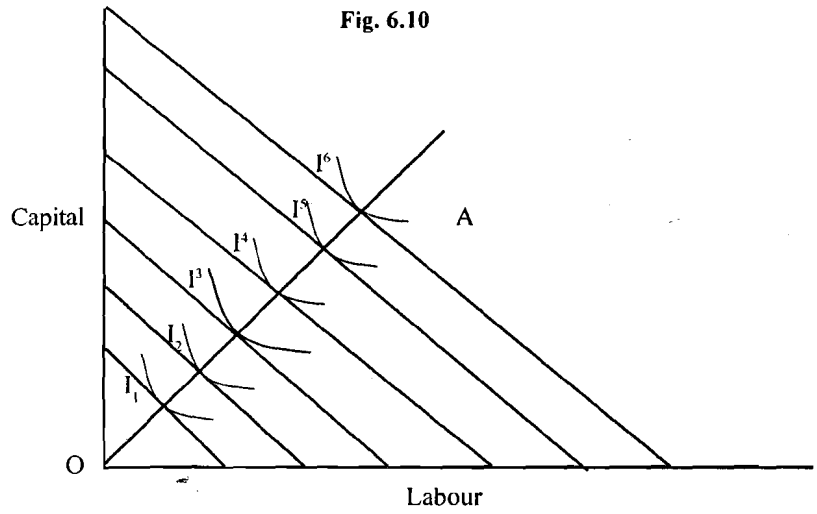


Fig.6.10 shows that change in inputs, capital and labour, does not lead to proportionate change in output.

However, I_4 and I_5 have drifted farther and farther. That means, larger and larger amounts of both the factors are being used but the rise in output remains 'fixed' at the old level of 100 units. Here the returns to scale are diminishing - the output rises at a rate, which is slower than the rate of rise of the use of inputs. Finally, we find that I_6 has come closer to I_5 yet increase in the output remains 100 units - in other words, smaller increase in the inputs is able to increase the output by 100 units now. This is the case of increasing return to scale.

Check Your Progress 2

1) Define an Iso-product curve in one sentence. What are shown along the axes of the figure to draw an Iso-product curve?

.....

.....

.....

2) If, for a particular combination of labour and capital, the marginal productivity of capital is 6 units of output and the marginal rate of technical substitution is 2 units of capital per unit of labour, calculate the marginal productivity of labour.

.....

.....

.....

3) Let the price of capital be Rs. 60 and that of labour Rs.50. Write the Iso-cost equation for the outlay of Rs.2000 and calculate the slope of the Iso-cost curve.

.....

.....

.....

.....

6.4 LET US SUM UP

In this unit, the concept of production function has been explained. It tells that inputs and outputs are related technically. The idea of Iso-product curve has been introduced to explain how a constant level of output can be produced by using two variable inputs like capital and labour in different proportions. The Iso-product curve is downward sloping. It is also convex to the origin just because the two inputs, though substitutable, are not perfect substitutes of each other. This curve is deduced on the assumption that the marginal rate of technical substitution of labour for capital declines with the increase in the proportion of labour to capital. This curve is used to analyse homogeneous, linear homogeneous and non-homogeneous production functions. An analysis of production function helps us to deduce the expansion path, and to explain the scale line under various returns to scale. In the homogeneous production function, the technique of production remains the same with the expansion of the scale of operation.

6.5 KEY WORDS

Constant Returns to Scale : If all factors of production are increased in a given proportion, the output increases in exactly the same proportion under constant returns to scale.

Homogeneous Production Function : If a set of Iso-product curves generated by a production function has the same slope along a ray through the origin, then the production function is called homogeneous.

Decreasing Returns to Scale : Under this law of production, output increases proportionately less than inputs.

Expansion Path : This is a locus of cost minimising proportions of two inputs to produce maximum output resulting from changes in the scale of outlays, keeping the input prices constant.

Increasing Returns to Scale : Under this law of return, output increases proportionately more than inputs.

Iso-Cost Line : This shows all the different combinations of two inputs that a firm can purchase or hire with given input prices and outlay.

Iso-Product Curve : This shows all the different combinations of two inputs that a firm can use to produce a specified quantity of output.

Non-Homogeneous Production Function : When a production function generates a set of Iso-product curves having different slopes along a ray through the origin, it is termed as non-homogeneous production function.

Production Function : This is a technical relation showing the maximum quantity of output that can be produced from a set of input combinations on the basis of existing technology.

Short Run : The time period when at least one of the inputs (plant size or building) is fixed.

6.6 SOME USEFUL BOOKS

Koutsoyiannis, A. 1979. *Modern Microeconomics*, Macmillan; New York Chapters 3, 4 & 5. pp. 67 to 163.

Salvatore, D. 1983. *Microeconomics Theory*, Schaum's Outline Series, Chapters 7, 8 and 10, pp. 124 to 174 and 196 to 220.

Samuelson, Paul A. and W. D. Nordhaus, 1985. *Economies*, McGraw- Hill Book Company, Chapters 21 and 22, pp. 461 to 501.

Lipsey, Richard G. 1979. *An introduction to Positive Economics*, English Language Book Society/Weidenfold and Nicoloon, Chapters 16, 17 18 and 19, pp. 201 to 259.

6.7 ANSWERS OR HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) i) (T); ii) (F); iii) (F); iv) (F); v) (T)

2) Include in your answer the Sub-section 6.2.4

3) See Sub-section 6.2.7

Check Your Progress 2

1) The curve shows a constant level of output produced by different combinations of two inputs such as capital and labour. Capital(K) is measured on the Y-axis and Labour (L) on the X-axis.

2) The Marginal Rate of Technical Substitution:

$$= -\frac{\text{Marginal Product of Labour}}{\text{Marginal Product of Capital}} = -\frac{MP_L}{MP_K}$$

$$\text{Or, } -2 = \frac{MP_L}{6}, \text{ or } MP_L = 12$$

3) Iso-cost equations:

Total outlay = Price of capital × quantity capital + Price of labour × quantity of labour

Or, Rs.2000 = Rs.120 × quantity of capital + Rs. 50 × quantity of labour

$$\text{Slope of the Iso - cost curve} = -\frac{\text{Price of labour}}{\text{Price of capital}}$$

$$= -\frac{50}{120} = -\frac{5}{12}$$

UNIT 7 THEORY OF COSTS

Structure

7.0 OBJECTIVES

After going through this unit you will be in a position to:

- distinguish between various types of costs;
- represent these cost concepts graphically;
- identify the relationship between various types of costs;
- distinguish between long run and short run costs; and
- identify the relationship between long run and short run cost curves.

7.1 INTRODUCTION

In our day-to-day life we come across the term cost in various contexts. The amount of money we pay for the purchase of a good or hiring a service is generally considered as cost to us. But for a production unit or firm cost means the amount paid towards purchase of inputs used in the production of good or service. We have seen in the beginning of this course that there are four factors of production, viz., land, labour, capital and organisation. Out of these labour (L) and capital (K) are two primary inputs for production (see Unit 6). Payment made for the use of these inputs is cost. Since the amount of output produced depends upon level of inputs the cost of production depends upon the level of output.

If we do not specify properly, the definition of cost may be ambiguous. Therefore, the concept of cost will be taken up in the next section.

7.2 COSTS IN THE SHORT RUN

By now you must be familiar with the term short run. You use the term for a period of time when at least one of the inputs cannot be varied. Examples of fixed inputs in the short run are machines, building, etc.

7.2.1 Fixed Cost

Let us consider a production process using two inputs K and L only. Here the level of capital input in the form of machines and tools cannot be changed in the short run while amount of labour can be. The machines once purchased by the firm cannot be disposed off easily. Similarly, the size of the building used for production and godown cannot be changed as and when wished. The firm would require a longer period of time to change such inputs.

The cost incurred on all fixed inputs by the firm is *total fixed cost or simply fixed cost (FC)*. Note that whatever be the level of output produced, the fixed cost remains same. Thus fixed cost is constant for a firm. If the amount of capital used by the firm is K_0 and its rental price is r, then

$$FC = r K_0 \qquad \ldots(7.1)$$

The fixed cost is also called the *overhead cost*. When represented in a graph FC curve is a horizontal straight line as given in Fig. 7.1. The figure indicates that fixed cost remains constant for all levels of output.

Fig. 7.1

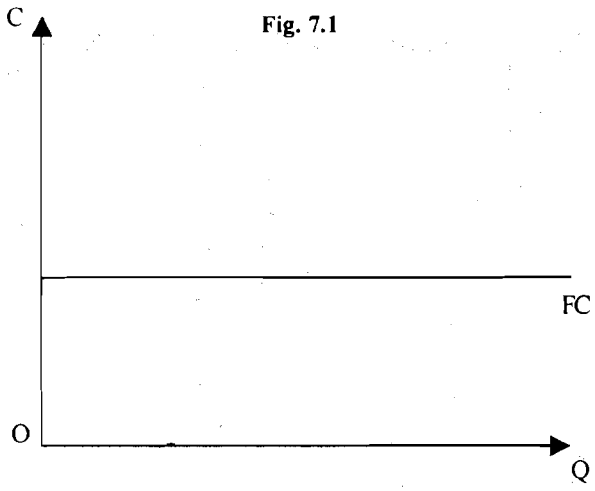


Fig. 7.1: Horizontal axis shows output in number of units. The vertical axis shows fixed cost in rupees. This figure depicts the information contained in 1st and 2nd columns of Table 7.1.

7.2.2 Variable Cost

Variable Cost (VC), on the other hand, is the cost incurred on variable inputs. As mentioned above labour is one example of variable input. To appreciate the idea of variable input consider the examples like: given the machines in a factory producing shoes an increase in output can be brought about by increasing the amount of labour or the quantity of raw materials.

Often, labour use of increases with increase in the level of output. Accordingly, the cost of hiring labour also increases. Hence, we see that *variable cost depends upon the level of output*. If labour is paid a wage rate w and L1 amount of labour is used for producing Q_1 units of output, then

$$VC(Q_1) = w L_1 \qquad \ldots(7.2)$$

At this point you may be wondering about the shape of the VC curve. Is it a straight line or curve? Does it slope upward or downward? To answer these questions we have to visualise the shape of the total product (TP) curve discussed in the previous unit.

You know from the previous unit that there are two types of increase in the TP curve. In the beginning it increases at an increasing rate and then slows down to increase at a decreasing rate. Since fixed cost remains the same in the short run production process, the increase in cost is entirely due to increase in variable cost. The shape of the VC curve reflects exactly the opposite behaviour of TP curve. When TP curve increases at an increasing rate VC curve increases at a decreasing rate and vice versa. Thus, variable cost increases at a slower pace in the beginning. In the second stage, after the point of inflexion, the TP curve shows diminishing returns. During this stage the variable cost will increase at a faster pace.

Fig. 7.2

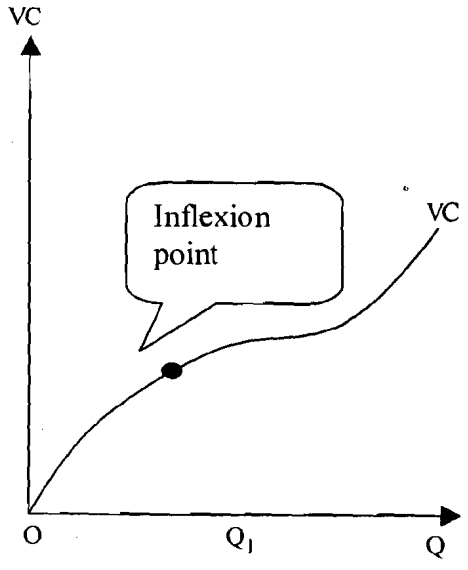


Fig. 7.2 shows relation between variable cost and level of output. Point Q_1 on the horizontal axis corresponds with the level of output 5 in Table 7.1.

Such a behavior can be seen from Table 7.1. Upto 5 units of output VC increases at a decreasing rate. Afterwards it increases at an increasing rate. In Fig. 7.2 Q_1 level of output is the dividing line between the two stages.

Table 7.1: Output and Costs

Q	FC	VC	TC	AFC	AVC	ATC	MC
0	128	0	128	∞	-	∞	-
1	128	56	184	128	56	184	56
2	128	90	218	64	45	109	34
3	128	108	236	42.7	36	78.7	18
4	128	116	244	32	29	61	8
5	128	120	248	25.6	24	49.6	4
6	128	126	254	21.3	21	42.3	6
7	128	140	268	18.3	20	38.3	14
8	128	168	296	16	21	37	28
9	128	216	344	14.22	24	38.22	48
10	128	290	418	12.8	29	41.8	74
11	128	390	618	11.6	35.5	56.1	100

7.2.3 Total Cost

Total cost is the sum of fixed cost and variable cost. Therefore, we can write,

$$TC(Q_1) = FC + VC(Q_1)$$

Total cost consists of two components FC and VC. Out of these FC is constant. But VC depends on level of output. Hence, TC also depends on level of output. For a higher level of output TC is higher.

Fig. 7.3 shows the shape of TC curve. There is an equal gap between TC and VC curves as can be seen from the figure. The difference is equal to the level of fixed cost, which is constant for all levels of output.

Fig. 7.3

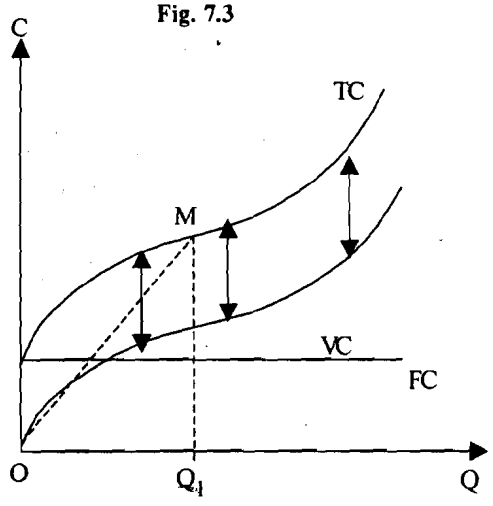


Fig.7.3: This figure brings together information of Fig. 7.1 and Fig. 7.2. We obtain TC curve by vertically adding up FC and VC.

7.2.4 Average Fixed Cost

Average fixed cost (AFC) is defined as fixed cost divided by level of output. Symbolically,

$$AFC(Q_1) = \frac{FC}{Q_1}$$

As you know FC is constant for any level of output. Hence AFC declines as Q increases. Such a feature can be seen from Table 7.1 also. AFC is 128 when Q = 1 and 25.6 when Q= 5. See that while FC is constant for all levels of output AFC varies according to level of output.

Let us look into the graphical representation of AFC curve in Fig. 7.4. It has the shape of a *rectangular hyperbola*. It does not touch either axis although moves quite closer. For zero output AFC is infinity. On the other hand for a large quantity of output AFC is negligible.

Fig. 7.4

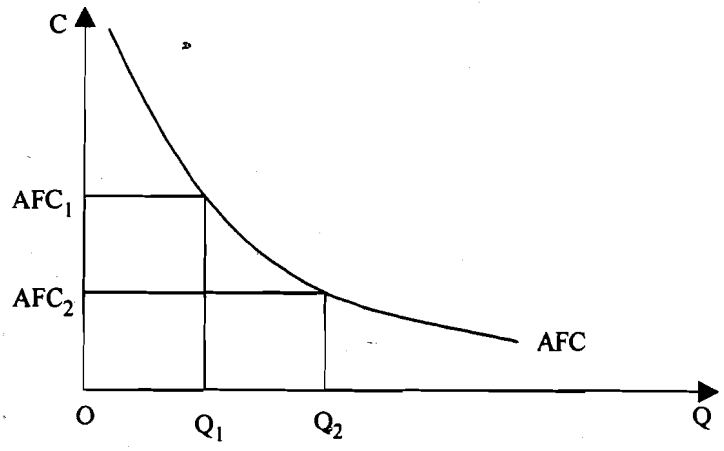


Fig.7.4 shows average fixed cost curve of a firm. Area under the curve is equal to level of output (Q) multiplied by AFC and therefore equals total fixed cost.

7.2.5 Average Variable Cost (AVC)

It is the variable cost divided by units of output, viz.,

$$AVC(Q_1) = \frac{VC}{Q_1}$$

Let us see the graphical presentation of AVC. In Fig.7.5 the upper panel shows the VC curve. In this curve AVC will be $\frac{MQ_1}{OQ_1}$ at output level Q_1. Similarly at output levels Q_2 and Q_3 AVC are $\frac{NQ_2}{OQ_2}$ and $\frac{PQ_3}{OQ_3}$ respectively. These levels of AVCs are shown directly in the lower panel of Fig. 7.5.

Fig. 7.5

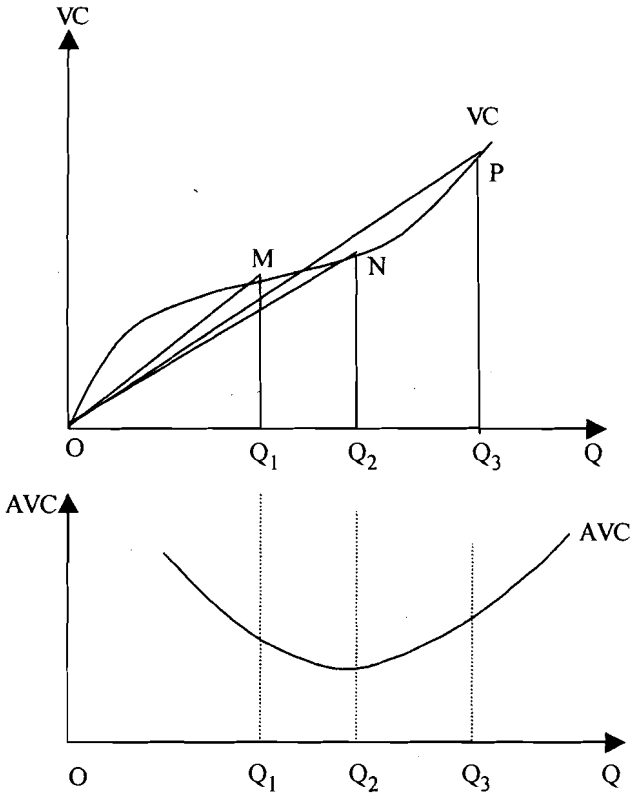


Fig. 7.5 shows derivation of AVC from VC. At each level of output, AVC is slope of a straight line joining origin with VC at that level of output.

Again from the upper panel of Fig. 7.5, we notice that the ratio $\frac{MQ_1}{OQ_2}$ (=AVC at Q_1) is equal to the slope of the line OM. Similarly for output Q_2 the AVC is equal to $\frac{NQ_2}{OQ_2}$ which is equal to the slope of the line ON. Thus, we can find out the AVC at any level of output by measuring the slope of the line from the origin to the corresponding point on the VC curve.

Notice another feature of AVC. The slope at output Q_1 is more than the slope at Q_2. Hence AVC is higher at Q_1 than at Q_2. On the other hand, AVC is less at Q_2 than at Q_3. Thus, it is obvious that AVC declines in the beginning, reaches a minimum and then starts increasing.

7.2.6 Average Total Cost

Average total cost (ATC) is the sum of AFC and AVC at the level of output under consideration. Suppose at the level Q_1,

$$ATC(Q_1) = AFC(Q_1) + AVC(Q_1)$$

As AFC and AVC both depend upon the level of output, ATC also changes according to output level. Fig. 7.6 shows the ATC curve, which is obtained by vertically adding the AFC and AVC.

Fig. 7.6

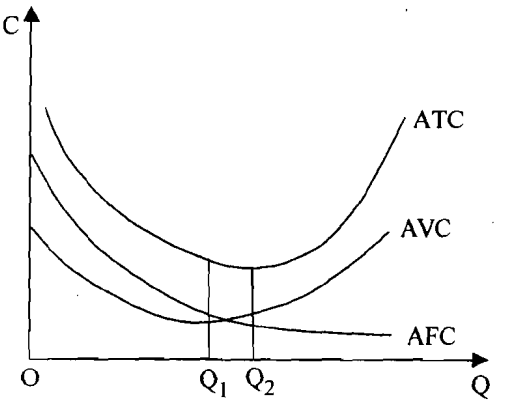


Fig.7.6: We add up AVC and AFC at each level of output to obtain ATC. This ATC is also called average cost of production, or simply AC. Note the U-shape of AC.

Just we derived the AVC curve from the VC curve by taking the slope of the line from the origin, we can derive ATC also in a similar way from the TC curve. Refer back to Fig.7.3 and Fig. 7.5. The ATC at output level Q_1 is given by $\frac{MQ_1}{OQ_1}$ which is the slope of the line OM from the origin. Slope of TC can be measured at all levels of output in a similar manner. This will provide us with the ATC curve and one is plotted in Fig. 7.6.

Two observations can be made from ATC curve above. First, it is U-shaped. It declines in the beginning, reaches a minimum and then increases. The minimum points of ATC and AVC need not be at the same level of output (see Fig. 7.6). Minimum level of ATC is at a higher level output compared to minimum point of AVC. This feature can be observed from Table 7.1. While the minimum point of AVC is at Q = 7, it is at Q = 8 in case of ATC. So, for some levels of output (the output range Q1Q2 in Fig. 7.6) ATC declines even though AVC is increasing. This is due to the fact that decline in AFC *outweighs* rise in AVC.

7.2.7 Marginal Cost

Marginal Cost (MC) is the increase in total cost due to production of an additional unit of output. Symbolically,

$$MC(Q_1) = \frac{\Delta TC}{\Delta Q}$$

In Fig. 7.7 we depict the change in TC as output increases from Q_1 to Q_2. With the increases of output from Q_1 to Q_2 (= ΔQ) we see that TC increases from MQ_1 to NQ_2 (=ΔC). Thus, at Q_1, $MC = \frac{\Delta TC}{\Delta Q}$ $(= \frac{PN}{MP})$. Similarly we can find out MC for any change in output level. But here MC is over a discrete change in output, not at a point of output. Hence in order to find out MC at any point on the TC curve, we have to consider an infinitesimal change in output.

Those of you who are familiar with calculus may say that we can obtain such a measure of MC by taking limit value of $\frac{\Delta TC}{\Delta Q}$. Thus MC can be found out by taking partial derivative of TC with respect to Q. Symbolically,

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

You have seen earlier that FC is constant and does not change if output increases by one unit. The increase in TC for production of an additional unit of output is entirely due to increase in VC. Therefore, MC = ΔVC/ΔQ. Going by the derivative formula we can define

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

Fig. 7.7

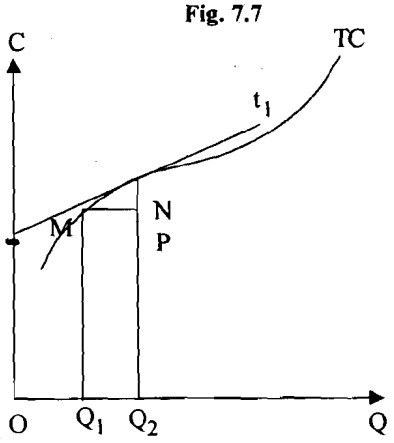


Fig. 7.7: Emphasises the point that slope of total cost curve at each level of output can be interpreted as marginal cost at that output.

Now, we will go a step further. You know that the MC is equal to $\frac{\Delta TC}{\Delta Q}$. At point Q_1 we found out that $MC = \frac{PN}{MP}$. But this is equal to the slope of the *tangent* line t_1 drawn on the TC curve at point Q_1.

Let us go back to Table 7.1. In the table see that when output increases from 0 to 1 unit TC increases from Rs.128 to Rs.184, that is, by Rs.56. Hence MC is Rs. 56. But when output increases from 1 unit to 2 units then TC increases from Rs. 184 to Rs. 218, i.e., by Rs. 34. Hence MC is Rs. 34 at this level of output. Similarly we can calculate MC at all levels of output.

Check Your Progress 1

1) Define the following concepts.

   i) Variable Cost

   .....

   .....

   ii) Marginal Cost

   .....

   .....

   iii) Average Fixed Cost

   .....

   .....

7.3 RELATIONSHIP BETWEEN AVERAGE COST AND MARGINAL COST CURVES

By now you are familiar with the shapes of cost curves. Let us look into the relationships between them.

Take the relationship between ATC and MC first. From the previous section we know that ATC is the slope of the line from the origin to the point on the TC curve. On the other hand, MC is the slope of the tangent to the TC curve on the point concerned. For the output level Q_1 (see Fig. 7.8), the corresponding level of TC is Q_1M. On the TC curve consider the point M. At this point ATC is given by the slope of the line OM. But MC at this point is slope of the tangent t_1. The slope of the line OM is greater than the slope of t_1. Hence ATC>MC at level of output Q_1.

Now consider the level of output Q_2. At this level the corresponding point on the TC curve is N. At N, ATC is the slope of the line from the origin OR while MC is slope of the tangent at point N, which is equal to OR. Thus we find that ATC is equal to MC at Q_2.

Fig. 7.8

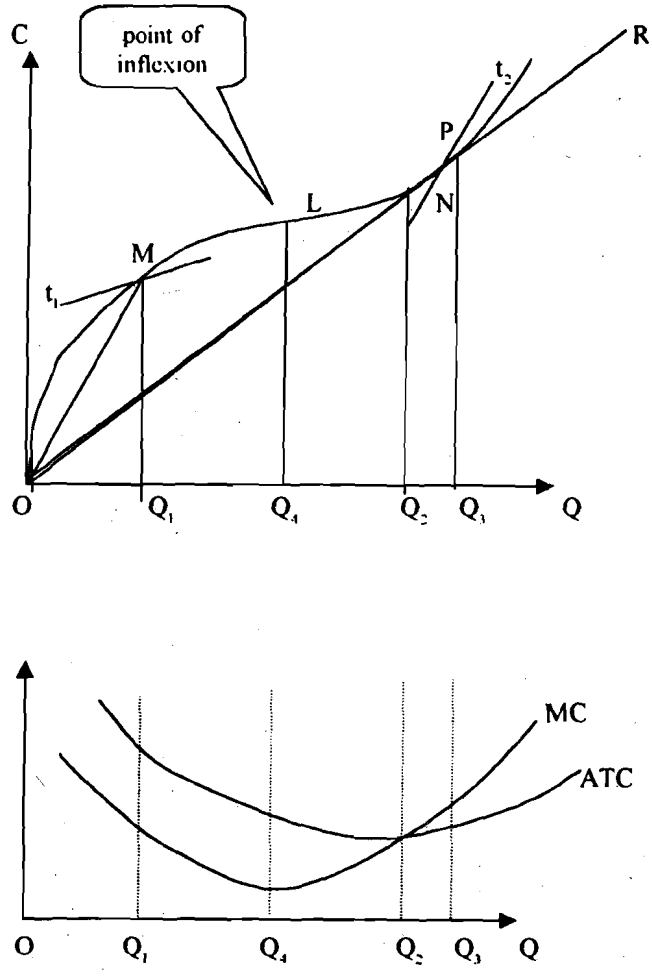


Fig. 7.8: is quitean interesting figure. We go through steps involved in Fig. 7.5 to get ATC. Then, we plot slopes of total cost curve at all those points. The resulting curve, as explained in Fig. 7.7 above is MC. Also note that MC has a shape similar to AC and cuts AC at its minimum point.

Take the level of output level Q_3. The corresponding point on the TC is P. At P we see that slope of OP is less than the slope of t_2. Thus ATC<MC.

Consider the point L. Here the slope of the tangent is zero and MC reaches a minimum at the corresponding level of output (Q_4). In fact, before Q_4 level of output, the slope of TC is negative. Hence MC is declining. On the other hand, after Q_4 level of output, slope is positive. Hence MC is increasing. At Q_4, the TC curve has a point of inflexion.

In the beginning ATC>MC (up to the level of output Q_2). At Q_2, ATC = MC. After Q_2, MC>ATC. Observe the relationship which has emerged. Note another fact that at Q_2 level of output ATC is the minimum. Thus ATC = MC at the level of output where ATC is minimum. Below this level of output, MC is lower than ATC and above this level MC is greater than ATC. Hence, MC crosses ATC from below at the minimum level of ATC.

Check Your Progress 2

1) For a cost function $TC = 53 + 0.5\,Q + 0.2\,Q^2$ find out

   i) Fixed Cost

   .....

   .....

   ii) Average Variable Cost

   .....

   .....

   iii) Marginal Cost

   .....

   .....

7.4 RELATIONSHIP BETWEEN COST AND PRODUCT CURVES

There is a close relationship between production and cost. From the theory of production discussed in the previous unit we have the following definitions in the short run (where labour is the only variable input and capital is fixed) for average and marginal product:

$$AP = \frac{TP}{L} \text{ and } MP = \frac{\Delta Q}{\Delta L}$$

In the short run, we know that $MC = \frac{\Delta TC}{\Delta Q}$. This can be represented as

$$MC = \frac{\Delta (wL)}{\Delta Q} = \frac{w \Delta L}{\Delta Q} \text{ (since w is constant)} = \frac{w'}{MP_L}$$

where $\frac{\Delta L}{\Delta Q} = \frac{1}{MP_L}$. Thus, MC will be minimum where MP_L is maximum. Higher will

be MC where MP_L is lower. Hence MC curve will be the inverse of the MP_L curve discussed in the previous unit. Similarly from the AP_L curve we can derive the AVC curve:

$AVC = \frac{VC}{Q} = \frac{wL}{Q} = \frac{w}{AP_L}$ where $\frac{L}{Q} = \frac{1}{AP_L}$. Thus AVC is the inversely related to AP_L. When AP_L is maximum AVC is minimum and vice versa.

Fig. 7.9

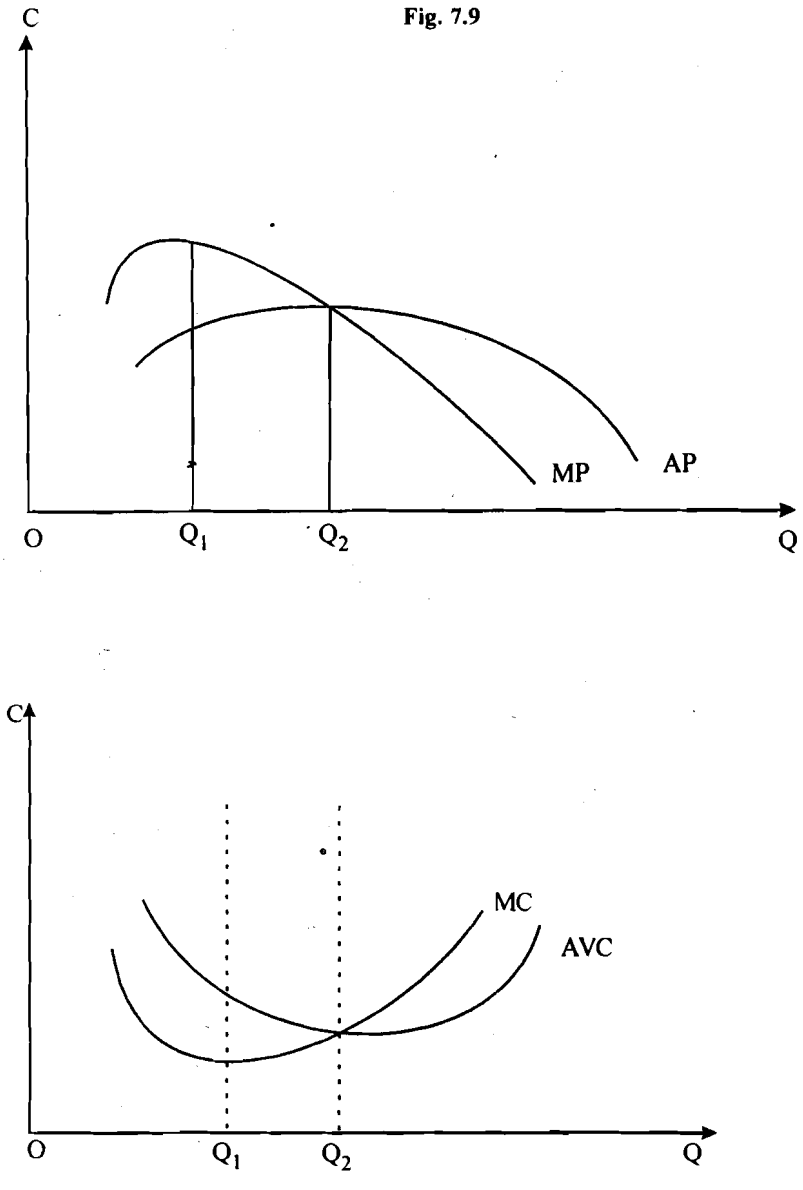


Fig. 7.9: MC and AVC curves are drawn as mirror images of MP and AP respectively. This is done because there exists an inverse relationship between these two curves.

In the previous unit we have seen that MP_L cuts AP_L from above at the maximum point of AP_L. In the upper panel of Fig. 7.9, we see that at output level Q_2, MP_L curve crosses AP_L from above. Thus at Q_2, $AP_L = MP_L$. But the maximum point of MP_L is at output level Q_1 that is less than Q_2. The minimum point of MC corresponds to the maximum point of MP_L, which is at Q_1. On the other hand the minimum point of AVC corresponds to maximum point of AP_L which is at output level Q_2. Thus in the lower panel of Fig. 7.9 we see that MC crosses AVC from

below at the minimum point of AVC. Recall that data in Table 7.1 gave us similar shapes of AC and MC in Fig. 7.8 above. MC did cut ATC at Q_2 level of output and at that point ATC was minimum. MC cuts both AVC and ATC at their minimum points.

Check Your Progress 3

1) Explain the relationship between average product and average variable cost curves.

.....

.....

2) Explain how MC is the inverse of MP.

.....

.....

7.5 COSTS IN THE LONG RUN

In the long run all inputs are variable. The firm can easily liquidate all the capital assets in the long run. Thus no cost remains fixed in the long run. There is only variable cost, which is equal to total cost. This change affects the shape of cost curves in the long run.

The long run total cost curve (LTC) starts from the origin. Recall that short run TC starts from the vertical axis. In Fig. 7.10 we have drawn an LTC curve. As in the case of short run cost curves we can draw long run average total cost curve (LAC) and long run marginal cost curve (LMC).

Fig. 7.9

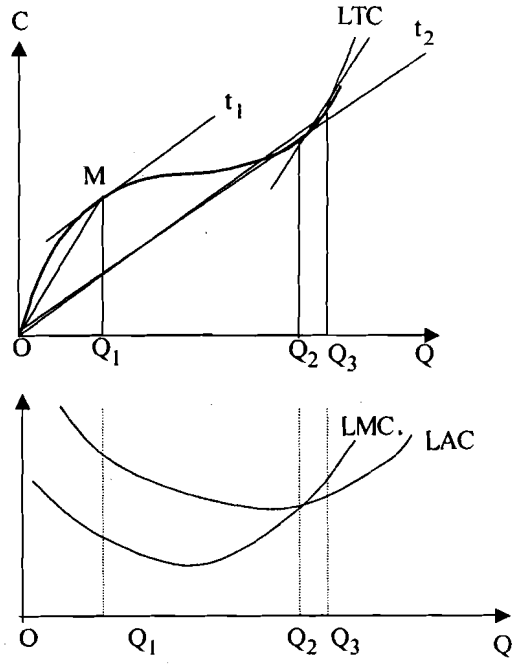


Fig. 7.10: Long run total cost curve is nothing but long run variable cost. Following the methodology of Fig. 7.8 above, we have drawn LAC and LMC in this figure. Again notice, both LAC and LMC are U-shaped but this U is rather wider than the U in Fig. 7.8. Also note that LMC cuts LAC at the latter's minimum point.

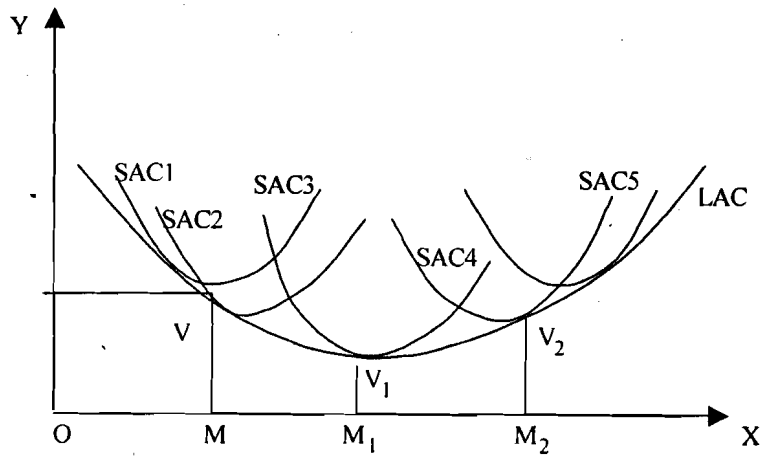
Again LAC is defined as LTC/Q and is given by the slope of the line from the origin to the point on the LTC. In Fig. 7.10, LAC at output $Q_1 = \frac{MQ_1}{OQ_1}$. This is equal to the slope of the line OM. Similarly LAC at other levels of output can be found out.

From LTC we can also find out long run marginal cost (LMC). This is given by the slope of the tangent line to LTC curve at the concerned point. In Fig. 7.10, LMC at output Q_1 is the slope of the tangent t_1. The relationship between LAC and LMC is similar to that in the short run.

In the beginning LAC>LMC upto the level of output Q_2. At Q_2, LAC = LMC. After Q_2 LMC>LAC. At Q_2 level of output see that LAC is the minimum. Thus LAC = LMC at the level of output where LAC is minimum. Below this level of output LMC is lower than LAC and above this level LMC is greater than LAC. Hence, LMC crosses LAC from below at the minimum level of LAC.

Now we turn to relationship between short run and long run cost curves. In the long run, as firms enter into the industry, the industry expands. This creates business environment conducive for economising costs through better utilitsation of transport facilities. When many industries come together in one place, we have an industrial concentration through the development of auxiliary industries supplying raw materials and encouraging skill formation. Again, high concentration increases cost of distribution, cost of living, and also creates scarcity of inputs specific to the industry. Due to external economies in the first stage of expansion and then the diseconomies, the long run average cost curve declines first then increases. Thus, the LAC is also 'U' shaped, like the SAC, but is flatter.

Fig. 7.11



In Fig. 7.11 for output OM in the short run, the average cost of production is VM in SAC_1. But in the long run, the firm will choose to establish a larger plant to produce this output and operate on SAC_2. This reduces the average cost to V', where V', is the point at which SAC_2 is tangent to LAC. In this case, economies of scale operate and each successive plant size will have a lower minimum average cost. Accordingly, SACs slide down to indicate diminishing minimum average cost. SAC_3 is the plant-size corresponding to the minimum long-run average cost, V_1M_1. On the other hand, when diseconomies of scale operate, each successive large plant will have higher minimum average cost e.g., for SAC_4, average minimum cost is M_2V_2. If we draw a curve tangent to every SAC, we get a long run average cost curve as an envelope of short run average cost curves. This long run average cost

curve is a 'U' shaped curve. This is tangent to the falling portion of the SAC while economies of scale operate, but it is tangent to the rising portion while diseconomies operate. The minimum point on LAC is a point of tangency with one of the SACs, here SAC_3, at its minimum point.

When both long run internal and external economies and diseconomies are absent and minimum average cost is the same for all plant sizes, the returns to scale are constant and the minimum average cost is also constant

Check Your Progress 4

1) In what respects long run cost curve is different from short run cost curve?

.....

.....

.....

.....

2) Explain the relationship between LAC and LMC.

.....

.....

.....

.....

3) State the reasons for the 'U' shape of the short run average cost curve in three sentences.

.....

.....

.....

.....

4) Why do the internal economies and diseconomies operate in the long run and how do they affect the shape of the long run average cost curve ?

.....

.....

.....

.....

5) Explain briefly the operation of external economies and diseconomies in an industry.

.....

.....

.....

6) Is the following statement correct? The marginal cost curve cuts the average cost curve at its minimum point.

.....

.....

.....

.....

.....

.....

7.6 LET US SUM UP

Cost for firm is the due to payments made towards purchase of inputs. Some of the inputs like capital may be fixed in the short run while others may be variable. In this unit we defined seven types of cost concepts, viz., Fixed Cost, Variable Cost, Total Cost, Average Fixed Cost, Average Variable Cost, Average Total Cost and Marginal Cost. We showed the shape of the cost curves for all these cost concepts and established the link between them.

Also we discussed the inverse relationship between cost and product curves. The difference between long run and short run cost curves is briefly discussed in the unit.

7.7 KEY WORDS

Average cost : Total cost divided by level of output.

Average Fixed Cost : Fixed Cost divided by level of output. It is a downward sloping curve in the shape of a rectangular hyperbola.

Average Variable Cost : Total Variable Cost divided by level of output.

Cost Curves : Representation of cost concepts graphically.

Fixed Cost : The cost incurred on inputs, which are fixed in nature in the short run. Capital is one example of fixed input.

Total Cost : Sum of Variable Cost and Fixed Cost is Total Cost of production.

Variable Cost : Cost incurred on variable inputs. Labour is a variable input. Hence expenditure on wage paid to labour is variable cost.

7.8 SOME USEFUL BOOKS

As in the previous unit (Production function). Add to the list

Frank, R. H., 1991, *Microeconomics and Behavior,* McGraw Hill Inc., Singapore.

7.9 ANSWERS OR HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) These concepts have been defined in the text (See Section 7.2).

Check Your Progress 2

1) In the cost function $TC = 53 + 0.5\,Q + 0.2\,Q^2$

   i) fixed cost is 53. This remains constant whatever be the level of output.

   ii) VC is $0.5\,Q + 0.2\,Q^2$. Hence AVC is $\frac{VC}{Q}$, which is, $0.5 + 0.2\,Q$.

   iii) MC is $0.5 + 0.4\,Q$

Check Your Progress 3

1) Section 7.4 explains the inverse relationship between average product and average variable cost.

2) Go through the text in Section 7.4.

Check Your Progress 4

1) Section 7.5 depicts the relationship between LMC and LAC. Go through the text.

UNIT 8 VARIOUS MARKET FORMS AND PRICING UNDER PERFECT COMPETITION

Structure

8.0 OBJECTIVES

After going through this unit you should be able to:

- identify the different market forms;
- explain the characteristics of firm's equilibrium in short run as well as in long run;
- explain the conditions of equilibrium of industry under perfect competition;
- distinguish between short run and long run equilibrium under perfect competition;
- distinguish between external economies and diseconomies of scale; and
- explain the supply curve under perfect competition.

8.1 INTRODUCTION

Until now we have learnt, how supply and demand determine an equilibrium price which rations limited outputs and allocates scarce resources in a market system. Previous block (Unit 7) discussed the way in which a firm's cost structure is influenced by diminishing returns in the short run and also explained how firms can produce at minimum cost in the long run when resources can be employed in the variable amounts.

In this unit, we aim to explore how a firm maximises its profit under perfect competition. As we know, profit is the difference between revenue receipts and costs. Revenue in the given demand- supply situation depends upon how much competition it faces in its industry (An industry is defined as the total of all firms selling a particular product). Industries fall under one of the four categories of market forms: Perfect Competition, Monopoly, Monopolistic Competition and Oligopoly.

8.2 DIFFERENT MARKET FORMS

A perfectly competitive industry satisfies the following five essential conditions, viz., (i) large number of buyers and sellers of the product; (ii) each individual seller or buyer has no control over the price at which industry sells or buys its product; (iii) free exit or entry into the industry; (iv) each firm produces identical products; and finally, (v) there is perfect information about the market.

A monopoly exists when there is only one seller of a product that has no close substitutes. In this market form, the firm is the industry and there can be barriers to the entry of competitors. Because of the complete control over the supply of the product, the monopolist normally a charges a price for her product, which maximises her profits. We will examine this form of market in Unit 9.

Between the two extremes of perfect competition and monopoly, there is monopolistic competition. In a monopolistically competitive industry, there are large a number of small sellers who are engaged in some slight amount of product differentiation. Firms have smaller degree of control over the price, entry into the industry is not as easy as in case of perfect competition, and lastly, there is ignorance about the market in this form of market structure. For a detailed discussion on this form of market we have to wait until we come to Unit 10.

Closer to monopoly (and often tantamount to it) is an oligopoly. An oligopoly industry has a few giant sellers, each of which controls a significant share of the market. Entry into an oligopoly industry is very difficult. The distinctive feature of this market form is the interdependence among its members. No firm can make independent decisions about the changes in price without reckoning to the reactions of its competitors.

8.3 ASSUMPTIONS FOR PERFECT COMPETITION

The theory of perfect competition stands on two critical assumptions: behaviour of individual firm and the nature of industry in which it operates. The **firm** is assumed to be the price-taker. It means that the firm can alter its rate of production and sales without having any effect on the price of the product it sells. Thus, the firm accepts the prevailing price in the market.

Industry is characterised by freedom of entry or exit. It means that any new firm can enter the industry if it so wishes and an existing firm is free to stop production and leave the industry if it so desires. Existing firms cannot ban the entry of new firms and there are no legal prohibitions on entry or exit.

Price-taking behaviour of firm can be better explained with the help of an example. Let us assume the probable situations faced by a wheat farmer and a car manufacturer. The latter is aware of the market power it has got and knows that if it raises the

prices of cars the sales will decline. On the other hand, it may reduce the price of the cars to attract more customers from its rivals in the market. Thus, the car-manufacturing firm is not a price-taker. It faces a downward sloping demand curve and selects a price-output combination on that demand curve.

The wheat farmer, on the other hand, differs from the car manufacturer in following respects that explain the price taking behaviour:

A homogenous product: Contrary to car manufacturers who sell differentiated products, a wheat farmer has no such choice. A particular variety of wheat is indistinguishable whether produced by one farmer or another. Therefore, we can say that wheat farmers sell homogeneous product. The buyers will shift to the other farmers if one of them raises the price of wheat.

Note that Price taking behaviour requires that firms sell a homogeneous product

Well-informed buyers: A necessary condition for price taking behaviour is that buyers are well informed and that they do not unknowingly pay more than is necessary for their purchases.

Large number of sellers: An important distinction between the car and the wheat industries is in terms of number of sellers. A wheat farmer's share in the total production of wheat is very small. It has no effect on price. The farmer will therefore, face a horizontal demand curve.

8.3.1 Pure vs. Perfect Competition

Economists often distinguish pure from perfect competition. The conditions given above deal with a theoretical model that operates under pure competition. For a competition to be perfect, certain additional assumptions required are:

a) There is complete mobility of factors of production between industries; and

b) There are no transport costs and producers of all the goods (intermediary and final) work sufficiently close to each other.

Check Your Progress 1

1) Explain briefly the prerequisites for perfect competition.

.....

.....

.....

2) Describe the situations when a firm earns a normal profit.

.....

.....

.....

3) Explain the difference between monopoly and oligopoly in two sentences.

.....

.....

.....

8.4 EQUILIBRIUM UNDER PERFECT COMPETITION

The perfectly competitive firm is a price taker, but what price does it take? The equilibrium price in perfect competition is the price at which the quantity demanded by all buyers in the market is equal to the quantity supplied by all sellers. Also at this point the market demand curve intersects the market supply curve. The market equilibrium is shown in Fig.8.1. At point of equilibrium E, OQ quantity of the commodity is bought (and sold) at OP price. The individual firm must take this price as given; in maximising profit, therefore, it can make only a decision about how much it will produce at that price.

Fig. 8.1

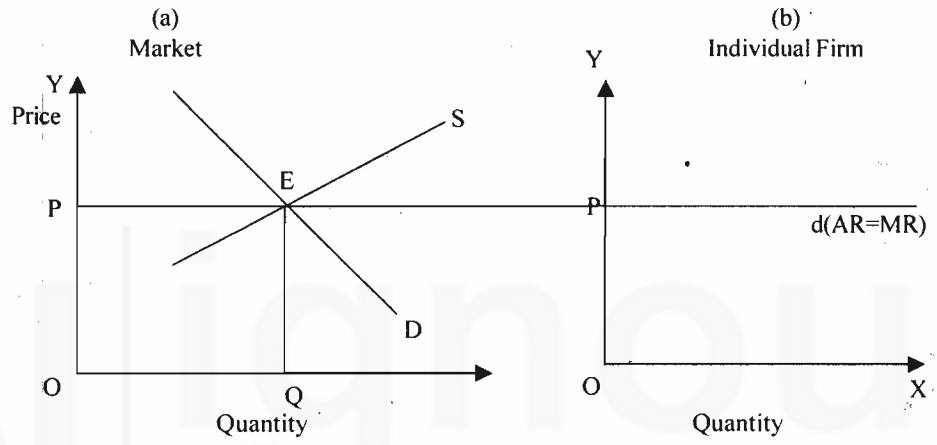


Fig 8.1: Market equilibrium occurs in segment (a) when D and S are equated. The market price P is determined. Individual competitive firm can sell whatever quantities it wishes to produce at this market price P. Thus, demand curve for a firm is horizontal line as shown in segment (b).

You know that the firm cannot control the price of its product. In other words, it is too small relative to the market to have any impact on the price and can sell any quantity of output at the equilibrium price prevailing in the market. So it faces a completely elastic demand curve for its product. However, the market demand for a product in a perfectly competitive market is not necessarily completely price elastic, as the market demand D in Fig 8.1 (a) is the sum of the quantities demanded by all the households.

Recall that we had discussed derivation of market demand curve from that of the individual's, in Unit.2. The market supply curve was also introduced in there. Since the market price is unaffected by variations in the output of the firm, it means that the marginal revenue resulting from an increase in sales by one unit is equal to the price of product, and the average revenue (Total revenue /No. of units sold) is also the same i.e., equal to price. All the three curves coincide in the same line and P = AR = MR, remains constant while the output varies.

Total Revenue (TR) that the firm receives by selling any given output at a given price is simply the price per unit (P) multiplied by the quantity sold (Q), or,

$$TR = P.Q$$

Since the firm is a price taker, every individual unit it produces adds the same amount to total revenue. Or, we can say, TR increases at a constant rate as output increases since price is constant for all outputs.

Average Revenue (AR) is the revenue per unit of output, or, the ratio of total revenue over output. If a firm sells its product at a uniform price, average revenue is simply another term for price. Algebraically,

$$\mathbf{AR = TR / Q = P.Q / Q = P}$$

Marginal Revenue (MR) is the change in total revenue that results from a unit change in output. It is the rate at which TR changes as output changes. If the firm increases its output by one unit, MR tells how much TR changes as a result of that one unit change.

$$\mathbf{MR = \Delta TR/\Delta Q}$$

For a firm, which is price taker, price and marginal revenue will be identical, or,

$$\mathbf{MR = P}$$

Every time price-taking firm raises its output by one unit, it increases its total revenue by Rs.1×P.

Since MR is the change in TR due to the change in output, MR must equal 1×P = P

8.5 EQUILIRIUM OF THE FIRM

The profit (π) is the difference between total revenue and total cost, i.e.,

$$\pi = TR - TC$$

Thus, profit reflects or sums up the effect of a change in output of the firm. A change in output raises cost of production on the one hand and affects the total revenue of the firm on the other.

The relationship between MR and MC is thus extremely important in determining the level of output the firm wants to produce in order to maximise the profit. Note that,

if MR is greater than MC (i.e., MR > MC) and MC is rising, the firm will increase its output;

if MR is smaller than MC (i.e., MR < MC) and MC is rising, the firm will decrease its output;

if MR equals MC (i.e., MR = MC) and MC is rising, the firm has reached its equilibrium level of output that maximise profit.

Thus, for equilibrium, in a perfectly competitive industry, two prerequisites have to be fulfilled, viz.,

(1) Each individual firm must be in equilibrium where its MR = MC; and (2) the industry as a whole is in equilibrium. This will occur when there is no tendency for firms to either enter or leave the industry. In other words, we describe this situation by saying that every entrepreneur in the industry is earning normal profits. Normal profits for any entrepreneur in an industry are therefore those profits, which are just sufficient to induce her to stay in the industry. It follows that if an industry is in equilibrium, with no movement in or out, no one outside the industry sees the possibility of being able to earn normal profits if is it were to enter the industry.

We know that average cost curve is roughly U-shaped (recall from the Unit 7). We also know that in perfect competition, each firm's average revenue curve is a horizontal straight line. So in perfect competition, the only situation in which the firm can be in equilibrium and earning normal profits is when the average cost curve is just tangent to the average revenue curve. Only then the firm covers all its costs and just earns normal profits. This is illustrated in (Fig. 8.2).

Fig. 8.2

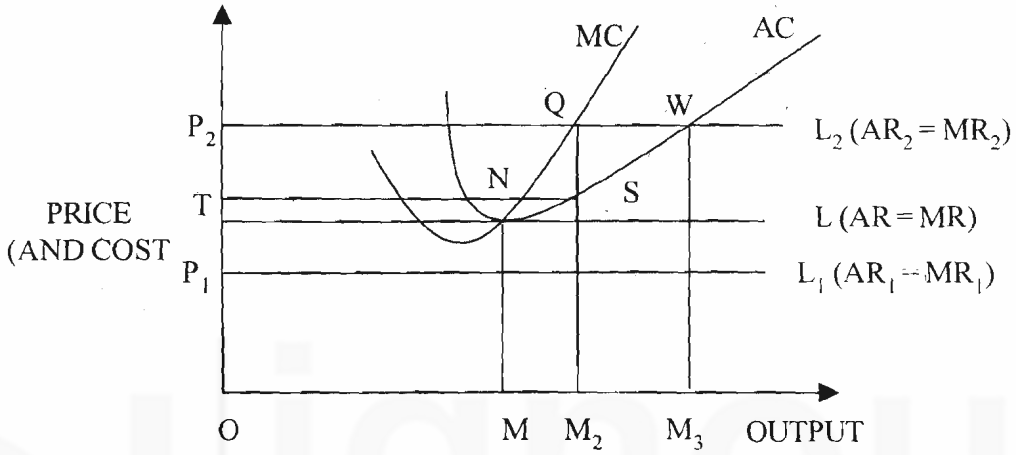


Fig 8.2: Equilibrium of the firm occurs at point N where its demand curve is just tangent to its average cost curve AC. Note that at point N, average cost of production is minimum and marginal cost curve MC cuts AC at this very point. Hence at equilibrium we find that MR = MC = AR = AC.

The above figure makes it quite clear that the firm cannot be both in equilibrium and earning normal profits for any position of the AR = MR curve below PL. For example, with the average revenue curve P_1L_1 the firm will earn less than normal profits whatever its output may be. At no level of output average cost is equal to price. On the other hand, if the average revenue curve is above PL, say, P_2L_2, it is possible for average cost to equal price (in this case at W). But at W the firm would not be in equilibrium. The firm sees an opportunity to expand output in the short run to take the advantage of available price conditions. However, perfect information assumption can induce new firms to enter the market as well. So, on the one hand, this firm raises its own average cost of production by moving up its AC curve, the new entrants on the other hand take away some of the possible customers. However, as we know, the firm will be in equilibrium only when its marginal revenue equals marginal cost. With the average revenue curve P_2L_2, this will happen when the firm produces OM_2 and sells it at the price OP_2. At this output, the firm will be in equilibrium but will be earning 'supernormal profits', equal to the area P_2QST. This again is a signal to all the firms in the industry to try to expand as well as to new firms to flock to this line of business.

The only situation, in which the firm can be in equilibrium while earning only normal profits, is the position where it is producing OM output. This happens when the price of the industry is OP. Then not only MR will be equal to MC, but average cost will also be equal to price or AR. Further, at such a point AC is tangent to AR.

8.5.1 Short Run Equilibrium

As already defined in Unit 7, short run is a period of time, which is long enough to allow the variable factors of production to be used in different amounts so that

maximum profits are earned; but during which the fixed factors cannot be altered. Moreover, when we are discussing equilibrium in short run, we must be quite explicit about the cost conditions. If costs differ between firms, the equilibrium position of the industry will not be the same as is the case when they are identical. We shall, therefore, consider the short run equilibrium of the firm and of the industry in the following three different situations.

First, we shall assume that all factors of production, including entrepreneurship, are homogeneous. Therefore, each entrepreneur will be able to combine the same factors of production in the same way. Now since we are assuming that every firm produces its output at the minimum possible cost, we may conclude that all firms will have identical cost curves when factors are homogeneous.

Second, when we assume that all factors of production except entrepreneuership are homogeneous, costs will differ between firms because some entrepreneurs are more efficient than others.

Third, when all factors are assumed to be heterogeneous, the cost differences between firms will be even greater because all factors are of different efficiency.

Short run equilibrium of the firm can also be explained graphically. There are three possible positions for a firm in the short run equilibrium. The following Fig. 8.3 explains it using the homogeneous factors in a situation of perfect competition in the factor market and where prices of all the factors are given.

Let us assume that to begin with the price of the product is OP_2. Therefore, all the producers will set their output accordingly to maximise profit. A firm produces the output OM_2 and sells at the perfectly competitive price of OP_2. The average cost is M_2R and the per unit supernormal profit is LR. The firm earns the supernormal profit shown by the area P_2LRS. All firms in the industry are in equilibrium since they are earning the maximum profit.

Fig. 8.3

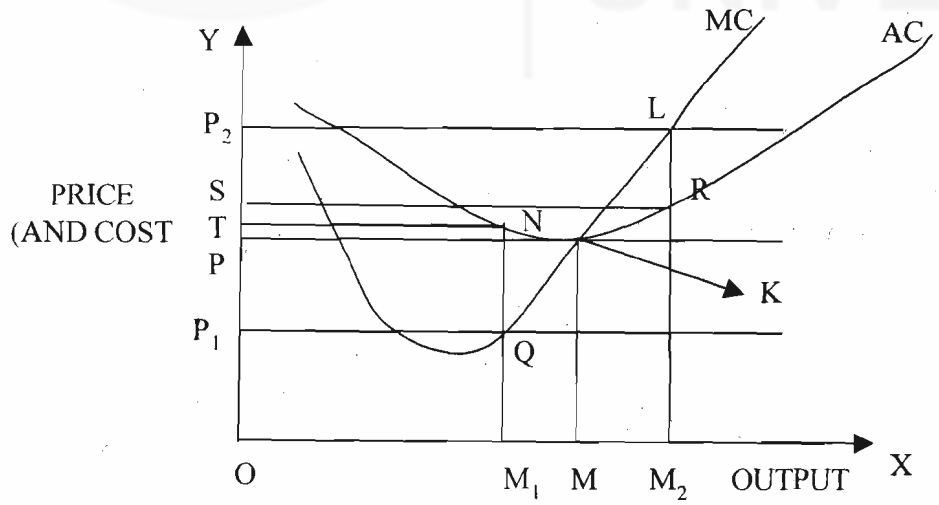


Fig 8.3: If short run price is P_2 >P, the firm produces OM output and makes extra normal profits =P_2LRS. If short run price P_1 <P, the firm is not able to cover all the costs, yet it continues to produce if at least the variables costs are covered. Output is OM_1 given by the equilibrium MR=MC.

In the short run, same situation will continue since the firm(s) cannot introduce new fixed cost equipment and no new firm can enter the industry. In the long run, however, new firms will enter the industry and they will compete away the supernormal profits.

Now if price were OP instead of OP_2 all the firms would be in equilibrium when producing the output OM. They would just be covering their normal costs and earning normal profits. Average cost curve is tangent to average revenue line at output OM. It thus reflects the situation of equilibrium even in the short run. The number of firms in the short run is just large enough to help each of them earn normal profits.

On the other hand, if price is OP_1, the firms in the industry would be in equilibrium, producing output OM_1. All would be losing money, losses being equal, to P_1QNT. The maximum profit each firm could earn would be a loss of P_1QNT. This is the smallest loss each firm could make in the short run. In the long run, however, the firms would leave the industry until those remaining there just earn normal profits.

The next obvious question to ask is, therefore, does it pay a firm, which is loosing money, to stay in the industry? The answer to this question, however, lies on the length of time. In the long run, if its efficiency does not improve, it will leave the industry altogether.

We have said above that, in the short run, the firm cannot alter its fixed assets. Such a constraint will influence the entrepreneur's actions to be able to stay in business. It, for sample, the variable costs of the firm is met in the short run; the firm would not close down. This is explained in Fig. 8.4. It shows the cost curves of a firm in our industry. The firm will remain in industry so long as it can cover its variable costs. If the firm closes down, it may be forced into bankruptcy, which may imply total loss of fixed investment. On the other hand, continuing to produce, even when fixed costs are not entirely covered, will help it maintain its presence in the market, and its plant and machinery will not be reduced to junk.

Fig. 8.4

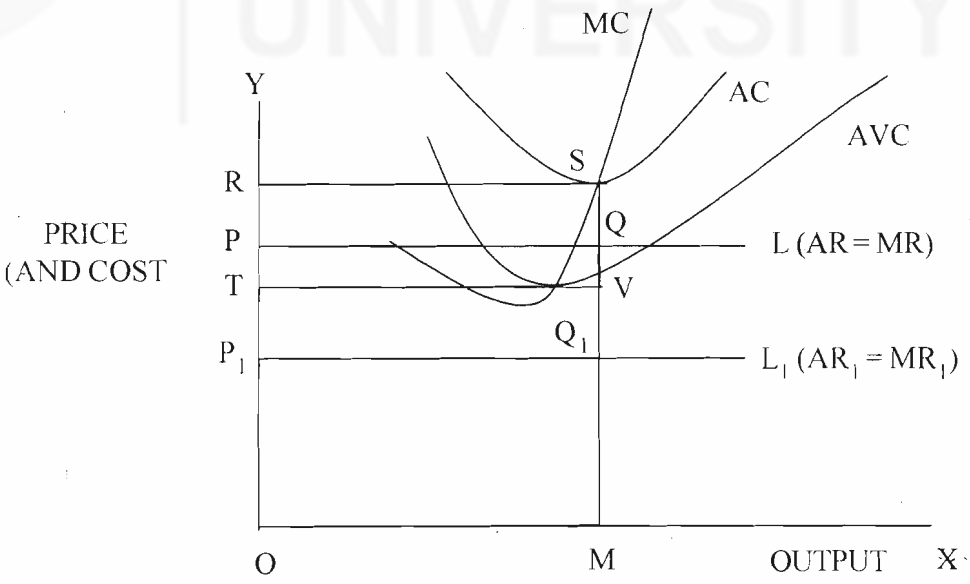


Fig 8.4: In short run, firms many find it better to keep producing at OM though the price OP does not cover all the costs. Its fixed cost TRSV is too large. By continuing production it makes a contribution towards fixed cost = TPQV. Therefore, at O price, its output OM is loss-minimising output.

Looking at the figure, we can see that the loss minimising output for the firm is OM at the price OP per unit of output. Price, therefore, exceeds average variable cost MV by QV. This amount is known as per unit contribution towards meeting the fixed costs. At this point, it will be worthwhile for the firm to continue running in the short run. Some fixed costs can be covered after variable costs have been met. In such a situation, PQVT is described as total contribution. This is the total amount at which the firm reduces its loss by staying open. The firm is losing RSQP, which is less than it would have lost, had it been closed down altogether. In that case it would have lost an amount equal to its total fixed cost, i.e. RSVT. Instead, it looses RSVT - PQVT.

If short run price is only OP_1, the cost per unit of output would have been MS and loses per unit SQ_1. In this case, the firm is not able to cover even the average variable cost. For example, at output OM, the firm is losing more than its fixed cost RSVT by P_1Q_1VT. Here, at this point it will therefore pay the firm to stop producing anything even in the short run. In fact, the firm shall not produce anything if it is not able to get price at least as great as OT - the minimum average variable cost covering price.

Now let us move to the second situation where entrepreneurs are not identical in efficiency. Costs then can be different between firms, despite the fact that all factor prices are the same to all firms and all factors except entrepreneur are homogenous. Some entrepreneurs will be able to produce more efficiently than the less efficient ones, and their firm's costs will, therefore, be lower. Even though all firms produce the same product, sell it at the same price and seek maximum profits, firms with different costs will maximise profits at different levels of output. Maximum profits will of course, differs too.

Let us assume that there are four firms in an industry. Firm A has the most efficient entrepreneur of the four and is in short run equilibrium. When producing the output at the given price it earns supernormal profits. Firm B with a somewhat less efficient entrepreneur produces an equilibrium output and earns only normal profits. Firm C with a still less efficient entrepreneur incurs losses. However, since firm C is covering variable costs it pays to go on producing the output in the short run. The losses are thereby minimised. Finally, firm D has the least efficient entrepreneur of the four firms. This firm is not able to recover its variable costs at any level of output. Therefore, it minimises losses by closing down, even in the short run.

Let us now look into a third situation where all the factors of production are heterogeneous. So, the difference in costs between firms is likely to be even bigger than where entrepreneurs only were of differing efficiency. The same kind of situation as described above will occur with the likelihood of difference between the costs of the firms being greater.

8.5.2 Long Run Equilibrium

In the long run we know that there is no restriction on entry (exit) into (out of) the industry and all factors of production are perfectly mobile. Firms under the perfect competition will be earning normal profits. In this period, firms will enter or leave the industry until no firm outside the industry thinks it could earn at least normal profits, if it were to enter the industry. And, no firm in the industry will leave unless it thinks that it could do better by leaving. Let us consider the process in more detail.

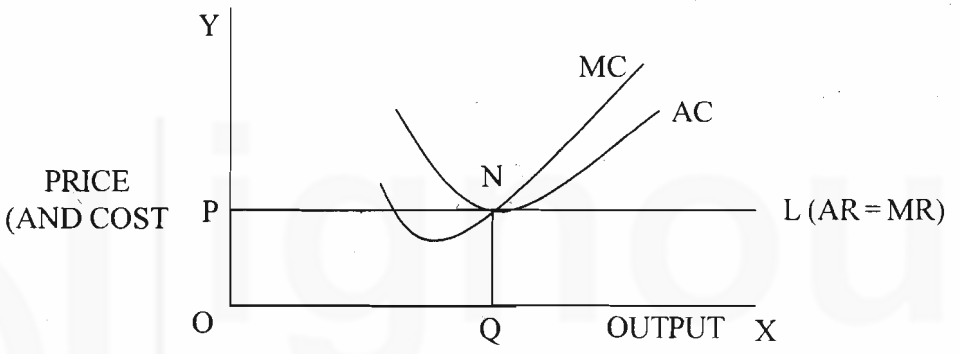
If all the firms in the competitive industry are in the position of earning more than normal profits, new firms will enter the market, and there will be more suppliers.

This will lead to the shift in the supply curve to the right. With unchanged demand conditions, the previous price will no longer prevail and the equilibrium price will fall. The falling prices will continue until all firms in the industry are just covering their total costs or are at the zero-profit equilibrium.

On the other hand, if the firms in the industry are suffering losses, some of the firms will exit. The supply curve will shift to the left. Supply decreases with no change in demand will result in a rise of the market price. Firms continue to exit and price continues to rise until the remaining firms can cover their costs, that is, until they are earning normal profits.

The equilibrium of the firm in the long run is shown graphically in the following Fig. 8.5 where it is earning normal profits.

Fig. 8.5



8.6 NORMAL PRICE

We must familiarise ourselves with the concept of normal price. Marshall defined normal prices as those prices, which may reasonably be expected to prevail in given conditions of demand and supply. Time is a very important factor here. A different price will be normal in the long period from the one in the short period. However, in practice, a long period normal price will never be arrived at. As we know the long run is like the tomorrow that never comes. There will always be a change in some of the conditions underlying the long period equilibrium before it has time to be established.

Check Your Progress 2

1) State 'true' for correct statement and 'false' for the wrong one.

   a) Marginal revenue is the revenue per unit of output.

   b) Under perfect competition, the firm is not a price-maker.

   c) There are no restrictions on the entry or exit of firms under perfect competition.

   d) Perfect competition assumes the existence of transport costs.

   (a).....(b).....(c).....(d).....

2) Explain the equilibrium attained by a firm in the short run when all factors of production except entrepreneur are homogenous.

.....

.....

.....

3) Explain briefly the long run equilibrium of the firm.

.....

.....

.....

4) When a firm is incurring losses, at which point will it leave the industry even in the short Run?

.....

.....

.....

8.7 ECONOMIES AND DISECONOMIES OF PRODUCTION

We will now learn to construct short and long run supply curves of the industry. But before that we shall discuss external economies and diseconomies of production. It may be useful to note that there are internal economies and diseconomies observed in the process of production, which can be obtained by altering factor proportions or by altering the internal organisation.

External economies and diseconomies in contrast to that of internal, depend on the increase in the output of the whole industry rather than on the output of individual firm. External economies of the industry occur where an increase in the size of an industry leads to lower costs for each individual firm in the industry. Of course, it is quite possible that the growth in the size of industry will lead to external diseconomies and to rise in the unit costs.

8.8 SUPPLY CURVE OF THE PERFECTLY COMPETITIVE INDUSTRY

Our main objective in the present section has been to get to a point, where we can build up a supply curve for a perfectly competitive industry. The shape of a competitive supply curve will not, however, always be the same. It will depend on the production conditions underlying it, which are based on certain assumptions. To begin with, it is assumed that all factors of production are homogeneous and are in perfectly elastic supply to the industry. Moreover, we shall continue to assume that there is perfect competition in the factor market so that the prices of all units of each factor are the same.

Now we will discuss supply curve of the perfectly competitive industry in two different situations: (a) Long run: a situation where neither the number of firms nor their scale

of production is fixed; (b) Short run: supply curve where not only the number of firms but the scale of production of individual firm is given.

We start with a situation where all firms in the industry have identical cost curves because factors are homogeneous. Each firm will, therefore, originally be in long run equilibrium, where it is producing an output of OM_1 (see Fig. 8.6) and selling it at a price of OP_1. Short run output given by SMC = MR_1 is OM_1. MR curve is equal to both short and long run marginal cost for all firms in the industry. All these firms are earning normal profits.

Fig. 8.6

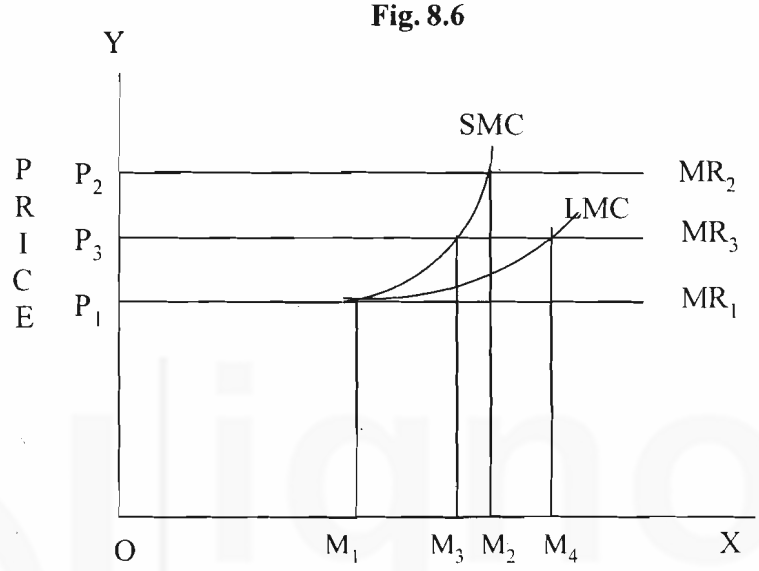


Fig 8.6 shows that the firm will be in the long run equilibrium. It will Produce output OM_1 and will sell at Price OP_1.

In order to find out the shape of the supply curve, we assume that price of the product rises to OP_2 and thus marginal revenue curve to P_2MR_2. In the short run, it is obvious that in perfect competition, the supply curve of the individual firm is its short-run marginal cost curve (SMC). This is the only way in which a firm can produce more by expanding output alongwith its short run marginal cost curve. If price rises to OP_2 as in Fig. 8.6, MR will be equal to MC at output level OM_2. for each firm in the industry. The firms' short run SS curve is its SMC. The industry SS will be lateral summations of SS's of all its constituent firms.

As each firm produces higher output, the industry SS shifts to the right. This will imply that given the demand conditions in the market, point of intersection of industry SS with market 'DD' takes place to the right. The industry can sell the increased output at a lower price only. Thus, as each firm tries to produce greater output, they tend to face a situation where they can realise a price OP_3 only. Thus, in long term, they all may produce say OM_3 level of output and sell it at OP_3 price. Notice that $OP_3 < OP_2$ and also that $OM_3 < OM_2$. It is also clear that the long term marginal cost curve (LMC) is less steep compared to short run marginal cost curve.

However, even when the number of firms remains constant, the long run equilibrium situation in Figure 8.6 will be, where output is OM_3 and price is OP_3. The long run marginal cost (LMC) curve of the industry will be less steep than the short run MC and output will expand more in the long run as compared to short run (by M_1M_3 instead of M_1M_4 for each of the firms. As we see this particular type of supply curve also slopes upwards to the right though less steeply than the short run supply

curve. In the new long run situation, it may be noted that with price OP_3, all the firms in the industry will be earning super normal profits. This can happen because the number of firms in the industry is assumed to be constant. But this situation cannot continue as information will spread soon and new firms will be attracted towards this industry.

8.8.1 Changing Number of Firms

Let us consider another aspect of long run situation, where not only the size of the firms in the industry alters, but the number of firms also changes. We assume that there is free entry into the industry and the new firms continue to enter, until super normal profits are eroded. In the short run as shown in Figure 8.6 the firm will produce output OM_2 at price OP_2. In the long run, however, the individual firm will be exactly at the same position as in the original equilibrium. Output will be OM_1 and price OP_1. Marginal revenue will be shown by the curve P_1MR_4 that is identical with the curve P_1MR_1. The output of each firm will be OM_1 but now the number of firms has increased. Consequently, the output of the industry will be more at the same price Firms, which are identical with those already there, have also entered to the normal level. When all factors of production are homogeneous, the long run supply curve of the industry with free entry will be horizontal straight line and supply price will be the same at every output. This is the simplest possible kind of long run supply curve.

However, in practice it is unlikely that all factors are homogeneous. A much more plausible case would be of heterogeneous factors, at least, due to the presence of entrepreneur. So far as the short-term supply curve is concerned, the fact that entrepreneurs are heterogeneous will make little difference. The short run supply curve will still represent a lateral summation of the short run marginal cost curve of the individual firms. Marginal cost curves will now be different for each firm and the process of summation more complicated. Whether or not factors of production are homogeneous, the short run supply curve is bound to slope upwards to the right. However, with heterogeneous factors it will slope upwards rather more steeply.

In the long run, with the free entry or exit in the industry, the supply curve is likely to slope upwards rather less steeply than it does in the short run.

Change in input prices, technology, quality of resources, taxes on the firm or its output, or any of the determinants of production and cost relationships will cause the perfectly competitive industry's long run and short run supply curves to shift. For example, higher fuel and energy prices will increase costs of production of almost everything. Firms respond to higher costs by reducing the output and thereby shifting the short run supply curve to the left until a new long run equilibrium is reached.

Check Your Progress 3

1) Explain briefly what do you mean by external economies and diseconomies. (Answer in three sentences)

.....

.....

.....

.....

.....

2) Find out the shape of the supply curve in an industry having firms with identical cost curves. (Answer in the space provided below.)

.....

.....

.....

.....

3) Explain the supply curve of perfectly competitive industry when not only the size but the number of firms also changes.

.....

.....

.....

.....

8.9 LET US SUM UP

In this unit, various market forms have been explained. Introducing the concept of perfect competition, it has been explained how such a market structure is different from pure competition and what are its operational the requirements. The firm under perfect competition is a price-taker. When marginal revenue is equal to marginal cost and the latter is rising, the firm reaches its equilibrium position. Equilibrium of the firm has been discussed for short run as well as long run situations. Equilibrium of industry has been discussed with reference to time. Here the time has been divided into short period and long period equilibrium.

The discussion of various aspects of the perfectly competitive industry has been extended to cover the short run and long run supply curves of the firm and industry under perfect competition.

8.10 KEY WORDS

External Economies and Diseconomies : External economies operate when an expanding industry has cost reducing effect on every firm due to better transport facilities, better supply of inputs etc. External diseconomies operate when and expanding industry gives rise to higher average cost for every firm due to scarcity of supply of inputs and the pressure on transport facilities, congestion etc.

Internal Economies and Diseconomies : Internal economies operate within a firm with increasing output, when cost reducing effects occur due to better utilisation of fixed inputs, greater division of labour, better and cheaper marketing facilities, etc. Internal diseconomies operate due to excess of fixed

inputs and reduction in the proportion of fixed to variable inputs.

Long Run : The time period when all inputs, including plant capacity, are variable.

Marginal Cost (MC) : Additional cost incurred for producing one extra unit of output is its marginal cost.

Marginal Revenue (MR) : Extra revenue that is added by selling one extra unit of output is its marginal revenue.

Short Run : The time period when at least one of the inputs (plan size or building) is fixed.

8.11 SOME USEFUL BOOKS

Koutsoyiannis, A. 1979, *Modern Microeconomics*, Macmillan, New York, Chapter 3, 4 & 5, pp. 67 to 163.

Salvatore, D. 1983, *Microeconomic Theory*, Schaum's outline Series, Chapters 7, 8 and 10, pp. 124 to 174 and 196 to 220.

Samuelson, Paul A. and W.D. Nordhaus, 1985, *Economics*, McGraw-Hill Book Company, Chapters 21 and 22, pp. 461 to 501.

Lipsey, Richard G. 1979, An *Introduction to Positive Economics* English Language Book society/ Weidenfold and Nicoloon, Chapters 16, 17, 18, and 19, pp. 201 to 259.

Stronier, Alfred W. and Hague, Douglas C. 1986, *A Text book of Economics Theory*, Macmillan and ELBS, Chapters 5,6 and 7, pp. 103 to 188.

8.12 ANSWER OR HINT TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) See Section 8.3

2) See Section 8.3

3) See Section 8.2

Check Your Progress 2

1) (a) False (b) True(c) True (d) False

2) See Sub-section 8.5

3) See Section 8.6

4) See Sub-section 8.5

Check Your Progress 3

1) See Section 8.8

2) See Section 8.9

3) See Sub-section 8.9.2

UNIT 9 THEORY OF MONOPOLY

Structure

9.0 OBJECTIVES

An earlier unit (Unit 8) has already covered pricing and output decisions in one extreme form of market, that is, perfect competition. In this unit, a study of such decisions is taken up in another extreme form of market, that is, in monopoly, where there is only a single seller.

After going through the unit, you should be able to:

- define monopoly;
- describe the causes why monopolies emerge;
- discuss the demand and the cost conditions, and the pricing and output decisions under monopoly;
- explain the short-run and the long-run equilibrium of a monopoly firm;

- describe and analyse price discrimination under monopoly; and
- describe various features and pricing techniques of a public monopoly.

9.1 INTRODUCTION

In general, monopoly is said to exist if there is one and only one seller (producer) of a product for which there is no close substitutes; this single seller is unaffected by the prices and outputs of other products sold in the economy. The stress in this definition is on absence of close substitutes. To understand this remember that one can often find substitutes, for goods and services. For example, if one wants to travel by train, Indian Railway is the only option available. In that sense, Indian Railways is a monopoly concern. However, there are other modes of transport, such as roadways and airlines to travel from one place to another. Substitutes to Indian Railways are available but close substitutes are lacking. Thus, Indian Railways meet our definition of monopoly. The following conditions prevail in monopoly market:

a) There is only one seller (i.e., producer) in the market but there are large number of buyers.

b) The product of the seller may be homogeneous, or, there may be differentiated products but without close substitutes. When close substitutes are not available, it implies that the cross-elasticity of demand between the product of a monopolist and products of other firms is very low.

c) There is no free entry to the market. Entry may be restricted by the natural factors such as control of raw materials, by legal and institutional factors such as patent rights or by technological factors such as the efficiency of large scale production.

If all these conditions are met simultaneously, the market structure is referred to as monopoly. The monopoly, like the perfect competition, is more or less, an hypothetical situation because some of the above conditions may not be met in practice.

9.2 CAUSES OF MONOPOLY

For monopoly to exist, there must be some way to keep competitors out of the industry. Among the barriers that perpetuate monopoly are (i) patents and licensing by government agencies, (ii) control of raw material supplies, (iii) the establishment of brand names, (iv) pricing policy designed to keep rivals out of the industry, (v) large capital investment necessary to enter the industry, and (vi) size of the market. The above factors may be grouped under three headings: (a) natural, (b) legal and institutional, and (c) technological. Natural factors relates to the ability of the monopolist to control the supply of raw material ; for example, ownership of sites having concentration of deposits. Now a days, legal and institutional barriers and technological barriers have acquired greater importance for the growth of monopoly. Permits and other legal rights granted by the government restrict the entry of others in the business. Patent rights to protect productive techniques or products are other examples of legal factors. Awarding of an exclusive franchise to serve a market leads to monopoly situation. The most important one in the category of technical barriers is the efficiency of large scale production, particularly, when the size of the market is not very large. The economies of large

scale production enable an existing firm to expand its level of output and to gain from the cost reductions. The minimum cost of production occurs at a rate of output more than sufficient to supply the entire market at a price covering full cost. New firms find it unprofitable to compete with an established large firm in such circumstances. Monopoly is inevitable in industries that require to install elaborate network of pipes or cables for taking its product or service to consumers such as water supply, electricity, telegraph and telephone.

9.3 DEMAND AND MARGINAL REVENUE CURVE UNDER MONOPOLY

You have noted earlier (in Unit 8) that under perfect competition, price and marginal revenue are same, since each firm accepts the price determined in the market (by the market forces of demand and supply). However, the demand curve facing the monopolist is the industry demand curve. Remember that there is a single firm producing the commodity and thus there is no difference between a firm and an industry. The market demand curve generally slopes downward to the right (except for the rare case of Giffen good). Demand, marginal revenue and elasticity have already been explained in earlier units (Units 4 and 5). There is need to probe a little farther into the relationships among them to get some useful propositions.

When demand curve is negatively sloped, marginal revenue curve lies below it suggesting that the marginal revenue falls at a faster rate than that of average revenue. If demand curve is linear, i.e., straight line, marginal revenue falls at twice the rate of average revenue. For example, as price goes down by Re.1 at a time, marginal revenue goes down by Rs.2 at a time (see Fig. 9.1)

Fig. 9.1

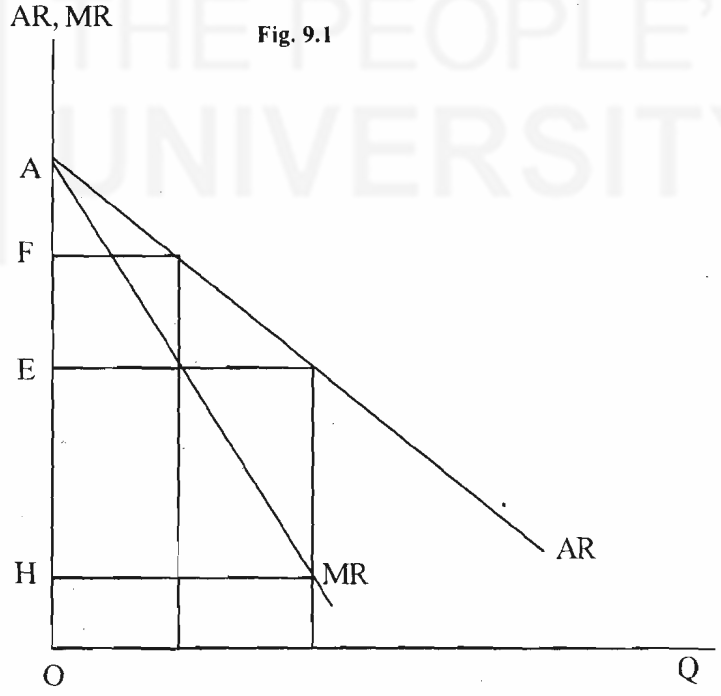


Fig. 9.1 shows inverse relationship between price of a commodity and its quantity demanded. The monopolist alone constitutes the entire industry. Therefore, her demand curve coincides with the industry demand curve, that is , it is downwards sloping. In other words, a monopolist can sell larger quantities at relatively lower prices only. Also note that as price falls from OF to OE, the marginal revenue declines from OF to OH But, EH = 2EF. Therefore, MR falls twice as sharply as AR.

Let us recapitulate the relationship among average revenue (AR), marginal revenue (MR) and elasticity (e), namely, MR=AR - AR / e = AR(1 - 1 / e). If e > 1, as we lower the price, total revenue increases. Since marginal revenue is the addition to total revenue, it follows that marginal revenue must be positive. If, on the other hand, total revenue falls with lower price, e < 1 and marginal revenue must be negative. If e = 1, total revenue does not change with a change in price, marginal revenue must be zero. Thus, we can say that, when MR > 0, then e > 1; when MR=0, then e = 1, and when MR < 0, then e < 1. This is shown in the Fig. 9.2.

Fig. 9.2

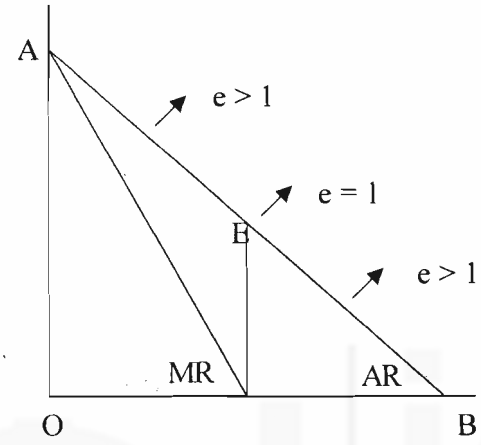


Fig. 9.2: Restates relationship between AR, MR and elasticity of demand. Recall that elasticity at any point on the demand curve is the ratio of its lower segment to the upper. Also note that when elasticity is unity, MR drops down to zero. If point E is on the demand curve AB, then over the range EB, e < 1, at E, e = 1 and over EA, e > 1.

Using the above formula, if e > 1, say 2, when MR = AR - AR / 2, and so MR is positive. If e = 1, then MR = AR - AR / 1 = 0. If e < 1, say ½, then MR = AR - AR / 1/2 = AR - 2AR < 0; and so MR is negative.

9.4 PRICING AND OUTPUT DECISIONS UNDER MONOPOLY

The pricing and output decisions of a monopolist are arrived at the same way as that of a perfectly competitive firm. That is, the monopolist wants to maximise profit in the long run. In the short run, the producer attempts to maximise the difference between total revenue and total cost, provided that variable costs are covered.

At the outset, it may be noted that monopolist will not determine the output of her product at any level where the elasticity of demand is less than unity. As discussed above, if e < 1, MR is negative. The total revenue can be increased by restricting output. Since marginal cost is always positive, a reduction in output will reduce total cost. Thus, profits must rise as output is reduced (or price is raised). Therefore, a monopolist cannot be in equilibrium if elasticity of demand is less than one.

To sum up the above discussion, a monopolist's equilibrium will always be where the elasticity of demand for her product is greater than one. A profit-maximising monopolist will never sell at a price where demand is inelastic. She will operate in

its elastic range. In the short-run, given the assumption of profit maximisation, the equilibrium position of the monopolist is deduced from the cost and revenue functions. The equilibrium conditions for a firm discussed in earlier units apply to monopoly firm also. A monopoly firm will be in equilibrium when (a) MR of the firm is equal to its MC, and (b) MC curve intersects the MR curve from below. In the short-run, marginal cost curve reflects the relationship between marginal cost and change in output from given plant and other facilities.

Fig. 9.3

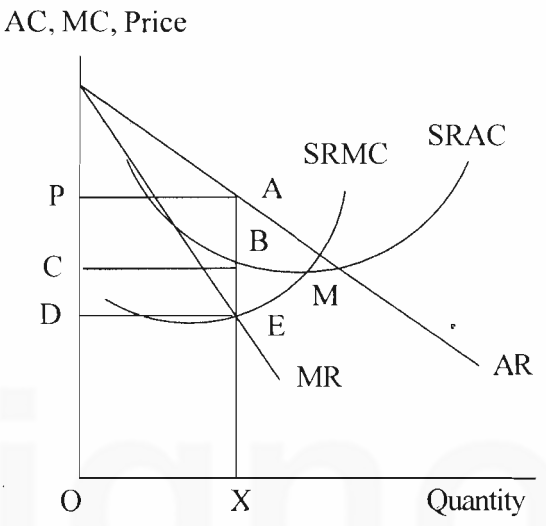


Fig.9.3 shows short run equilibrium of the monopolist firm. The usual condition of equilibrium, rising MC cuts MR from below applies to the monopolist as well. However, this firm is able to charge the highest price that the consumer may be willing to pay. The level of output is determined at OX, the point that coincides with the ordinate of the point of intersection of MR and MC curves. At this level of output, the maximum price that consumers are willing to pay is OP and the monopolist charges this very price to maximise his profits.

Fig. 9.3 depicts average revenue (AR) and marginal revenue (MR) curves, average total cost curve (SRAC) and marginal cost curve (SRMC) respectively. The equilibrium occurs at point E where both the conditions of equilibrium, i.e. MR equals SRMC and SRMC intersects MR from below are met. The equilibrium output and price are OX and OP. This shows that price is greater than average costs, and so a profit is being made. It also tells that profits are being maximised, since MC and MR are equal. If output were to increase beyond OX, the addition to total cost would be greater than the additions to total revenue, resulting in a decrease of profit. If output were less than OX, the additions to total revenue would be greater than the addition to total cost and profit would increase by expanding output. So, profits are maximised at OX and are shown by the rectangle PABC in the Fig. 9.3 above.

9.5 LONG-RUN EQUILIBRIUM OF A MONOPOLY FIRM

In the long run, a monopolist has to determine whether a plant of different size and thus a different price and output combination will enable him to earn larger profit. For this purpose, she will consider long-run marginal cost which reflects the cost

associated with a change in output when all factors, including the scale of plant, vary. A profit-maximising monopoly will be in equilibrium at that level of output where long-run marginal cost and long run marginal revenue are equal and long-run marginal cost intersects corresponding marginal revenue curve from below. Such long-run equilibrium for the firm must also imply short-run equilibrium. The firm, therefore, makes the choice of the size of its plant in such a way that (i) AR $\geq$ AC, i.e., it gets positive profit, or, (ii) AR = AC, it gets no 'extra', profit. Situation like AR < AC is ruled out in the long period (Why?). In a pure monopoly, no potential competitor can enter the market, by definition, and thus economic profit cannot be eliminated even in the long-run.

Check Your Progress 1

1) What do you understand by monopoly? What conditions must prevail for monopoly to be established?

.....

.....

.....

2) What are the conditions that might give rise to a monopoly?

.....

.....

.....

3) The demand curve facing a monopolist is downward sloping, Explain.

.....

.....

.....

4) Discuss the relationship between average revenue curve and marginal revenue curve of a monopoly firm.

.....

.....

5) Draw a linear demand curve and its associated marginal revenue curve. State the formula that relates marginal revenue, price and elasticity of demand, and explain how the curves illustrate the relationship identified by the formula.

.....

.....

.....

6) A monopolist is not free to determine both price and quantity to be sold according to her whims and fancy. Explain.

.....

.....

7) Explain how price is determined under simple monopoly?

.....

.....

.....

8) Discuss the determination of price and output of a monopoly firm in the long-run.

.....

.....

.....

9.6 SOME QUESTIONS REGARDING THE BEHAVIOUR OF MONOPOLIST

Some questions might arise in your mind regarding the behaviour of a monopolist. For example, (a) does a monopolist always make profits? (b) does a monopolist always benefit from a price rise? (c) does a monopolist have unique supply curve? (d) does a monopolist, since she is the only producer in the industry, produce with the optimum scale of plant at optimum rate of output? (e) is monopoly compatible with a downward-sloping/constant marginal cost curve? (f) is monopoly an inefficient type of market structure? These questions will be discussed in turn.

9.6.1 Does a Monopolist Always Make Profits?

There is certainly no reason to believe that the monopolist always makes a profit. If demand is sufficiently low relative to cost, the monopolist may incur a loss. For example, in the following Fig. 9.4, AC curve is above the demand curve (AR) at the point of equilibrium. The firm is incurring losses as shown by the rectangle ABCD.

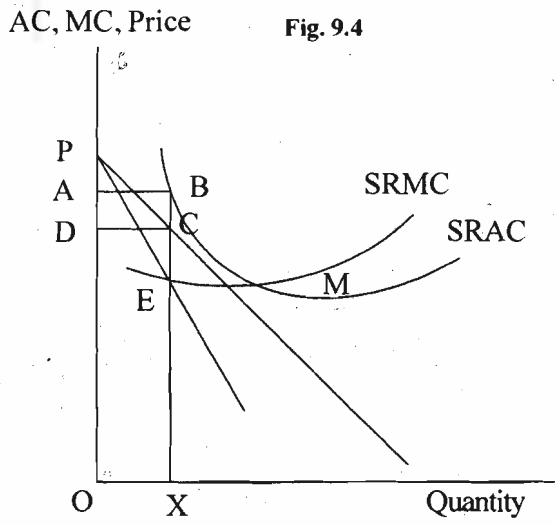


Fig. 9.4

Fig. 9.4 shows a situation where the monopoly firm may be forced to tolerate some 'loss' in the short run. It will continue to produce so long as AR > AVC. If demand curve slides lower, the firm will be forced to shut down even in the short run. If, in the long run, the firm is able to adjust its plant to the size of the market demand, well and good. Otherwise, it has no option but to quit.

In such a situation, a monopoly firm will go out of the business in the long-run unless it switches over to an appropriate sized plant. In the short-run, it stays in the business if it is able to cover its variable cost. If the equilibrium price is less than the average variable cost, the monopoly firm will close down its plant even in the short-run just like a perfectly competitive firm.

9.6.2 Does a Monopolist Always Benefit from a Rise in Price?

As explained above, a price rise (and thus, a reduction in output) would invariably benefit the monopolist if she operates over the range of demand curve where the elasticity of demand is less than unity. In case elasticity is less than one, marginal revenue is negative, rise in price causes fall in demand which results in rise in total revenue on the one hand and fall in total cost (assuming marginal cost is positive) on the other. Thus, rise in price is beneficial to the monopolist if $e < 1$. However, if the monopolist is operating in the elastic range of demand, the marginal revenue is positive when price is raised in this range, total revenue as well as total cost falls. Rise in price would be beneficial if fall in total cost is more than fall in total revenue. Otherwise, the producer may have to lower prices to increase profits.

9.6.3 Does a Monopolist have a Unique Supply Curve?

The monopoly firm does not have a unique supply curve since price is not equal to the marginal cost. Given a marginal cost, the same quantity may be offered at different prices depending on the price elasticity of demand. This can be said as the same quantity will be sold at different prices depending on the elasticities of various possible demand curves. Similarly, at the same price, different quantities would be supplied depending on the elasticities of the demand curve. Thus, there is no unique relationship between price and quantity.

9.6.4 Does a Monopolist Produce with Optimum Scale of Plant at Optimum Rate of Output?

Whether a monopolist produces with optimum scale of plant at the optimum rate of output depends upon the position of the market demand curve relative to the cost curves of the firm. Given cost curve, the scale of plant depends upon the position of the market demand curve and its associated marginal revenue curve. Since the entry of new firm is blocked, there is no pressure for the monopolist to reach an optimal scale or to use her existing plant at optimum. It is not necessary for the monopolist to reach the minimum point of LAC. The size of plant and degree of utilisation of any given plant size depend entirely on the market demand. A producer may reach the minimum point of LAC (optimal scale), remain at falling part of LAC (less than optimal scale), or expand beyond the minimum LAC (greater than optimal scale) depending on the market conditions.

It may be noted that because of entry barriers under monopoly, there are absence of market forces which compel the firms to operate at optimum plant size and utilise it at its full capacity in the long run. Any of the above case can occur depending on the size of the market.

9.6.5 Is Monopoly Compatible with Falling/Constant Marginal Cost ?

Perfect competition is incompatible with a continuously downward-sloping / constant marginal cost curve. The marginal cost curve of the perfect competitive firm must rise at the point of equilibrium output. But this need not be so in case

of monopoly. A monopolistic firm can be in equilibrium with rising, falling or constant marginal costs provided that (a) marginal cost is equal to marginal revenue, and (b) marginal cost curve shall cut marginal revenue from below.

Fig. 9.5 depicts the equilibrium of a monopolist firm whose marginal costs are constant and are equal to average costs.

Monopoly equilibrium is also possible with falling marginal costs to ensure that the marginal cost curve cuts the marginal revenue curve from below. Moreover, the marginal cost must not be falling more rapidly than marginal revenue. To sum up, equilibrium under monopoly can occur whether marginal costs are rising, falling or constant. The only situation in which monopoly equilibrium is impossible is where marginal costs are falling more swiftly than marginal revenue.

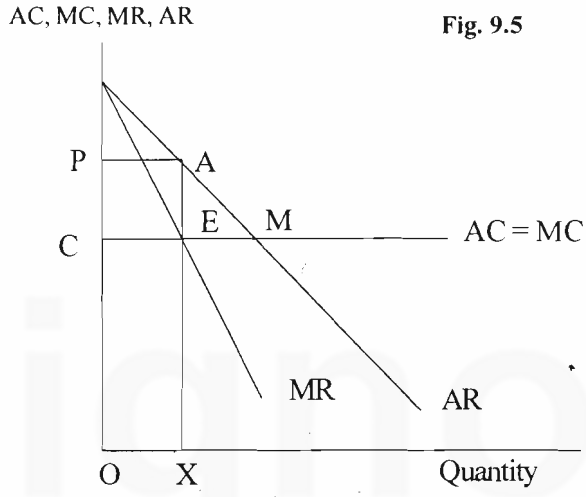


Fig. 9.5

Fig. 9.5: Shows a special case where MC =AC = constant. The monopolists equilibrium is still at point E where MC cuts MR from below. The firm produces OX output and sells it at OP (= AR) price.

9.6.6 Is Monopoly an Inefficient Type of Market Structure?

Monopoly is an inefficient type of market structure in the sense that gains to the firm from a monopoly position, its ability to charge higher price than the marginal cost, would be less than the loss of consumer's surplus due to rise in price. The difference between the gains to the firm and the loss of consumer's surplus on account of higher price under monopoly is termed as dead-weight loss.

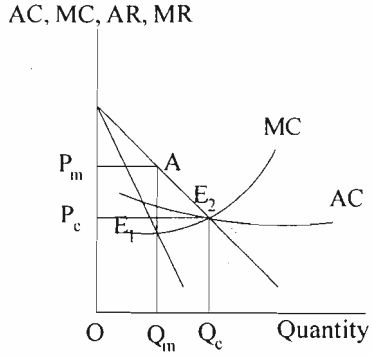


Fig. 9.6

Fig.9.6: There are two prices, the competitive price, P_c and the monopolist's price, P_m. The loss of consumer's surplus attributable to the higher monopoly price is the sum of rectangular area, $P_m AE_1 AP_c$ and the triangle E_1E_2A. But gain to the monopolist is confined to the rectangle only. The consumer 's loss equal to the triangle referred to above is 'dead weight loss'. Also note that the consumer is worse off in one more sense : she has to remain contented with a smaller quantity under monopoly than was available in perfect competition.

9.7 PRICE DISCRIMINATION UNDER MONOPOLY

A monopolist can increase profit through price discrimination. Price discrimination occurs when the monopolist charges different prices for different units of a homogeneous commodity. That is, in general, price discrimination means that a firm charges two or more prices for the same product at the same time. If the product is not homogeneous, it can also mean that the differences in the prices of a firm's products are greater than differences in their cost of production.

9.7.1 Degrees of Price Discrimination

How far can a monopolist go on charging different prices for the same product? What is the limit? The degree of price discrimination sets the limits within which a monopolists can charge different prices. We can distinguish between 3 degrees of price discrimination. Look into the following discussion in this regard.

First Degree Price Discrimination

Discrimination of the first degree defines the maximum possible limit of a monopolist in charging different prices for the same product. By setting the price accordingly, the monopolist extracts from each consumer the entire amount of consumer surplus. For this reason, first degree discrimination is also called perfect discrimination.

Second Degree Price Discrimination

In case of **second degree of price discrimination**, the monopoly firm sells to the consumer a block of output at one price and additional block of output at a different price. In this way, the monopolist captures a part of the consumers' surpluses, but not all of them. The schedules of rates charged by public utilities provide examples of second-degree discrimination. Observe the rates you pay for electricity and water supply. Charges are in different slab rates. Such structures come close to the underlying idea behind second degree price discrimination.

Third Degree Price Discrimination

Third-degree price discrimination means that the monopolist divides customers into two or more groups, charging a different price to each group of customers. Each class is treated as a separate market. This is the most common type of price discrimination and form the subject matter of study in the following pages.

9.7.2 Conditions for Price Discrimination

Let us consider the conditions under which price discrimination it is possible.

The fundamental condition which must be fulfilled, if discrimination is to take place, is that there can be no possibility of resale of product from one consumer to another. The discrimination may be owing to consumers' peculiarities, may be based on the nature of the good, or may be because of distances and tariff or other barriers. This may happen if consumers in one part of the market do not know that prices are lower in another, that is, lack of communication in different parts of the market or if the consumer has an irrational feeling that she is paying a higher price for a better good. The nature of the good may facilitate discrimination. For example, if the good in question is a direct service, hair cuts or manicures, services of a doctor or of a teacher. It is not possible for these services to be resold by a consumer who is charged a lower price to one who pays a higher price.

Conditions under which it is Profitable to Charge Different Prices

If above conditions are met, it is possible for the monopolist to engage in price discrimination. However, it will be profitable to discriminate only if elasticity of demand in the one market (e_1) is different from elasticity of demand in the other (e_2), that is, $e_1 \neq e_2$. This needs further elaboration.

Assume a monopolist who divides her market into two distinct segments, market 1 and market 2 and has a fixed quantity of the product to be sold in one or both of the two markets. How should the given quantity be allocated between the markets so as to maximise the monopolist's total revenue? The total revenue will be maximised from the sale of a given total quantity of a commodity if marginal revenue in market A is equal to the marginal revenue in market B, that is, $MR_A = MR_B$.

Now it is simple to understand that, if the coefficient of price elasticity of demand (e) is same in both the markets, it will not be profitable to attempt to separate the markets. Since marginal revenue in the two markets is to be equalised for optimum allocation of given output and when elasticity is the same in both the markets, that is, $e_1 = e_2$, it follows that

$$MR_1 = P_1 (1 - 1 / e_1) = MR_2 = P_2 (1 - 1 / e_2)$$

Thus $P_1 = P_2$ and so, there is no price discrimination.

However, if elasticity of demand is different, price discrimination will pay. When $e_1 < e_2$, a rise in price in market 1 will not cause much fall in demand, whereas, a reduction in price in market 2 will add more to revenue than it does to cost. It will pay the monopolist who is charging the single monopoly price to transfer goods from market 1 with the inelastic demand to market 2 with the elastic demand. The loss of revenue from reducing sales in market 1 will be smaller than the gain in revenue from expanding sales in the market 2.

Consequently, price will rise in market 1 and fall in market 2. Higher price occurs in the market having more inelastic demand curve. The result may be shown with the help of formula :

$MR = P (1 - 1/e)$. For optimum allocation of given output in the two markets, $MR_1 = MR_2$, that is,

$$P_1 (1 - 1 / e_1) = P_2 (1 - 1 / e_2)$$

Thus if e_1 is greater than e_2, then P_1 must be less than P_2 in order to maintain the equality between the two sides of the above expression. So, the price is higher in the market which has less-elastic demand. However, the above result hold good subject to the following conditions:

a) that the coefficient of elasticity of demand is greater than one in both the markets, and

b) it is profitable to sell in both the market, that is, marginal revenue in both the market is greater than marginal cost.

9.7.3 Equilibrium under Price Discrimination

To determine the profit maximising total output to be produced by the price discriminatory monopolist, the marginal revenue should be equal to marginal cost of producing the whole output. This is explained with help of Fig.9.7. Let the average and marginal revenue curves of the firm for two separate markets having different

elasticities of demand be AR_1, AR_2 and MR_1 and MR_2 respectively. The total demand curve D is obtained by horizontal summation of individual AR curves in the two markets. The aggregate marginal revenue (MR) is the horizontal summation of MR_1 and MR_2. MC is the marginal cost curve.

Fig. 9.7

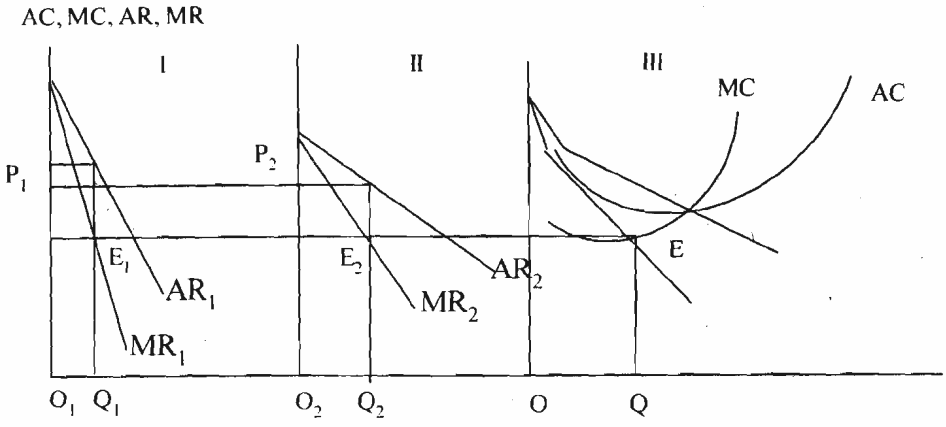


Fig. 9.7: Panel I of the figure $AR_1 = D_1$ and MR_1 curves in the market A. The corresponding curves for market B are shown in panel II. Notice that demand curve in market B is more elastic of the two. The third panel shows aggregate demand curve $D = D_1 + D_2$ and aggregate marginal revenue curve associated with D. The marginal cost curve of the firm, MC cuts the MR at point E and therefore, firms total output is OQ in panel III. The firm equates marginal revenue in each market to this equilibrium level of MC = MR = QE. Thus, it sells OQ_1 and OQ_2 quantities in markets A and B respectively. The prices in two markets are OPA and OPB. The market A, where elasticity of demand is lower, pays a higher price compared to market B.

To maximise her profits, the price discriminating monopolist has to take two decisions:

a) How much total output to produce?

b) How much to sell in each market and at what price?

The equilibrium level output of the price discriminating monopolist is at point where MC equals aggregate MR. Since elasticities in two market segments are different, her profit are maximised if $MR_1 = MR_2 = MR = MC$. The producer charges less price in market where demand is more elastic.

Check Your Progress 2

1) What is price discrimination ?

.....

.....

.....

2) What are the necessary conditions for price discrimination ?

.....

.....

.....

3) Explain how will a profit maximising discriminating monopolist distribute output in different markets? What prices will she charge in those markets?

.....

.....

.....

4) Discuss the relationship between price elasticity of demand and price discrimination.

.....

.....

.....

9.8 EQUILIBRIUM PRICE AND OUTPUT OF A PUBLIC MONOPOLY

So far we have discussed the behaviour of a private monopolist whose objective is to maximise her profits, given the economic and technical constraints. In this section, we analyse the behaviour of a public monopoly - a firm owned and controlled by the government. The objective of a public monopoly is to provide more output and charge lower price so as to increase the welfare of people. The optimal pricing and output decisions by such an undertaking is not based on profit or sales maximisation principles but on maximisation of welfare.

Average-cost-pricing and marginal-cost-pricing are the two possible options for the determination of output and price by a public utility firm. In fact, these two options can become policy guidelines for the government for price regulation of a private monopoly firm as well.

There is a need to regulate monopoly because monopolists have ability to restrict output and raise prices of their product and this way earn super normal profits. Such behavior increases inequalities in the distribution of income and wealth, leads to exploitation of the consumers and also causes inefficiency in allocation of resources. A net result of all these actions is reduction of consumer welfare in the society. Therefore, the main objective behind regulation of monopoly is the maximisation of welfare. A monopoly may be regulated either through fixation of a maximum price that a monopolist may charge or appropriate taxation policy. In this unit, we are concerned with price regulation of monopoly only. The issues involved in average and marginal cost pricing discussed here are useful in fixing of prices in a public utility as well.

9.8.1 Marginal Cost Pricing

As discussed above, a monopolist sets price of its product higher than marginal cost, that is , $P > MC$. As shown in Fig.9.6, monopolist maximises profit at OQ_m level of output and charges OPm price. The government may decide to regulate a monopoly by fixing maximum price that equals marginal cost of production. Thus, the monopolist will be forced to raise the output to OQ_c and the price corresponding to this level of output will be OP_c - price, which would have prevailed had the market been perfectly competitive. Such a price would ensure efficiency in allocation of resources as well, since it is equal to marginal cost. It also enhances welfare of

the consumers, as they get larger output at lower price. The consumer's surplus under non-regulated monopoly is shown by area of triangle AP_mE_1. When marginal cost pricing is enforced, it rises to AP_cE_2 (Fig. 9.6).

It may be noted that given the conditions of demand and cost as presented in the above figure, marginal cost pricing may still allow a monopolist to earn super normal profits as the price may still be higher than the average cost. This is a case of 'capacity-constrained situation', that is, the demand for the product is quite high as compared to the production capacity. But, in different a situation, when there is excess capacity, marginal cost pricing results in direct loss to the firm as its average cost is higher than marginal cost. Thus the firm will produce marginal cost price output only if it is compensated by the government for the direct loss at this level of production.

9.8.2 Average Cost Pricing

The aim of the public policy is to regulate monopoly in such a manner that it is possible to provide maximum output at minimum price. One policy option in such a context is to fix price according to the average cost, that is, at a point where AR =AC. This allows the firm to earn normal profit. In case of capacity-constrained situation, average cost pricing leads to higher output and lower price. This means there will be higher level of consumer's surplus compared to marginal cost pricing. However, in excess capacity situation, there shall be a somewhat higher price with average cost pricing but there shall be no direct loss to the producer as P =AC. Marginal cost pricing is adopted to reach full economic efficiency or maximum social welfare. But in case of excess capacity, where AC > MC, marginal cost pricing necessitates state subsidies to induce the monopolist to stay in the market.

9.8.3 A Note on Mark-up Pricing

It is suggested that, in practical life, prices are not fixed by marginal analysis, that is, by the use of marginal revenue and marginal cost concepts. An alternative approach is to set the prices in accordance with the average cost principle.

The firm sets a price equal to its total average cost which includes a certain net profit margin, that is,

$$P = AVC + GPM,$$

where P is the price, AVC is the average variable cost, and GPM is the gross profit margin which include average fixed cost and net profit margin.

The purpose of this note is to show that average cost principle and marginal analysis would give the same long-run profit maximisation solution. The setting of the price on the basis of the average cost principles incorporates an estimation of the elasticity of demand in the long-run equilibrium. Recall that the necessary condition for profit maximisation is MC = MR. It has already been proved that MR = P (1 -1/e). Given that MC > 0, MR must be positive for profit maximisation. This implies e >1. Provided that AVC is constant over the relevant range of output, that is, AVC = MC. For equilibrium, AVC = MR, that is, AVC = P (1-1/e) = P{(e-1)/e}. In other words, P = AVC{e/(e-1)}. Given that e >1, we may write {e/(e-1)} = (1+k), where k >0. Therefore, P = AVC (1+k), where k is the gross profit margin. For example, if the firm sets a 20 per cent of AVC as its profit margin, we have (1+k) = 1(t) 0.20 = {e/(e-1)}. Thus, the elasticity of demand is 6. Setting a gross profit margin is equivalent to estimating the price elasticity of demand and applying marginalist analysis. So when the businessman establishes

a mark-up on average costs, he is guessing at the coefficient of price elasticity of demand.

Check Your Progress 3

1) In what respect a public monopoly is different from a private monopoly? How does a public monopoly firm make pricing and output decisions?

.....

.....

.....

.....

2) Write a short note on mark-up pricing.

.....

.....

.....

9.9 LET US SUM UP

We commenced this unit with a discussion of factors which give rise to monopoly in the market. In some cases it may be a result of 'march of technology', but more often, it arises on account of legal provisions like patents, if it has not been created by law (public monopoly). The subsequent discussion was followed with a description of demand and marginal revenue curves of such a firm and its equilibrium. Then we took up questions like whether a monopoly always makes profits or, benefits a rise in price. Nature of the supply curve, efficiency in production etc., were the other aspects we touched upon.

Price discrimination is another aspect of a monopolist's behaviour and it has been dwelt upon in Section 9.7. We have examined two alternative approaches to price determination in a public monopoly, which can be used to regulate prices in private monopolies as well. We have concluded the unit by demonstrating equivalence of mark-up pricing with prior estimation of price elasticity of demand.

9.10 KEY WORDS

Monopoly : Existence of only one producer / seller of a good in the market.

Legal monopoly : Monopoly created by some legal provisions.

Technological Monopoly : Technology is such that it makes feasible for the production by one producer only.

Price Discrimination : Ability of a firm to charge different prices from different consumers for the same product.

Public Monopoly : A government sets up monopoly to ensure greater availability of some good or service at a reasonable price.

Marginal Cost Pricing : Pricing policy which covers only the marginal cost of production.

Average Cost Pricing : Policy under which the price covers average cost of production.

Mark-up Pricing : Pricing technique which fixes price at average variable cost plus a proportion of the fixed costs and a profit margin.

9.11 SOME USEFUL BOOKS

Varian, Hal (1995). *Intermediate Microeconomics*, W.W. Norton & Co, New York

Pindyck, Robert S. and Daniel Rubinfeld (1989) *Microeconomics*, Collier Macmillan, London

9.12 HINTS / ANSWERS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) Read Section 9.1 and 9.2 and answer.

2) Read Section 9.1, 9.2 and answer.

3) Read Section 9.3 and answer.

4) Read Section 9.3 and answer.

5) Read Section 9.3 and answer.

6) Read 9.4 and answer.

7) Read Section 9.3, 9.4 and answer.

8) Read Section 9.5 and answer.

Check Your Progress 2

1) Read Section 9.7 and answer.

2.) Read Sub-section 9.7.2 and answer.

3) Read Sub-section 9.7.3 and answer.

Check Your Progress 3

1) Read Sub-section 9.8.1, 9.8.2 and answer.

2) Read Sub-section 9.8.3 and answer.

UNIT 10 THEORY OF MONOPOLISTIC COMPETITION

Structure

10.0 OBJECTIVES

This unit aims at acquainting you with two market forms, which lie between the polar cases of perfect competition and monopoly. The cases we intend to deal with are: one where a large number of firms exist, and another where there are only a few firms. After studying this unit, you should be able to:

- define monopolistic competition;
- explain the concept of product differentiation;
- describe the main decision variables of monopolistic competition;
- compare monopolistic competition with perfect competition;
- analyse the economic effects of monopolistic competition; and
- define oligopoly and distinguish it from monopolistic competition proper.

10.1 INTRODUCTION

In earlier units (Units 8 and 9), we discussed the pricing and output decisions under two important market structures — perfect competition and monopoly. Both these are extreme forms in the sense that they are somewhat far from reality. The conditions

of perfect competition such as large number of producers producing homogenous products having perfect knowledge of the market are seldom fulfilled in the real world. Similarly, monopoly, or a single producer producing a commodity, which does not have close substitute is also rarely found. The intermediate situation described as imperfect or monopolistic competition is more prevalent in a modern economy. Imperfect competition includes all forms of market structure other than perfect competition and monopoly.

In today's world, generally, a firm is neither a price maker in the sense that it has full control over the price and need not worry about competitor nor a price taker, in the sense of no control over the price. The theory of monopolistic competition combines the elements of perfect competition and pure monopoly to reflect a realistic market structure. As will be explained below, firm operating under monopolistic competition enjoys some degree of monopoly. But it faces intense competition from other firms producing products that are close substitutes.

The monopolistic competition recognises the fact that many policy options such as changing expenditure on advertisement and sales promotion are available to firms. This market structure not only explains pricing and output decisions of a firm but also takes into account other important business policy variables such as advertising and sales promotion, product development and quality variation of the products.

Imperfect competition covers a wide range of market structures, which may be broadly divided into:

i) Oligopoly, or competition among few, and

ii) Competition among many sellers, or the 'large group' case, popularly known as monopolistic competition.

In case of oligopoly, there are few firms producing a product. The most significant feature of oligopoly is the interdependence among different firms in respect of the price and other policy decisions. In other words, demand for the product of one firm depends not only on its price but also on the prices charged by other competing firms. Any change in the policy of one firm affects the rival firms significantly and they react. Without making some assumptions regarding the behaviour or reaction of the rivals, it is not possible to predict the impact of change in the price or other variables on the demand for the product of the firm. As different types of behaviour or reactions by different firms can be expected among oligopolists, there cannot be just a single model of oligopoly.

But the type of interdependence seen is absent in the case of 'large group' or monopolistic competition. In such kind of a market structure, there are a large number of producers, change in the price or other policy variables do not affect rival firms significantly. Thus, they do not react. Consequently, the policy decisions of firms such as determination of price and output can be treated as independent of one another. This unit primarily deals with this type of "large group" case.

10.2 MEANING OF MONOPOLISTIC COMPETITION

Monopolistic competition refers to a market structure in which there are many firms selling closely related, similar but not identical products. The main assumptions of monopolistic competition are the same as those of pure competition (see Unit 8) except the homogenous product. These are listed below:

i) **Large Number of Sellers and Buyers**: In this market, there are large number of producers and buyers. The number of producers is large enough to allow individual producer taking policy decisions regarding pricing, output and other variables, independent of others, without fear of any reaction. In other words, action of one producer does not affect others significantly. As a result they do not react. This requirement is similar to that of the perfect competition.

ii) **Differentiated Products that are Close Substitutes**: The producers under monopolistic competition produce differentiated products. But the product offered by each producer is a close substitute of another. By product differentiation we mean that consumers have their own rankings or preferences for the products produced by different firms based on physical attributes of products or on her own perception about products. This point is discussed in more detail later in this unit (see Section 10.3).

iii) **Free Entry and Exit from the Market**: There is freedom for potential sellers to join the market. Similarly, there are no restrictions on the existing firms to leave the market.

iv) **Perfect Knowledge of the Market**: The firm is assumed to behave as if it knew its demand and cost curves with certainty.

v) **Uniform Cost and Demand Conditions for all the Firms:** This characteristic is due to Chamberlain who assumed that both demand and cost curves for all 'products' are uniform throughout the group. The implication of this assumption is (a) that consumers' preferences are evenly distributed among different sellers, and (b) that the differences between the products do not give rise to differences in costs.

10.3 PRODUCT DIFFERENTIATION

In comparison to perfect competition, monopolistic competition is characterised by heterogeneity of products. Thus, consumers have definite preferences for the particular variety or brand of products offered for sale by sellers. The products are said to be differentiated, if consumers distinguish the product of one producer from that of another while making their purchases. For example, consumers have preferences for a particular brand of toothpaste. They are willing to pay somewhat higher price for their preferred brand. In other words, the products of different firms under monopolistic competition are treated by consumers as different commodities. This particular characteristic lends a type of 'monopoly power' to the firms.

Product differentiation can be 'real' in the sense that the inherent characteristics of the products are different. For example, materials used, workmanship, durability and strength may not be alike. The product may be different in size, colour, shape, style and packing. The location of the seller, the qualities like promptness, courtesy, good delivery of service, favourable credit terms could be sources of real product differentiation.

Fancied differentiation is created in the mind of the consumer through advertisement, differences in packaging and design, or by brand name. The purpose of product differentiation is to make the product unique in the mind of the consumer. This special status of the commodity manufactured by a producer, ultimately, persuades the consumer to pay a relatively higher price for it.

The element of product differentiation causes some consumers preferring the products of a particular firm to that of others. Each firm obtains a kind of limited monopoly for its unique version of the product. It gives the firm limited influence over the price it charges for the output and in a restricted sense makes it a 'price maker'. However, the monopolistically competitive firm faces stiff competition from the close substitutes offered by others. So, its control over the price is limited. The price elasticity of demand between products of monopolistically competitive firms is quite high, though not infinite, as under perfect competition.

On account of product differentiation, monopolistic competition provides fertile ground for non-price competition among firms. They compete with one another partially upon price and partially upon non-price modes like product quality, service and other conditions of sales, and sales promotion. A firm can simultaneously undertake three strategies for influencing its sales volume. First, it can change the price charged - the strategy of price competition. Second, it can modify the nature of its product - the strategy of product variation. Third, it can revise its sales promotion tactics - the strategy of promotional competition. The first strategy represents an attempt to move along the demand curve whereas the last two involve an attempt to shift the demand curve faced by a firm.

Check Your Progress 1

1) Define monopolistic competition and give a few examples of it.

.....

.....

.....

2) Identify the competitive and monopolistic elements of monopolistic competition.

.....

.....

.....

3) Which of the following assumptions of perfect competition does not apply to monopolistic competition?

   i) Many buyers and sellers

   ii) Free entry and exit

   iii) Homogenous product

   iv) Both (ii) and (iii)

4) In monopolistic competition, we have

   a) few firms selling a differentiated product;

   b) many firms selling a homogenous product;

   c) few firms selling a homogenous product, or

   d) many firms selling a differentiated product.

5) Product differentiation means that consumers do not view the product of one firm as exactly identical to the product of another firm. Elaborate.

.....

.....

.....

10.4 PRICE AND OUTPUT DECISIONS OF MONOPOLISTICALLY COMPETITIVE FIRMS

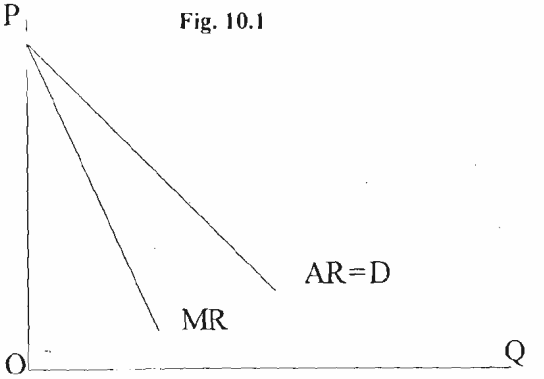
10.4.1 Demand Curve of Monopolistically Competitive Firms

The demand curve faced by a monopolistically competitive firm slopes downwards and is flatter than that under pure monopoly. This is because of the brand loyalty of consumers created by product differentiation. As stated above, the monopolistic competition combines the elements of both perfect competition and monopoly.

On account of product differentiation, a monopolistically competitive firm has a small measure of discretion, that is, monopoly power in determining the price of its product. If the firm sets its price somewhat below the prices of other firms, some customers are induced to buy its product when it is a good substitute for the products of its competitors. But some consumers continue to buy the products of other firms even at a relatively higher price because of their brand loyalty. On the other hand, increase in price by a firm results in significant decline in sales as many of its customers may switch over to low-priced substitutes. But some continue to buy the product of this firm even at a relatively higher price because they rate its product higher. Therefore, it may be said that the firm cannot raise prices without losing sales. However, it cannot gain sales without charging a lower price. Of course, it does not lose all the customers by raising the price. Thus, the firm's demand curve is downward sloping but not perfectly elastic. Moreover, the presence of close substitutes makes the demand curve of a particular firm highly elastic. The demand cure for such a firm is shown in Fig. 10.1.

It may be recapitulated that, when the firm's demand curve is downward sloping, its marginal revenue curve does not coincide with average revenue (AR) curve and lies below it. If the firm's AR curve is assumed to be linear and downward sloping, MR curve will also be linear and downward sloping and will have a slope twice that of AR curve as shown in the above figure.

Fig. 10.1



As depicted in Fig. 10.1, a firm in a monopolistically competitive industry faces a highly elastic but negatively sloped demand curve for its product. The larger the number of competitors and weaker the product differentiation, the greater the price elasticity of demand for a firm's product.

10.4.2 Short - run Equilibrium of a Monopolistic Competitive Firm

A firm is to decide the fixation of price of its product in order to maximise profit. While doing this it has already decided about the type or quality of the product to be produced and the amount of selling cost to be incurred. The average cost curve of the firm is taken to be U- shaped. Given AR, MR and AC and MC curves, the profits are maximised at output level where MC and MR are equal, and MC intersects MR from below. This is illustrated in Figures 10.2 a. and 10.2b. Short-run profits are maximised at output level OQ as given by the intersection of MR and MC curves. It is a stable equilibrium because MC exceeds MR for an output greater than OQ. At this equilibrium output, the maximum price as given by AR curve is OP. Since AR is greater than AC, the firm is earning super-normal profit equal to the area $PRNP_1$.

Fig. 10.2 (a)

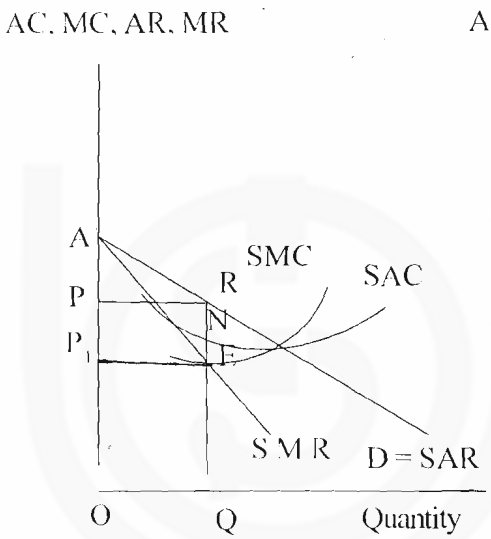


Fig. 10.2 (b)

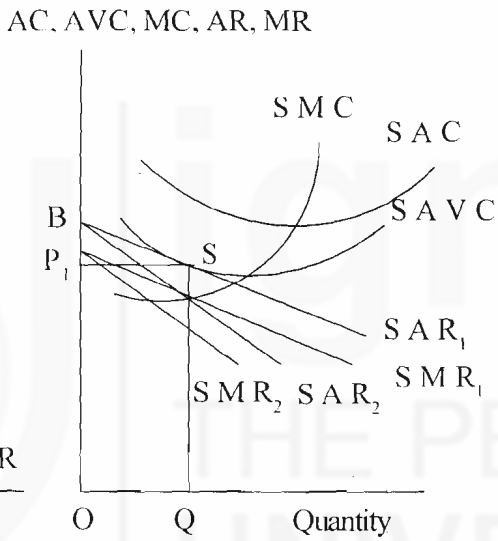


Fig. 10.2 (a): Shows short run equilibrium of a monopolistically competing firm. Short run AR = D can lie above short run AC. But, equilibrium is, as usual, determined at point E where rising MC (SMC) cuts falling MR (SMR). Output at that point is OQ. The firm is able to charge OP price for this quantity. Here, total revenue exceeds total costs by $PRNP_1$, which is called 'extra normal profit'.

Fig.10.2 (b): Shows the case when firm is not able to cover average cost in the short run, but continues to be in the market so long as its average variable costs are covered. If however, the demand curve falls below the SAVC, the firm will have to shut down. This is the reason we call this output level as the 'shutdown point' S.

However, if the demand is weak, the monopolistically competitive firm may not be able to make a super-normal profit, or even a normal profit. In case the demand is too weak to recover variable costs, the firm would prefer to shut down its operations. Like perfect competition, a monopolistically competitive firm remains in the business in the short period if:

i) AR > AC, earning supernormal profits, or

ii) AR = AC, making just normal profits, or

iii) AR< AC but greater than, or equal to the AVC, thus, minimising losses associated with shut down.

10.4.3 Long-run Equilibrium of a Monopolistic Competitive Firm

You have learnt in Unit 8 that with free entry and exit, firms in a perfectly competitive industry are able to earn only normal profit in the long period. Similar forces are present under monopolistic competition also. The process of long-run adjustments in a monopolistic competitive 'group' is analogous to that of perfectly competitive industry. When short-run super-normal profit (economic profit/extra profit) exists, new firms are motivated to enter the 'group'. The entry of new rivals, assuming constant market demand for the product of the industry (group) will cause the demand curve of each firm to shift to the left. Each firm will have a smaller market share since more firms will be dividing the relatively constant total market among themselves. Moreover, the demand curve is likely to become somewhat more elastic owing to the presence of a larger number of close substitutes. These shifts in demand tend to narrow profit margins and cause economic profit to dissipate. There exists a tendency for economic profits to be eliminated in the long run.

Fig. 10.3

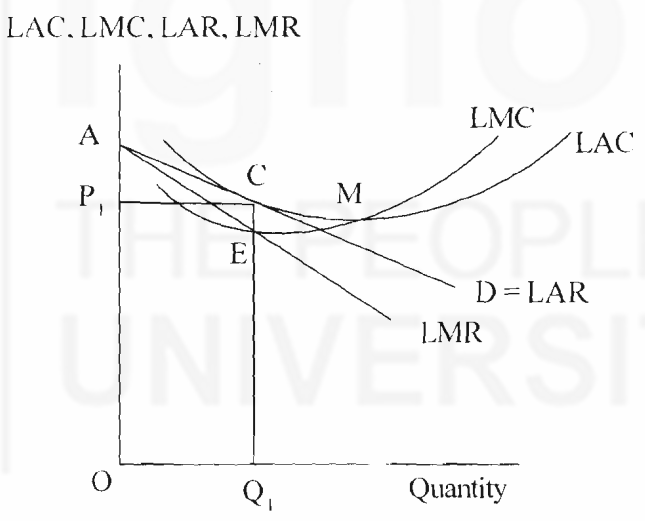


Fig.10.3 illustrates the long-run equilibrium position of the representative firm. AR curve is shown tangent to LAC curve at the profit-maximising output. Output Q_1 is the long-run equilibrium output, and price P_1 is the long-run equilibrium price. Since price equals LAC at Q_1 units of output, the firm is just covering all its costs, including implicit and opportunity costs. Since profits are just normal, there will be no further entry into the industry. The equilibrium is stable, since any deviation from an output of Q_1 and price of P_1 will yield revenues, which are insufficient to cover all production costs including a normal profit.

Check Your Progress 2

1) Will the demand curve for a firm under monopolistic competition be horizontal or downward sloping?

.....

.....

.....

2) Show the short-run equilibrium position for a firm under monopolistic competition such that the firm makes some above normal profits. Explain why it is a temporary equilibrium.

The short-run equilibrium level of output for a monopolistic competitor is given by the point where

a) P = SMC

b) P = SAC

c) MR curve intersects SMC

d) MR curve intersects SMC curve from below and P > or = AVC

3) Show that in the long-run equilibrium in monopolistic competition each firm makes zero profits? (Reference is to economic or super normal profits)

.....

.....

.....

.....

10.5 COMPARISON OF MONOPOLISTIC COMPETITION WITH PERFECT COMPETITION

The long-run equilibrium of monopolistic competitive firm as well as perfectly competitive firm is defined by the point of tangency of demand curve to LAC curve. At this point, in the case of monopolistic competition, MC = MR and AC = P, but P > MC. On the other hand, in perfect competition, the long-run equilibrium condition is MC = MR = AC = P. An important implication of the long-run equilibrium of the firm under monopolistic competition as given by the tangency between AR and AC curves is that equilibrium output will necessarily be less than the least-cost output of the firm. This is because a downward-sloping AR curve can be tangent to U-shaped cost curve only at some point to the left of the lowest point of AC curve. This implies that even in the long-run equilibrium position, all the economies of scale are not fully exploited by the firm. The firm will always have some unutilised or 'excess capacity'. In Fig. 10.4, excess capacity is represented by $Q Q_1$ amount of output.

A firm operating under monopolistic competition will have some excess capacity in the long run. The firm may be using the most economical plant for producing the output that it can sell but without exploiting fully the economies of scale, i.e., the plant will not be used to its optimum capacity.

Fig. 10.4

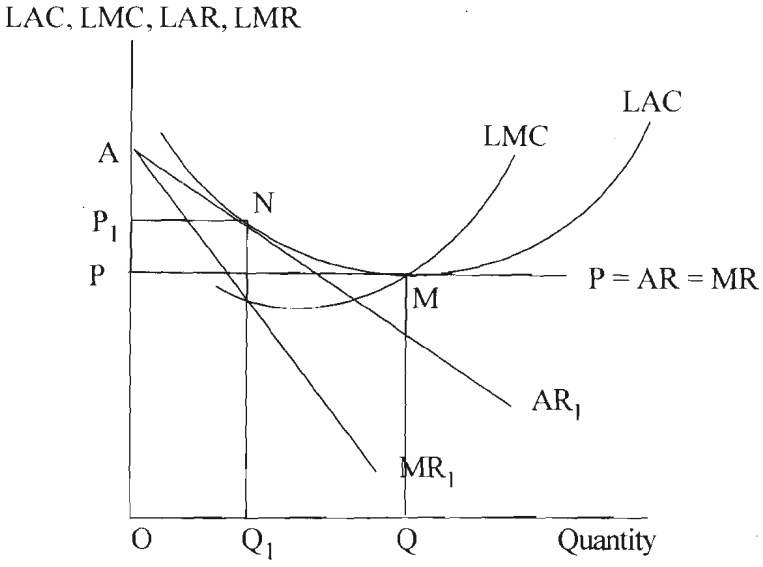


Fig. 10.4: **Line PP = AR = MR shows the perfect competition demand curve, while the line AR_1 shows the same for monopolistic competition. The former is tangent to the long run average cost curve LAC at its lowest point M, AR_1 is tangent to this LAC at a point which is higher, the point N.Thus, the firm under monopolistic competition produces an output OQ_1 that is smaller than the competitive output OQ.**

10.6 ECONOMIC EFFECTS OF MONOPOLISTIC COMPETITION

It will be useful to point out two major economic consequences of monopolistic competition at this point. First, the allocation of resources in the economy is non-optimal since price is greater than marginal cost. For optimal allocation of resources, price must be equal to marginal cost in the case of every product.

Second, monopolistic competition gives rise to a great deal of social waste arising out of selling cost or advertisement expenditure. Advertisement, which is essential and profitable from the point of view of an individual firm, is a waste from the social point of view. The social welfare would have gone up had the resources devoted to advertisement been used towards increasing the quantity or improving the quality of the goods supplied by the monopolistic competitive firms.

10.7 POLICY VARIABLES FOR A MONOPOLISTICALLY COMPETITIVE FIRM

The policy-related variables for a monopolistically competitive firm are price, product variation, and selling expenses. Product variation and selling expenses are referred to as non-price competitive efforts on the part of the firm to increase sales or make the demand curves that they face less elastic.

10.7.1 Selling Costs as a Policy Variable

Product differentiation is the basic feature of the monopolistic competitive market structure. Consequently selling costs occupy an important place for creating and

maintaining product differentiation in the minds of consumers. Let us remember that selling costs are incurred to adapt demand to what is being produced by the firm. On the other hand, production costs are incurred to adapt production to demand. Selling costs are incurred to bring the product to the notice of the consumers and to systematically make place for the product in the mind of consumers. The basic purpose and effect of advertising is to increase demand, that is, to shift the demand curve rightwards and also to make it less elastic. Initially, increases in selling costs may result in more than proportionate increase in demand, because, some minimum advertising outlay enables the firm to reach the public at large and to attract them to the product. However, subsequent increases in advertising expenditure may not induce the buyers as much, thus, may yield less than proportionate increase in demand. The effect of selling cost on the demand for the product is illustrated in Fig. 10.5.

Fig. 10.5

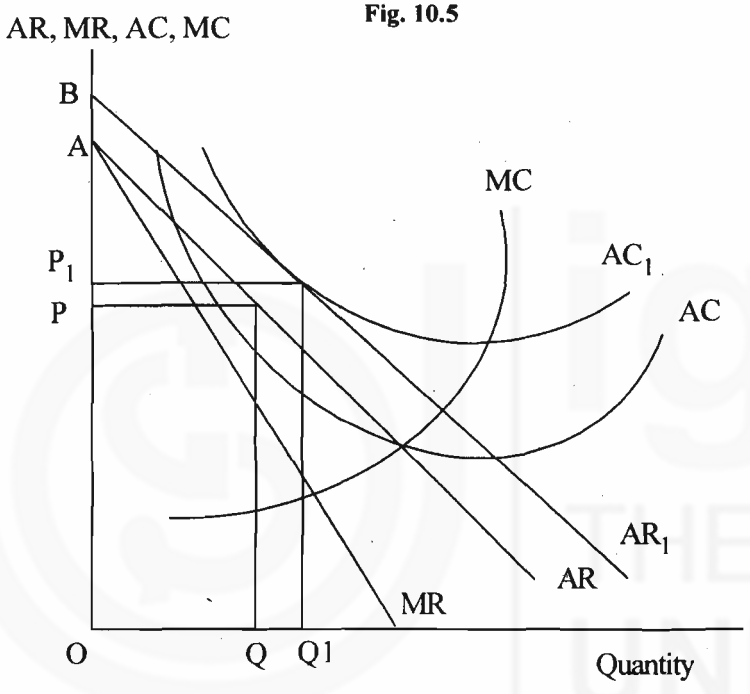


Fig.10.5 shows that selling costs shift the demand curve to the right. AR was the demand curve before the firm started incurring expenditure on sales efforts. It shifted to new position AR_1. You can check that AR_1 is less elastic at any given price OP_1.

Sales promotion adds to firm's costs on the one hand and shift its demand (AR curve) to right on the other. As long as the revenue increases by more than promotional expenditure, it would improve the profit position of the firm and advertising may be considered to be a wise move. This is depicted in Fig. 10.6. AR is the demand curve before any advertisement is undertaken and AC is the average costs, which do not include any advertisement costs. The initial equilibrium position where profits are being maximised is at output level OQ and price OP. Let us assume that firm spends some amount on advertising. The advertising cost is treated as fixed cost. AC_1 is the new average cost curve, which includes both average production cost and average advertising cost. As a result of sales promotion, the new demand AR_1 has shifted to the right of AR. It would be wise to incur expenditure on advertisement if profits are larger in the new equilibrium than they were in the original one. On the same line, the firm would repeat the examination whether increase in advertising expenditure leads to a further increase in profit or not. It can go on

increasing selling costs so long as each increment of advertising expenditure adds more revenue than costs. The profit will be 'maximum' only when additional revenue generated equals the extra amount spent on selling costs.

Please note that in the figure, the equilibrium output is determined at the point of intersection of MR and MC curves, whether the firm advertises or not. Selling costs do not affect MC, but a shift in AR (to AR_1) definitely implies that there shall be a new marginal revenue curve MR_1 and post-advertising equilibrium will take place at a point where old MC cuts MR_1.

Fig. 10.6

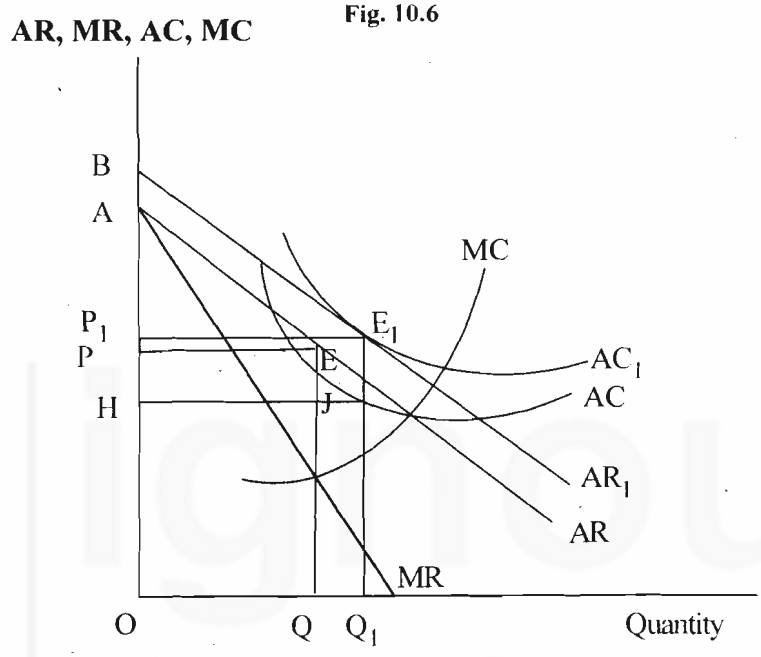


Fig. 10.6: As the firm incurs some selling expenditure, its average cost shifts from AC to AC_1. These selling costs are treated as fixed costs, therefore, the firm's marginal cost curves remains unchanged. The selling costs shift the demand curve for the firm's output upwards to AR_1. New total revenue of the firm will be $OP_1E_1Q_1$, which is larger than the old total revenue given by OPEQ. The firm can go on increasing the promotional outlay so long as this rise in total revenue exceeds the outlay HJE_1P_1, and not beyond that.

10.8 SUMMARY OF MONOPOLISTIC COMPETITION

In this section, we attempt to capture the monopolistic competition in brief. The distinguishing features of monopolistic competition are:

1) many firms selling differentiated, yet similar products;

2) the ability of each firm to influence its sale by changing its price;

3) downward sloping, but highly elastic, demand curves faced by the firms;

4) firms can enter and leave the industry with relative ease;

5) the actions of any one firm have a small effect upon rival firms; and

6) firms seek to maximise profit.

When large numbers of small firms sell a homogenous product, the market structure tends to be perfectly competitive, where large numbers of small firms sell differentiated product, the model of monopolistic competition prevails.

Monopolistically competitive firms may earn profits or incur losses in the short-run but free entry and exit limits the possibilities of earning economic profits and economic losses in the long run.

To maximise profits, monopolistically competitive firms pursue three strategies: price changes, product variation and variation in selling costs. Each possible combination of price, product variation and promotional outlay poses a different demand and cost situation for the firm. Therefore, to arrive at an output level and price which maximise profit is a complex issue in monopolistic competition.

If consumers accept homogenous products, a perfectly competitive market structure drives prices to the level of minimum average cost. Each firm operates at the minimum point of its short-run and long-run average cost curve and the firms, which fail to attain the lowest possible unit cost, are thrown out of the industry. Consequently, resources tend to be employed at maximum production efficiency in a perfectly competitive industry.

However, if consumers have preference for differentiated products, then monopolistically competitive conditions produce consumer's benefits in the sense that competition is strong enough to keep prices down close to cost and to eliminate excess profits. But monopolistically competitive firms tend to produce at less than optimum level since the demand curve is tangent to the downwards sloping portion of the long-run average cost curve.

10.9 OLIGOPOLY

Let us have a brief overview of the market structure known as **oligopoly**. Oligopoly is synonymous with competition among few. Markets are said to be oligopolistic wherever small number of firms supply the dominant share of an industry's total output.

The principal effect of there being a few firms is that it gives each firm a prominent market position such that its decisions and actions have significant repercussions on rival firms. If one firm announces a price change, introduces a new product, changes product design or steps up its advertising, other firms take note and chalk out their strategies to match the rivals' action. Note that each firm recognises that its best course of action depends on the action of its rival. Thus producers' actions become interdependent which extends to all facets of competition in an oligopoly market structure. The anticipation of actions and reactions of rival firms introduces a new and exceedingly complex dimension to the firm's decision process. The interdependence of the firms' actions, and consequently indeterminate demand curve is the key feature of oligopoly.

If the firms in an industry produce a standardised product, the industry is called a pure oligopoly. On the other hand, if few firms dominate the market for a differentiated product, the industry is called a differentiated oligopoly.

It is difficult, but not impossible, to enter an oligopolistic industry. Presence of substantial economies of scale, complex technology, production of highly advertised product by existing firms, work against the successful entry of new firms.

In an oligopolistic market structure, there is no clear-cut equilibrium position towards which all firms tend to move. In this market structure, several courses of action are open to firms selecting a competitive strategy. What firms will decide to do and how their rivals will react are open-ended. A number of outcomes are possible, and no single theory of oligopolistic behaviour can be postulated. It may be interesting to note that these days, most of the consumers products' markets may appear to be oligopolistic in nature. Analysis of strategic options available under such market condition will take us closer to understanding of actual working of the markets. But you are not exposed to such analyses presently.

Check Your Progress 3

1) In what respects the monopolistic competition differs from perfect competition? Explain with help of some suitable diagram.

.....

.....

.....

.....

2) What are the policy variables for a monopolistically competing firm? Bring out the effect of selling / promotional costs / advertising outlay on price and output under monopolistic competition.

.....

.....

.....

.....

3) What are the main characteristics of the oligopoly form of market?

.....

.....

.....

.....

10.10 LET US SUM UP

This unit was the second successive one that dealt with market forms, which are different from perfect competition. Unlike the preceding unit that was on monopoly, this one dealt with market forms that lie between the polar cases of perfect competition and monopoly. Here we consider two more market forms viz., monopolistic competition and oligopoly. Monopolistic competition was considered in greater detail than the latter and consequently more space was devoted to the former. The main characteristics of monopolistic competition were spelt out, and it was also seen as to how monopolistic competition compares with perfect competition. We then looked at some policy variables concerning monopolistic competition. We also touched upon, briefly, that usually a unique equilibrium does not exist under oligopoly.

10.11 SOME USEFUL BOOKS

Varian, Hal (1997), *Intermediate Microeconomics*, W.W. Norton & Co, New York

Pindyck, Robert S. and Daniel Rubinfeld (1989) *Microeconomics*, Collier Macmillan, London

10.12 ANSWERS / HINTS TO THE CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) Read Section 10.1, 10.2 and answer.

2) Read Section 10.2, 10.3 and answer.

3) (iii)

4) (iv)

5) Read Section 10.3 and answer.

Check Your Progress 2

1) Read Sub-section 10.4.1 and answer.

2) See Sub-section 10.4.2 and answer.

3) (iv)

4) Read Sub-section 10.4.3 and answer.

Check Your Progress 3

1) Read Section 10.5 and answer.

2) Read Section 10.7 and answer.

3) Read Section 10.9 and answer.

UNIT 11 MARGINAL PRODUCTIVITY

Structure

11.1 OBJECTIVES

This unit gives you an overview of marginal productivity theory of distribution. It is explained with reference to determination of reward for labour. After going through this unit, you will be able to understand and explain:

- marginal product of a factor;
- factor price determination as a special case of general theory of pricing;
- 'complete' distribution of output (or product exhaustion) under perfectly competitive market system;
- influences and actions of trade unions and intervention of the state in labour market; and
- differential earnings.

11.1 INTRODUCTION

We know that social output is a result of combined efforts of its members. Different members of the community participate in process of production in some way or other. They all expect to receive some reward for the efforts put in. What should be the basis for determining individual's share in total output? This question has engaged the attention of economists, social scientists and philosophers over the ages. Various views have been put forward. Marginal productivity theory of distribution is an operational version of functional approach to the problem of distribution of social product. Under this theory we try to equate marginal product of a factor with the rewards for that factor. Operation of various markets for us ensures that in the long run, each factor receives reward equals to and commensurate with its productivity. Moreover, if perfect competition prevails in product as well as factor market, theory of marginal productivity ensures that total output is completely distributed leaving no residuals. Though we have developed the concept of marginal productivity theory of distribution with reference to labour, determination of wage is discussed separately (Sections 11.7 – 11.11) to highlight some special features of labour market.

11.2 FACTORS OF PRODUCTION

Production needs cooperation of different people. Some of them work together at the work place. Others may not be physically present in the work place (factory) but make it possible for those in the work place to initiate the process of production in the first instance. Some contribute capital that buys plant and machinery and buildings, which house them. Still others make it possible for workers and machines to come together. Somewhere, some arrangements are made to procure raw materials, which are to be used to produce output in the factory. This way, we can say that production in today's world is possible when someone organises labour, land and capital to come together and produce. The organiser (entrepreneur) decides what those essential things are to start production. We in economics call them, "Factors of Production" by defining them, broadly, along the following lines:

i) **Land**

It is sum total of what physical resources have been endowed upon us by the nature. It includes all minerals, forestry, soil and water resources of the economy.

ii) **Labour**

The foremost factor of production is defined to include all human endeavour, both physical as well as mental, which transforms things given to us by nature into commodities that satisfy some human want.

iii) **Capital**

During the process of transformation of free gifts of nature into commodities, human labour modifies materials in such a manner that it becomes easier to produce commodities. All such modified materials are given the common name, capital. We customarily regard capital as a separate factor of production.

iv) **Enterpreneurship**

Broadly speaking, this is the special human ability to organise the production and bear all the attendant risks.

We shall have an opportunity to discuss more about these factors production later on. We pay more attention to labour in Section 11.7.1, while land, capital and entrepreneurship are discussed in Unit12.

11.3 CONCEPTS OF MARGINAL PRODUCT

Marginal product of factor is the increment in output, which realised as one more unit of the factor in question, is used (added) while amounts of all other factors remain unchanged. In other words, in an establishment with given number of workers and machines one more worker is added, the change in output because of the additional worker will be called marginal product of labour.

We can similarly define marginal product of any of the factors of production. The factor whose amount is varied is called the variable factor. Table 11.1 shows what happens to total output in a hypothetical factory with given plant and machinery etc., as the number of workers is engaged one by one.

Table 11.1 : Production in a Factory

Units of Variable factor Labour	Total Product	Marginal Product
0	0	0
1	10	10
2	18	8
3	24	6
4	28	4
5	30	2

11.3.1 Marginal Physical Product

Col. 3 of Table 11.1 above shows that as the first worker is engaged, 10 units are produced. The second worker adds 8 units to output and so on. Finally, the 5th worker adds just 2 units. These units of output are in physical terms. Therefore, we can call the Col. 3 as the Marginal Physical Product (MPP) of labour.

11.3.2 Marginal Value Product

When marginal physical product of a factor is sold in the market at going market price the money thus realised is called value of marginal product or marginal value product (VMP). Thus, it is

$$VMP = MPP \times \text{Price (of the commodity)}.$$

11.3.3 Marginal Revenue Product

Firm's interest may not be centred on units of physical output. Firm could be more interested in what happens to total revenue it earns. Thus we come to the notice of marginal revenue product (MRP), which is defined as addition to total revenue as an additional worker is engaged. Thus,

Marginal Revenue Product = Marginal Revenue × Marginal Physical Product,

or, MRP = MR × MPP

It can also be defined as rate of change of total revenue as the utilisation variable factor changes, i.e.,

MRP = (change in total revenue) ÷ (change in variable factor)

When perfect completion prevails in product market, firm can sell any number of units of a commodity it decides to produce at the going market price. The demand curve for its output is horizontal. Therefore, average revenue and marginal revenue are constant and equal to the price. This fact leads us to Table 11.2, which is a further development of Table 11.1. The market information about the price of a commodity helps us to calculate total revenue, VMP and MRP.

Table 11.2 : Various Magnitudes of Marginal Product

Units of Labour	Total Product	MPP	Product Price	Total Revenue	VMP	MRP
0	0	0	20	0	0	0
1	10	10	20	200	200	200
2	18	8	20	360	160	160
3	24	6	20	480	120	120
4	28	4	20	560	80	80
5	30	2	20	600	40	40

Note: Notice that VMP = MRP (compare Co.6 and Col.7). However, this is a result of our assumption of perfect competition only as will be shown in Section 11.5.3.

Check Your Progress 1

1) Explain the following concepts in about 50 words each.

   a) Marginal physical product

   b) Marginal value product

   c) Marginal revenue product

   .....

   .....

   .....

   .....

   .....

2) By employing 5 workers a firm produces 10 chairs a day and sells each chair at Rs. 50. If the firm employs one more worker, total production goes up to 12 chairs a day but it has to sell them at Rs. 45 each.

a) What is MPP of labour?

b) What is VMP of labour?

c) What is MRP of labour?

.....

.....

.....

.....

.....

11.4 DEMAND FOR A FACTOR OF PRODUCTION: THE DERIVED DEMAND

A firm demands factors of production to produce commodities, which it hopes to sell in the market. Therefore, one can say that firm's demand for a factor of production will depend upon the demand for what that factor can produce. When the firm is hiring, say, labour from the market, it always takes into account the implication of its action on total revenue as well as total cost. The second worker also adds a net surplus of Rs. 120 (Rs. 160 – Rs. 40) to the firm's accounts. The firm still goes on to hire more and more workers. The 5th worker too costs the firm Rs. 40. However, her MPP = 2 and firms revenue also rises by Rs. 40 only. Clearly the firm has no incentive to hire more workers. Why? Change in cost equals change in revenue. Hence the firm attains 'equilibrium'. If the wage rate had been higher, the firm would not have engaged even the 5th worker. We can draw a diagram showing MRP = VMP curve of the firm (Fig. 11.1)

Fig 11.1

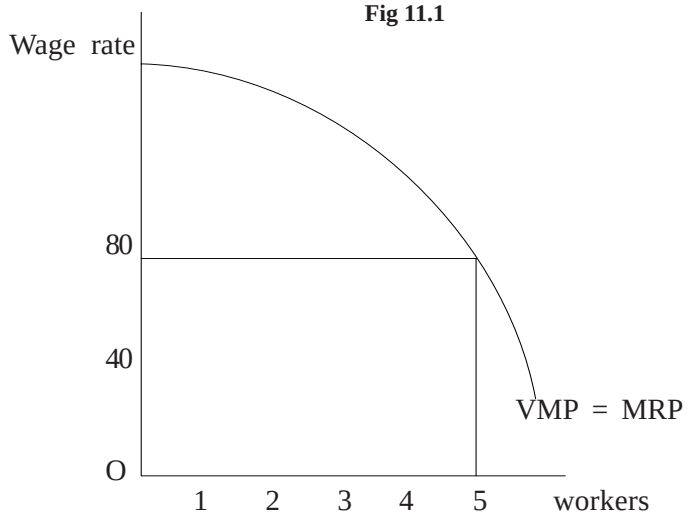


Fig 11.1: At wage rate of Rs. 40 per worker firm hires 5 workers. But if the wage rate were to rise to Rs. 80 the firm will have to stop at 4th worker. Why? Because at this wage rate, 5th worker will add Rs. 80 to the cost while her MPP of 2 units are worth only Rs. 40 in the market. The producer does not like to bear the 'loss' of Rs. 40.

We can gain an additional insight into producer's behaviour through MRP curve. The rectangular region under the curve represents wage cost to the firm. But the total area under the curve is the total revenue that has been earned by employing different units of labour. Therefore, triangular area (area between rectangle and the curve) can be called producer's surplus. Such profit maximising behaviour is another name for maximising this surplus. We can also say that when factor prices are brought into picture, VMP curve gives us the maximum number of units of the

factor that a firm will hire. Thus, MRP becomes the firms demand curve for the factor.

We can say, in general, the firm continues to hire a factor till its MRP drops down to the prices of the factor in the market. In the example above, so long as MRP exceeds the wage rate (i.e., Rs. 40), the firm will continue to hire more and more workers as every worker adds to its producer's surplus. However, it will not expand employment beyond the point MRP = W, i.e., 5 workers in our example above. Such a choice is because of sixth worker's MRP, which falls short of the wage rate. Employing the 6th worker will mean that the firm has to pay out of producer's surplus accumulated upto 5th worker. Thus, its net producer's surplus with six workers will be less than what it was when only 5 workers were employed. So firm earns maximum producers' surplus (at wage rate of Rs. 40) when it employs 5 workers.

We know that MRP curve becomes demand curve for a factor of production. Therefore, factors that affect MRP will also affect the demand curve for the factor. These factors are:

i) substitutability of a factor by some other factors;

ii) change in demand for the finished product. We know that the price of a product and hence its marginal revenue determines its MRP. Therefore, if demand for a product rise, so will its price. An upward shift of demand for product curve will raise marginal revenue curve. Given the marginal physical product of a factor, its MRP curve will shift to the right. This indicates a rise in demand for the factor;

iii) percentage of total cost incurred on the factor in question also affects price elasticity of its demand.

11.4.1 Market Demand for a Factor

We know that factors of production possess the virtue of being able to put to various alternative uses in the economic system. Labour is employed in all kinds of industrial activities. So what is the market demand for labour?

We can approach market demand in two steps: First, all the firms producing a commodity X, have their demand curves for, say, labour, determined by their respective MRP curves. We can aggregate MRP curves of all the firms to arrive at industry demand curve.

Fig. 11.2

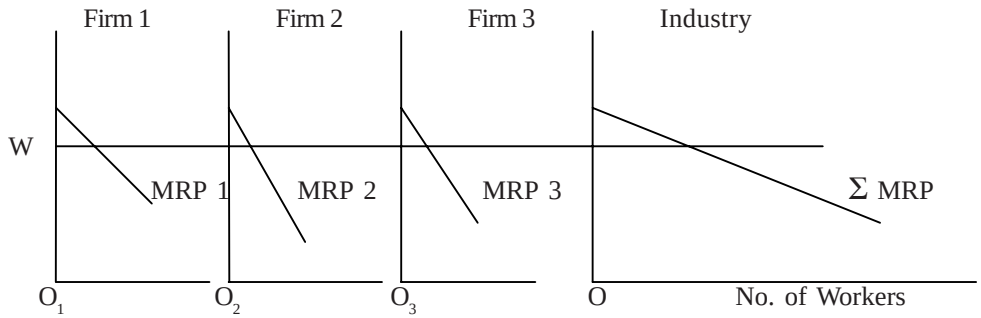


Fig. 11.2 Shows the derivation of one industry's demand curve for a factor. This is arrived at by horizontally adding up MRP curves of all the firms, at every wage rate, such as W. This industry MRP demand curve appears to be much flatter than MRP_s of its constituent firms.

Second, we aggregate, in similar manner, demand curves of various industries, in the economy to arrive at aggregate market demand curve for labour (or, for that matter, any other factor of production). This is shown Fig. 11.3 below.

Fig. 11.3

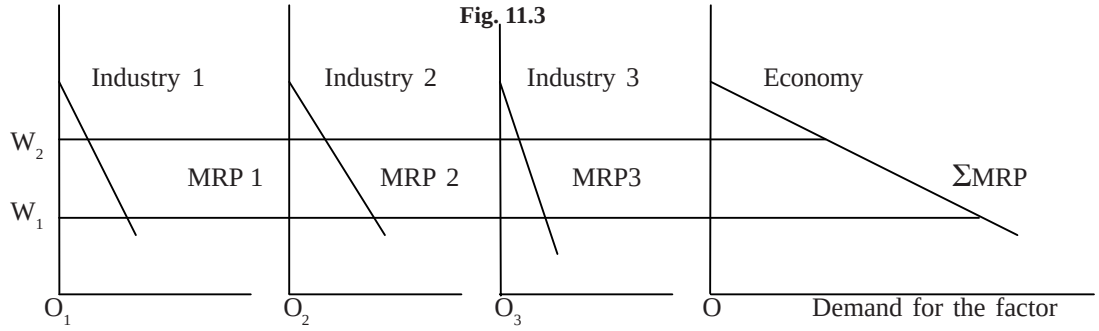


Fig. 11.3: Shows derivation of market demand curve by horizontally summing up demand curves of different industries.

11.5 FACTOR SUPPLY

Prices in factor market are determined by interaction of demand and supply. In Sections 11.1 to 11.4 we have outlined the demand curve for a factor of production. Now we turn to considerations that determine the supply of a factor of production. We can find out the equilibrium rate of remuneration for the factor. Some special features of factor supply curves must be pinpointed right in the beginning. We have divided the factors of production into four categories, viz., labour, land, capital and entrepreneurship. Their intrinsic nature is quite different. We cannot say that supply of labour and supply of land will be akin to, say, supply of sugar and supply wheat. In case of commodities, a buyer takes the quantity with her. But a buyer of land's productive services will have to come to the land and work on it. Similarly, when one buys labour, she pays for the work, be done by the labour. Here, the labourer comes and works at the place decided upon by the employer, completes the work and goes back. So, buying a commodity may imply an absolute transfer of ownership whereas hiring of factor services may not involve any such happening. Land, after raising of a crop by hirer, reverts back to the owner. Similarly, owner of labour himself has to work to make the supply of labour effective. Still, we can define the notion of supply curve of a factor of production.

11.5.1 Remuneration / Reward for a Factor and its Supply

Given the physical occurrence of a factor in the society, its supply will depend on what reward or remuneration can be earned. So higher the reward, higher the supply of a factor. Fig. 11.4 shows the supply curve for a factor of production.

Fig. 11.4

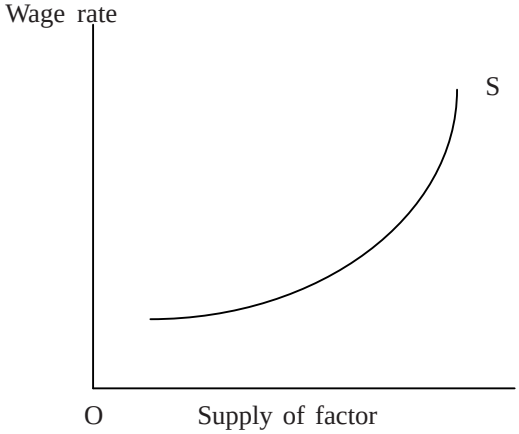


Fig. 11.4: Shows a supply curve for factor of production. Higher remuneration induces the owner of the factor to supply larger quantities, which amounts to letting some one make greater use of that factor's productive powers.

Reward or remuneration offers inducement to the owner of a factor to let some one use the productive power/ capacities under its possession. Higher the inducement, larger the quantity the owner will be willing to offer.

We give different names to rewards for different factors. Remuneration for labour is called wage. The reward for use of land is called rent. Interest is received for letting someone use capital, while entrepreneurs get profits.

11.5.2 Product Exhaustion Theorem

Marginal productivity theory of distribution shows that each unit of a factor of production receives remuneration equal to its marginal product. This prescription implies that each unit receives equal rewards. Reconsider our Fig. 11.1. The market wage rate is Rs. 40. Here, reward is equal to MRP of the last worker employed. We know that value of marginal product of the first worker was Rs. 200. Similarly, 2nd workers' efforts produced output valued as Rs. 160 and so on. Reward equals to marginal revenue product does not mean that each unit is rewarded its own marginal revenue product. It has to accept reward equal to MRP of the last unit of its kind that is employed. We have defined the difference between wage bill and total product as "Producers' Surplus". This amount is not gobbled up by producers. They arrange for compensation of other factors of production out of this surplus.

In the present section, we examine the implication of marginal productivity theory of distribution. One way to assess the implication is through assumption of constant returns to scale. You are familiar with Cobb–Douglas production function. Remember, for example,

$Q = AL^{\alpha}K^{\beta}$, where $\alpha + \beta = 1$ and

L : Labour
K : Capital
A : Technical Constant
Q : Output

Here, let us take marginal product of labour and capital as MP_L and MP_K respectively. We can derive these magnitudes by differentiating the above production function with respect to labour (L) and capital (K). Thus,

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

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

The wage rate (W) is equal to MP_L and the rate of reward for capital (r) is equal to MP_K. Therefore, the total wage bill will be $MP_L \times L$, and total reward for capital will be $MP_K \times K$.

Hence we can say,

share of labour $= L \times \alpha AL^{\alpha-1}K^{\beta} = \alpha AL^{\alpha}K^{\beta}$

share of capital $= K \times \beta AL^{\alpha}K^{\beta-1} = \beta AL^{\alpha}K^{\beta}$

We know that only two factors labour and capital have been used here. Now add up their shares:

Share of L + share of K $= \alpha AL^{\alpha}K^{\beta} + \beta AL^{\alpha}K^{\beta}$

$= AL^{\alpha}K^{\beta}(\alpha+\beta)$

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

(We know that $\alpha+\beta=1$)

We find that share of two factors add upto total output. Hence the name, 'Product Exhaustion Theorem', that is, if each unit of a factor is given rewards equal to its marginal product, the total output is exactly divided between the participating factors of production. Nothing is left over!

11.5.3 Factor Remuneration and Employment under Imperfect Commodity Market

So far we have assumed perfect competition in both factor and commodity markets. Let us relax this assumption in respect of commodity market and examine the consequence. We know that under perfect competition, each producer is a price taker in commodity market. The demand curve for the commodity is horizontal and the price, or average revenue equals marginal revenue (i.e., AR and MR curves overlap). Therefore, MRP and VMP curves coincide as we have shown in Fig. 11.1. However, if commodity market was not perfectly competitive, the demand curve will be sloping downwards to the right and the corresponding marginal revenue curve will be completely under it, half way between AR curve and vertical axis. What happens to marginal revenue productivity now? Engaging one more worker means producing larger output. Increase in output means selling total output (not only additional output) at a smaller price. So the total revenue does increase, but at a smaller rate. Hence MRP, and not VMP, determines where the employment will stop. Fig. 11.5 shows this situation.

Fig. 11.5

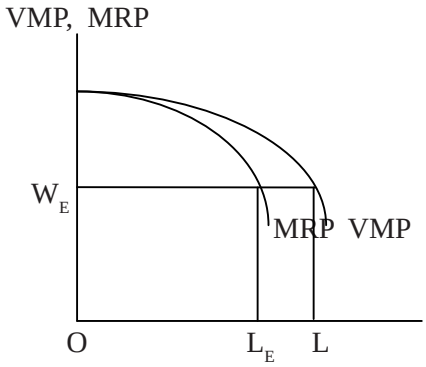


Fig. 11.5: Re-emphasises the fact that under imperfections in product market MRP will no longer coincide with VMP. It shall be wholly inside VMP. Therefore, at any wage rate W_E employment of labour will be OL_E, which is substantially less, then OL, and the employment under perfect competition.

Had perfect competition prevailed, MRP would have coincided with VMP and employment at wage rate OW_E would have been OL_E. But imperfection of the commodity market forces MRP to depart from VMP. Hence the equilibrium level of employment is OL_E, which is substantially lower than the perfect competition employment. Therefore, it can be said that imperfection in commodity market affects the employment level adversely.

We now turn to impact on wage rate in the market with imperfect competition. Had the perfect competition been there, OL_E number of workers would have been entitled to receive a wage rate equal to OW_1 and not OW_E.

So, we can say that as a result of commodity market imperfection, labour has to remain satisfied with a smaller wage rate and contend with lesser employment as

well. This points to exploitation of labour. Such an argument applies equally to remuneration and employment of all other factors.

We shall examine the effect of factor market imperfections when we discuss individual factor price determination. For labour, these aspects are examined in Section 11.10.

11.6 LIMITATIONS OF MARGINAL PRODUCTIVITY THEORY OF DISTRIBUTION

Sometimes, on the strength of product exhaustion theorem, an impression is sought to be created that since marginal productivity theory leaves no 'remainder' after rewarding the factors of production, it must be inherently fair as well. However, fairness of a theory cannot depend upon accounting calculation. The rewards do not go to factors. These are received by the owners.

Marginal productivity theory does not tell us whether the person getting reward on the basis of MRP actually deserved the same or not. For instance, MRP of an acre of land may be Rs. 2000, whosoever controls that acre gets the rent equal to Rs. 2000. But does that person really deserve that amount? The landlord might have acquired control over the land through fraudulent means, or through sheer brute force of hired goons, or may simply have occupied public land in the riverbed.

Again note that MRP does not depend upon how hard, say, different workers work on the land. As MRP = MPP $\times$ MR, it depends on how many workers are employed. More the employment, lower the marginal physical product and lower the MRP. Thus, MRP does not reflect the efforts of workers who are given remuneration on its basis. As we pointed out in Section 11.5, employment and remuneration are also adversely affected by happenings in product market. Suppose a firm grows very large on the basis of hard work put in by its employees. It tends to acquire some 'power' in the market. Its AR and thus MR tend to slope downwards. As a result, MRP deviates from VMP. The firm tends to employ fewer persons at lower wage than it would have under perfect competition. Like any other theory, this one too derives its strength from the assumptions of constant returns to scale and perfect competition, with no externalities. But all the three assumptions are to be suspected in the present day world.

In any case, this theory treats human recipients of remuneration as inputs in the productive system. What happens to, say, aged and infirms? Their marginal revenue productivity drops to zero and hence their earnings too should be nil. Such mechanical approach to the problem of distributing output among the members of society cannot be justified and need not be regarded desirable on social considerations.

So, by way of conclusion, we can say, marginal productivity theory provides us with some insights into working of system and policy matters. Though it suffers form several shortcomings, yet it's a better model to explain the working of economy.

Check Your Progress 2

1) Which concept of marginal productivity is taken into consideration by firm while employing a factor? (Answer in one sentence)

   .....

   .....

2) When MRP > W, expansion of employment leads to increase / decrease in firms point of view.

.....

.....

3) State whether the following statements are true or false:

a) Product exhaustion theorem holds good under increasing returns to scale.

b) Product exhaustion theorem holds well when externalities are present.

.....

.....

4) The firms' VMP can become its (demand /supply) curve for labours and is based on (marginal/average) product of labour and the (cost/price) of firms output. (Retain the correct phrase).

.....

.....

11.7 DETERMINATION OF WAGES

In the present section, we examine the determination of wages rate. Before we touch upon specifics of demand for and supply of labour, let us define some notions and general concepts:

i) Wage is the general name for price paid for services of labour.

ii) Salary: Wages paid to white collar workers are called salary. Generally it is shown on monthly basis. Sometimes, even annual salary may be mentioned.

iii) Cash and Kind: Generally wages are paid in cash. However, under certain situations, some categories of workers may be paid wages in the form of, say, food or cloth. For example, draught relief workers receive food grain for the work done by them. Persons employed for the household work, often get food, shelter and clothes as part of payment for their services.

iv) Piece wages are payment according to work done by a worker or a group of workers.

v) Time rate is paid where quality matters but generally it is not possible to measure the work done by each worker. Supervision is a must in this kind of payment considerations.

vi) Nominal wages are monetary wages received by the workers.

vii) Real wages refer to purchasing power of the wages – after all their real worth is in terms of goods and commodities, which can be acquired with the money received.

11.7.1 Labour and Supply of Labour

The term labour denotes human endeavour to transform raw inputs into commodities. The transformation may be physical, say, making cloth out of cotton; spatial, say, making apples grown in Kashmir available to consumers in Kerala; or in time, say, making available in winter the produces of summer harvests. All the three

transformations add utility to things and are known as production (includes making/ manufacturing, storing and transporting). Human efforts involved in these activities are called labour. Some of the labour may be paid for while others may not be paid at all. For instance, if members of family help in family business, wages are not even calculated.

Overall supply of labour in society depends upon:

i) Size and age composition of population,

ii) Norms of and participation in work force.

This way, supply of labour undetermined by a set of complex economic-social-biological and even legal forces.

11.7.2 Supply of Labour by an Individual

Generally, one would expect that an offer of higher wage rate should induce a worker to put in longer hours of work. However, such a view ignores the fact that a worker has to remain physically present and active while supplying her labour. She needs relaxation and time to regenerate her ability to work. Therefore, it is unrealistic to expect a person to keep working all the twenty four hour a day continuously. Thus, while a person gets very low hourly wages, she tends to work longer at a higher wage rate. However, once higher wage rate is fixed, she may prefer to have longer hours of leisure so that she can enjoy fruits of her labour. So her supply curve tends to have a backward bending at a some sufficiently high wage rate. The wage rate at which such a feature is encountered will differ from worker to worker, depending upon her won circumstances and needs.

Fig. 11.6

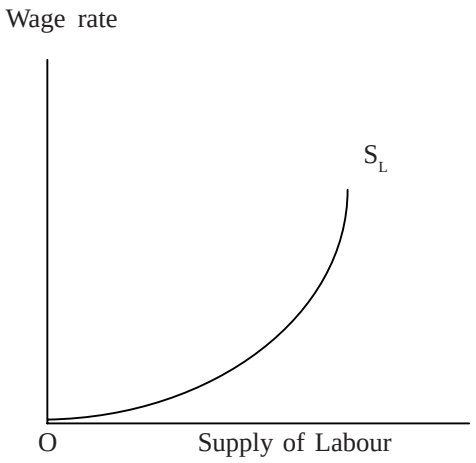


Fig. 11.6: Shows backward bending supply curve of lobour. The ben occurs at a sufficiently high level of hourly wage rates, as the worker want time to enjoy the fruits of their labour. In other words, as they are able to meet certain requirements they tend to place a higher vluation of the leisure.

11.7.3 Market – Supply of Labour

We can derive the market supply curve of labour as we do with supply curve of any commodity. This is nothing but horizontal summation of individuals supply curves. Fig. 11.7 shows supply curves of 3 workers, marked S_1, S_2, and S_3 respectively. These are added up to obtain the market supply of labour.

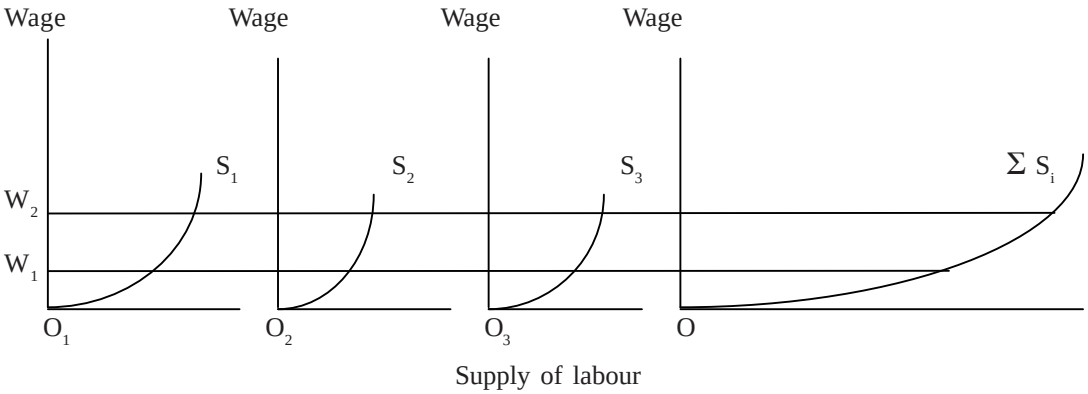


Fig. 11.7: We have shown three individual workers' supply curves of labour in the first panel. The market supply curve shown in the second panel is horizontal summation of this individual supply curve.

11.8 SHORT RUN DEMAND FOR LABOUR

Producer of various goods and services need labour to produce these for consumers. If consumers demand larger quantities of commodity X, manufacturers of such a commodity sense the opportunity to earn more profit through expansion of output. Therefore, they demand larger quantities of labour. In short run, they may engage the existing workers to work for extra hours over and above the normal schedule so that production can help them meet additional market demand. They may set up new factories, expand capacities in their existing ones only where rise in market demand becomes permanent. If, on the other hand, in industry Y the demand for commodity declines, some temporary or casual workers may be discharged. Alternatively, overtime production may stop. In both the cases, demand for labour undergoes a change in response to changes in demand for commodities, which are manufactured by labour. For this reason that we say demand for labour is derived demand.

An employer will keep employing more workers as long as marginal revenue exceeds marginal costs. We know that addition to revenue that results from employment of one more worker is called marginal revenue product (MRP). Addition to cost, in similar manner, is called marginal wage cost (MWC). Marginal productivity theory says that a producer goes on employing more and more workers so long as MRP exceeds MWC. When MRP is just equal to MWC, i.e.,

MRP = MWC, equilibrium is attained.

We know that MRP depends on marginal physical product of labour and marginal revenue in the product market, which have been discussed in Section 11.4. When labour market is in perfect competition, the producer in question can hire as many workers as she needs at the going market wage rate. Thus, MWC will be equal to wage (W). Hence, we can say that producer's equilibrium occurs when

MRP = W.

If MRP>W, the producer can increase her profit by expanding employment. If MRP< W she can contain her losses by reducing employment. When MRP= W, it is not in her interest to move away from that level of employment. Thus, downwards-sloping portion of MRP becomes demand curve for labour. It shows that higher the wage rate, lower will be the employment of labour. Conversely, if the wage rate declines, the producer will hire larger number of workers.

MRP tends to slope downwards as MPP declines due to employment of larger number of workers. This happens because of law of diminishing returns to the variable factor. Hence,

as MRP = MPP × MR, it tends to decline as the MPP declines.

We also know that under perfect competition in product market MR = AR

Therefore,

MRP = MPP × MR = MPP × P = VMP. Though MR = P is constant for the firm,

MRP = VMP still tends to fall as a result of fall in MPP.

If, however, there exists some imperfection in the product market, MR tends to lie under AR. Hence, MRP = MPP × MR tends to fall rather sharply.

11.8.1 Labour Market: From Individual Firm to Market Demand for Labour

We can obtain demand curve for labour, as discussed in Section 11.4.3 above, by adding together individual firms' demand curves. However, we have to make allowance for one fact, i.e., as the wage rate falls, all the firms employ more workers. This increased employment of labour and leads to production of much larger output. Though, under perfect competition, a single producer could have sold any amount of the product at its going market price, yet, when all of them want to sell larger quantities, market supply of the commodity shifts to the right. This lowers the market price. Hence, the increased output forces producers to sell their products at a lower price. As a result MRP = MPP × MR tends to fall sharply. So the industry demand for labour curve will appear to be far steeper then what would simple horizontal summation of individual firms' demand curves suggest.

11.8.2 Elasticity of Demand for Labour

Responsiveness of demand for labour to changes in wage is called elasticity of demand for labour. Thus,

E = (Per cent change in labour demanded) ÷ (per cent change in wage rate)

Just like consumers' demand for a commodity, we say demand for labour is elastic if E >1. It will be called inelastic if E <1.

When elasticity of demand for labour is unity, the total wage bill remains unchanged. When E >1, a fall in W will lead to increase in wage bill. However, when E <1, a fall in wage rate does lead to a fall in wage bill as well.

11.9 SUPPLY AND DEMAND BROUGHT TOGETHER: THE MARKET EQUILIBRIUM

11.9.1 Perfect Competion in both Product and Labour Market

Buyers and sellers of labour constitute the labour market and intersection of demand for and supply of labour gives the point of equilibrium. We have already derived the supply curve of labour in Section 11.7.3 above. The considerations outlined above in Section 11.8.1 give us a downward sloping market demand curve for labour. We

can put them together in Fig. 11.8. The two curves intersect at point E. The market wage rate is OW and OL quantity of labour is bought and sold.

Fig. 11.8

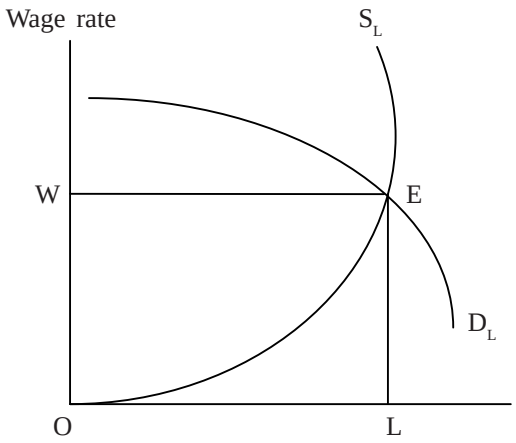


Fig. 11.8: Shows equilibrium of labour market with intersection of demand and supply curves for labour. The wage rate thus determined is OW and the level of employment is OL.

Individual producers of different commodities can make adjustments in their production plans, given the market wage rate, using the principle of equating their MRP to the MWC as discussed in Section 11.8 above.

Please note that so far we have not deviated from our standard assumption of perfect competition in labour market. What happens when labour market fails to meet the tests of perfect competition? What causes the labour market to move away from perfect competition? We discuss these and related quarries in Section 11.10 below.

Check Your Progress 3

1) Individual supply curve may be backward bending. Why? Explain in about 50 words.

   .....

   .....

2) Explain Marginal Wage Cost in one sentence.

   .....

   .....

3) What will happen to wage bill when wage rises and

   i) demand for labour is inelastic?

   ii) demand for labour is elastic?

   iii) demand for labour is unitary elastic?

   .....

   .....

   .....

11.10 IMPERFECTIONS IN MARKET AND THE WAGE RATE

We can consider several types of imperfections in the market.

1) The product market may be imperfect, that is, the producer may not be just one of the numerous perfectly competitive players.

2) There may be imperfection in the labour market. This can result from several sources: we know that today, an industrialist is not just like any small employer of the nineteenth century. Control over massive amounts of financial resources and a very large scale of operation imply that in any area that employer may be the only employer, or at the best one of the few employers. Thus, her market power will be disproportionately large compared to any single job seeker.

Attempts to counter the 'market power' of employers have led to organisation of trade unions. These organisations of workers create their own imperfections in labour market.

In the present section, we discuss all these possibilities and their implications for employment and wage rate.

11.10.1 Product Market Imperfections

When the producer has some market power in the product market, she cannot sell any quantity that she desires to produce at going market price. She can sell larger quantities at smaller prices only. Thus, demand curve for her output slopes downwards. She gets lower prices if she produces more. In the process, VMP is no longer the important decision factor. She is more affected by MRP. Hence, her demand for labour curve is not VMP but MRP curve. Therefore, at the given market wage rate, she will employ fewer workers as shown in Fig. 11.5 above. Recall that, at WE , the wage rate under perfect competition, firms demand for labour would have been OL. But as MRP tends to drift down VMP, firm's demand at this wage rate is down to OLE.

Thus, the monopolistic employer in product market will employ fewer workers compared to a perfectly competitive market.

11.10.2 Labour Market Imperfections

a) **Trade Union**: The organisation of workers tries to better the working condition of workers. It may try to extract higher nominal wages and other benefits from an employer. In fact, trade unions attempt to raise earnings would be more successful when demand curve cuts supply of labour curve near the backward bending portion and the leaders would bargain accordingly to reap the maximum benefit. Further, the union may try to restrict supply of labour to employers. Sometimes they are able enforce 'closed shop' recruitment policy, that is, vacancies in the establishment will be filled up from amongst the members of the trade union, or, their families only.

b) **Monopolistic Employer**: When only one firm is there in the market, which can employ workers, it tends to have tremendous clout. The poor workers have no option but to accept what the firm offers. In such situation wage rate and employment both suffer.

c) **Bilateral Monopoly**: In this situation a monopolist employer has to negotiate with a powerful trade union. The employer tries to keep wage rate and employment as low as possible. But the union tries to attain the highest possible

level of wages. The ultimate outcome will depend upon relative bargaining powers and skills of the firm and the union. In such situation, we say, wage rate is indeterminate. It can be anywhere between the lowest, the minimum which workers may accept, and the highest, beyond which employer will never pay.

11.10.3 State Intervention

Sometimes, the government intervenes in the labour market. It may fix a certain minimum wage level and ultimately employers are forced to pay at least that much to labour. This minimum wage can benefit workers only if it is higher than the equilibrium wage rate and if large-scale displacement of labour does not follow such wage level fixation.

The state may alternatively act as a model employer. The wages and other facilities offered by public sector tend to become benchmark levels all across the economy.

11.11 WHY DO SO MAY KINDS OF WAGE RATES PREVAIL IN THE MARKET?

We have been trying to determine the wage rate in our analysis so far. But any observer of the market will immediately point it out that there are any number of wage rates, which exist in the market. Such wage rates differ not by insignificant amounts. How do we explain these differentials? We can say that wage rates differ because:

1) jobs are heterogeneous

2) workers are heterogonous

3) information is imperfect and costly

4) labour mobility is restricted and involves huge establishment costs.

Sometimes we try to explain away the wage differences in terms of differences in quality of human capital. For example,

i) abilities of workers differ. Therefore, they receive different levels of earning;

ii) family background ensures that some people are able to acquire abilities to earn more. For example, a rich person's sole daughter can afford to pay for higher/ technical education, while a poor, yet brighter, child may have to help increase family's meagre income and in the process discontinues studies;

iii) some people take chances or risks and are in some of the cases, duly rewarded for that.

We may emphasis again that to some extent, trade unions are able to raise the wage levels. Yet, those increases may remain confined to their members and organised sector of the economy. The unorganised sector, which is much larger than the organised portion and where unionisation of labour may not be feasible, remains outside the trade union negotiated wage increases.

Check Your Progress 4

1) When a producer has monopolistic power in product market, she employs fewer worker. Why? (Explain in about 50 words.)

.....

.....

.....

2) When a monopolistic producer bargains with monopolistic trade union, wage rate becomes indeterminate. Why? (Explain in about 50 words.)

.....

.....

.....

3) What are the factors that explain differentials in earnings of labour? (Write in about 100 words.)

.....

.....

.....

.....

.....

11.12 LET US SUM UP

The present unit has introduced you to factor pricing. It deals with marginal productivity theory, which has been one of the most dominant theories in economics. It asserts that if competitive condition prevails in both product and factor markets, then rewards earned by different factors will be equal to their respective marginal products. The theory is further strengthened by product exhaustion theorem which asserts that if all the factors are rewarded according to their marginal products, the total product will be so divided among the various participating factors of production that there remains no left over.

However, marginal productivity theory is not above criticisms. It tends to concentrate only on the demand side of factor market ignoring supply side influences, where the determination of market price must take into consideration both demand and supply. Further, excessive dependence on perfect competition also leads to inherent shortcomings in the theory, as this type of market conditions may fail to exist in real life.

Even the equilibrium wage rate determined through demand and supply interaction is subject to modification by institutional factors like state intervention and actions of trade unions.

11.13 KEY WORDS

Backward Bending Supply Curve of Labour : Relationship between supply of labour and wage rate may reverse after a certain wage level. It may happen due to domination of income effect over substitution effect.

Competitive Market : A market with large number of buyers and sellers operating freely and no one is able to control the market price.

Income Effect : A wage rise increases income of workers and they tend to consume more of every good including labour (in the form of leisure). Thus, labour supply in terms of hours of work tends to decline.

Institutional Factors : The social, political and organisational factors, which affect economic decision-making.

Marginal Physical Product : Change in quantity produced as one additional unit of the variable factor is engaged, all other factors remaining constant.

Marginal Revenue Product : Marginal physical product multiplied by marginal revenue.

Marginal Value Product : Marginal physical product multiplied by price of the commodity.

Minimum Wage Act : Government passing a law fixing minimum level of wages payable.

Mobility of labour : Willingness of workers to take up employment at places far off from their normal residence.

Non-Competitive Markets : A market where some condition or other of perfect competition is not satisfied.

Nominal Wage Rate : Wage rate in term of current prices.

Product Exhaustion Theorem : If each unit of a factor is given a reward equal to marginal product of that factor, total output will be completely divided among the factor of production.

Substitution Effect : An increase in wage rate makes leisure costlier in terms of income foregone. This induces worker to work more.

Trade Union : A recognised organisation of workers that seeks to protect their rights.

Wage Differentiations : Differences in average earnings of different groups of workers.

11.14 SOME USEFUL BOOKS

Baumol, W.I. and Blinder, A.S., 1988, *Economics: Principles and Policy*, Harcourt Brace Jovanovich, Chicago

Stonier, A.W. and Hague, D.C., A *Text Book of Economics*, Macmillan and ELBS, London

11.15 ANSWERS/HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) Read Section 11.3 and answer

2) (a) 2, (b) Rs. 90 (c) Rs. 40

Check Your Progress 2

1) MRP

2) increase, decrease

3) (a) False, (b) False

4) demand, marginal, price

Check Your Progress 3

1) Read Sub-section 11.7.1 and answer

2) Read Section 11.8 and answer

3) i) increases
   ii) decreases
   iii) remain the same

Check Your Progress 4

1) Read Sub-section 11.10.1 and answer

2) Read Sub-section 11.10.2 and answer

3) Read Sub-section 11.11 and answer

UNIT 12 RENT, INTEREST AND PROFIT

Structure

12.1 OBJECTIVES

This unit deals with determination of remuneration of non-labour factors of production. On going through this unit, you will be able to:

- understand and explain the origin of concept of rent;
- extend the notion of rent to other factors of production;
- understand how the rates of interest are determined and why do they vary; and
- explain how the profits arise in the economy.

12.1 INTRODUCTION

In this unit, we focus our attention on determination of rates of remuneration/ rewards for non-labour factors of production. The treatment of various ideas and theories will highlight several limitations of the marginal productivity theory of distribution, which is supposed to be applicable to all factors of production. You will note, how the determination of rates of remuneration for land, capital and entrepreneurship tends to deviate from marginal productivity based approach used for labour. In fact, virtually each of these – three factors has some special features, which necessitate a special and different treatment for each one of them.

12.2 REMUNERATION FOR SERVICES OF LAND: RENT

Although today we generally use the term 'rent' for any contractual payment for use of an asset, say, rent for a shop, a house, a piece of land or a machine etc., economists have traditionally associated rent with the use of services of land only. Not only this, the notion of rent has its origin in feudal societies, where most of land was owned by big landlords or noblemen. They used to charge some payment from the farmers who cultivated those tracts of land. That portion of payment for use of productive power of the soil was called rent.

Theory of rent has two variations: (i) the classical theory or Ricardian theory and (ii) the modern theory.

12.2.1 Ricardian Theory of Rent

David Ricardo propounded original theory of rent. His idea was that 'nature has been very generous to human beings and has endowed the land with some 'original and indestructible powers' for their use. Due to those powers, output from the land would exceed the total of payment to all kinds of inputs. Thus, after compensating all other factors of production involved in agriculture, there would remain a surplus. The owner of the land can rightfully claim that 'surplus'.

Thus we can show this kind of rent as surplus arising from agricultural activities. We can show it in two ways:

i) Intensive Cultivation and

ii) Extensive Cultivation

i) **Intensive Cultivation:** A farmer with a given piece of land keeps on engaging more and more workers as long as the marginal revenue product (MRPL) exceeds market wage rate.

Fig.12.1

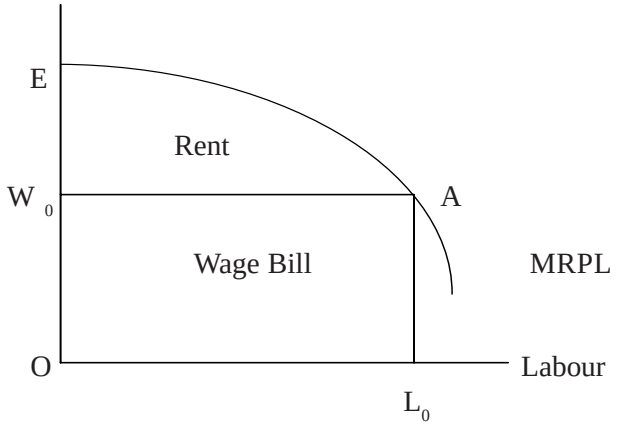


Fig.12.1: Recall from Unit-6 that if we increase the number of workers engaged on a fixed piece of land, extra output generated by each additional worker keeps on falling. Also, recall that in Unit-11, we had shown that under the given market conditions, the marginal revenue product curve of the variable factor becomes its demand curve. This figure uses that demand curve for variable factor labour to determine its share in the total product, that is, the wage bill. The rest is surplus, called rent.

In Fig.12.1, OL_0 is the number of workers employed at market wage rate of OW_0. Therefore, total product of labour equals $OEAL_0$. But, the wage will be equal to the rectangle OW_0AL_0 only. Thus, when worker are paid at the going market rate, the landlord is still left with a surplus equal to W_0 EA. This will be regarded as share of land – the rent.

ii) **Extensive cultivation:** This approach to the determination of rent is significantly different from that of intensive cultivation example seen above. Suppose we have land of different grades, say, A, B, C, D and E. Initially, the farmer will plough the best grade land. As the population rises, demand for the farm produce grows. Therefore, inferior grade land is also brought under the plough, gradually. Suppose, value of the produce from the E grade is barely equal to input costs. Then we can say that plot E generates no surplus. Hence, it is called marginal piece of land. Rest of the pieces marked A, B, C and D generate output greater then the costs – which are same for all the pieces. This excess of value of output over and above the costs is rent.

Fig.12.2

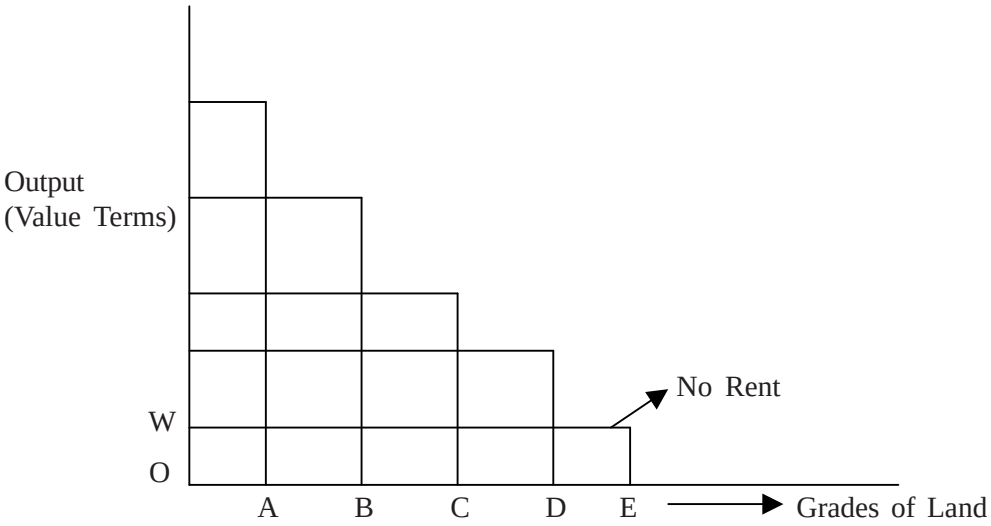


Fig.12.2: We have 5 pieces of land, A, B, C, D and E. The piece A is of the best grade, B is inferior to A, C is inferior to B and so on. Thus, the piece E is of the worst grade. Heights of bars above A, B, C, D and E indicate the value of output produced on the respective pieces. We find that value of output on piece E is just enough to cover the cost of cultivation. Also note that this cost = OW is same for each of the pieces. So the Grade E land generates no surplus, while grade A generates the maximum. Such surplus is appropriated as rent.

Hence, it can be said that superior grade land generates a surplus. That surplus is called a rent and is claimed by owners of land as differential surplus. If market value of produce on the marginal land E increases above the cost of cultivation, the farmer will have an incentive to move on to the so far sub-marginal land F as well – thus extending the cultivation. If, on the other hand, market value falls below cost of cultivation, the farmer may be forced to stop using the marginal land. Therefore, we can say that only the lands superior to the marginal land can generate surplus or rent. Ricardo asserted that rent of land tended to depend upon the price of the produce of land. Rent was not regarded as a cost of production for the farm output, as wage bill represented the labour cost.

Criticism: Ricardo's theory has been criticised on several grounds:

a) Historically, best piece of land may not have been cultivated first. Yet, it is true that order of cultivation does not affect the nature of rent as a surplus.

b) Ricardo attributed the surplus to 'original and indestructible' powers of land. We know that fertility of soil is neither original nor indestructible. It can be changed with the help of technology. But, to be fair to Ricardo, technical developments were not known in his time. In any case, this assumption can be reinterpreted to suggest inelasticity of supply of land with respect to changes in its prices.

c) Ricardo assumed perfect competition in the farm sector. However, this assumption is most unrealistic at best of times in agriculture. The ownership of land is concentrated in a fewer hands, which does not promote competition – it rather restricts it.

d) Rent may not be an unearned surplus after all! The superior land does not come into being on its own. The superiority and higher productivity could be a result of cleaning of stones and boulders, shrubs and roots and arrangements for proper drainage and irrigation Therefore, the land supply may not be given. It may have a cost function .

e) Ricardo had considered only one use for land, namely, agriculture. Alternative uses of land will mean a certain minimum payment, for, the same will be needed to retain it in the existing use. This, however, may not remain limited to only land. It can be extended to any factor of production.

Despite these criticisms, Ricardian description of rent as unearned differential surplus that rises along with economic progress has had tremendous influence on policy makers around the world. Land reforms like Zamindari abolition have their roots in an urge to eliminate unearned surplus.

12.2.2 Marshall's Theory

Alfred Marshall tried to extend the Ricardian theory of rent to other factors. He thought that if the supply of any factor was limited and could not be raised through efforts of human beings, then income derived by this factor would be regarded as rent. Such short period rent was given the name quasi-rent by Marshall. Ricardo's 'rent' was a long-term phenomenon while Marshall's quasi-rent was confined to essentially a short period of time. We know that land supply may remain fixed over very long periods of time. But supply of other factors may change in the long run.

12.2.3 Modern Theory of Rent

Modern economic analysis develops Marshall's ideas still further. Here, rent is sought to be determined like any other factor price through interaction of demand and supply curves. Even when all the land units are equally fertile, rent will still arise on account of its scarcity.

Fig.12.3

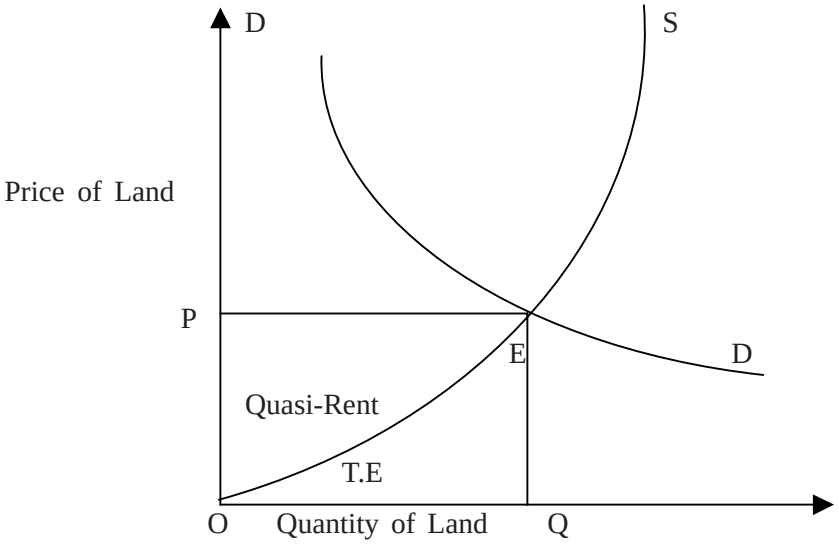


Fig.12.3: Is simple demand and supply intersection diagram. Supply curve shows the asking price of each factor unit, that is, the minimum price for which a given number of factor units will be willing to work. Similarly, demand curve shows the maximum price the buyer is willing to pay for each given amount of the factor. Intersection of the two curves determines usual equilibrium. The area under supply curve - OEQ is called 'transfer earnings'– the minimum payment needed to retain the given factor units in the present employment. The area OPE represents the 'quasi-rent' - rent type payment over and above the transfer earnings which arise because of relative scarcity of the factor, or, inelasticity of factor supply.

Fig. 12.3 shows such a situation. We have usual demand and supply curve for land (or, for that matter, any factor of production). Point E is the point of intersection of the demand and supply curves. Those who want to hire land, will be willing to

take OQ amount at OP price (rent). The owners are also ready to supply OQ units of land at this price. Hence the equilibrium. But notice one point. The supply curve also shows the prices at which different units of the factor are willing to work. This is their asking price to retain them in the present employment or use. This can be called transfer earnings – shown by area marked T.E in Fig 12.3. These factor units, of course, receive rent equal to OP. The excess payment received by these units is a rent – like payment, or, the quasi-rent that arises because of inelasticity of the supply.

We can illustrate this aspect of inelasticity of factor supply curve with help of Fig. 12.4. Here demand curve DD is shown to intersect with four supply curves S_1, S_2, S_3 and S_4.

The curve S_1 is perfectly inelastic. Therefore, entire payment to the factor is rent. This is also called 'pure economic rent'.

Fig. 12.4

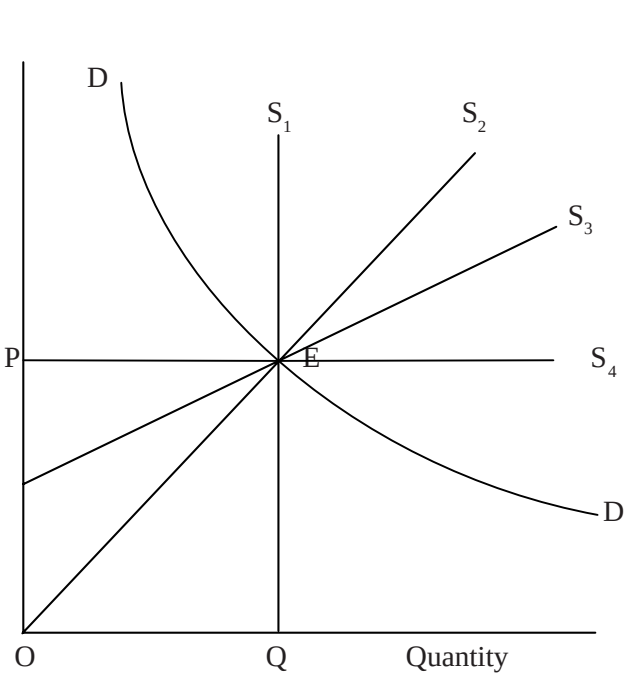


Fig. 12.4: Shows four supply curves, S_1, S_2, S_3 and S_4. They intersect with the demand curve at a single point E. Thus at each of the supply curve OP is the factor price and OQ is the equilibrium quantity. Area under each of the supply curve upto point Q is called transfer earning (TE). The total factor payment is OPEQ in all the four cases. Note that with supply curve S_1, TE is zero, while with S_4 entire factor payment becomes transfer earning.

Note that S_2 is somewhat elastic. Immediately, we find that a significant part of the total factor earnings became transfer earnings (denoted by the area under the curve); yet a substantial part is above the supply curve and it is called quasi-rent. As supply conditions improve further and elasticity of supply tends to increase, share of TE tends to rise and that of QR tends to decline. When supply is perfectly elastic, as shown by S_4, the entire payment is in the nature of transfer earning and quasi-rent simply vanishes.

It may be pointed out here, that, when Marshall described rent as a surplus, he had this excess over the transfer earnings in mind. Thus, his concept of rent was substantially different from that of Ricardo.

Check Your Progress 1

1) Explain the notion of extensive Margin.

.....

.....

.....

.....

2) What is Intensive Margin?

.....

.....

.....

.....

3) Explain Rent as a Surplus.

.....

.....

.....

.....

4) Explain the notion of Quasi Rent.

.....

.....

.....

.....

12.3 REMUNERATION FOR THE SERVICES OF CAPITAL: INTEREST

Capital is a factor of production made by human beings. They transform some free gifts of nature into such tools and implements, which help them in producing things that satisfy some of their wants. Had they not diverted some time and worked to perfection of those implements, they could have produced more goods. Thus, if some are making implements, they is forgoing some opportunity to produce some consumption goods. However, at a later stage, when human beings used implements, they were able to produce a larger quantity of consumers' goods quickly. Capital enhances productivity of labour and this is helpful in the process of production. Therefore, we regard it as a factor of production. We have further enlarged the notion of capital to include all machinery, buildings, factories etc., which help in the process of production. Even the money needed to continue operations of an enterprise smoothly, called working capital, may include some stock of raw materials and some inventory of finished goods as well. Raw materials ensure smooth functioning of manufacturing while some inventories are needed to ensure regular supply to the customers. We can say that initially capital was defined with respect to tools and implements, but now it encompasses buildings housing those implements/machinery as well as whatever stock etc., may be regarded as necessary for smooth functioning of the enterprise. The reward for the services of capital is called interest.

12.3.1 Various Notions of Interest

You must have noted in the introductory note above that capital comes into existence as someone abstains from consumption or saves. He may use his savings himself, or, allow someone else to use them. He expects the user to reward him for this facility. This reward will be interest which may be in kind as well. For example, if a farmer borrows some grain from a neighbour, she returns a larger quantity when she harvests her crops.

In other segments of the economy, some people who need more money than their current income, borrow from those who do not want to spend their entire income. The borrower uses such funds for a given period of time and returns the original amount, called principal, along with some additional sum called interest. The 'price' for borrowings is generally designated as a percentage and chargeable per unit of time (usually one year). For instance, suppose the rate of interest is 10 per cent per annum. Then the meaning is that if one hundred rupees are borrowed for one year, the borrower has to return Rs.100 + Rs.10 (10 per cent of Rs. 100) at the end of said one year. Therefore, interest works out to Rs.10. There are a number of notions of the rate of interest. Some of the prominent types are discussed below:

Simple and Compound Interest

When someone borrows P units of money for 'n' units of time at 'r' rate of interest, then the amount of interest will depend on 'n', 'r' as well as P. You may have done calculations of interest in your school level arithmetic as follows:

Interest = Principal $\times$ n $\times$ (r / 100). Thus, if Rs.1000 were borrowed at the rate of 10 % for one year, interest will be 1000 x 10/100 = Rs.100. If the loan was for 10 years, interest would be (1000 x 10/100) 10 = Rs. 1000

This type of interest is called **simple interest**.

Sometimes, the amount of interest that becomes due at the end of one accounting period is added to the initial principal. For the second accounting year, a new principal, which is old principal plus interest for the first accounting period, is used for calculation of interest. In each subsequent period, interest due over the previous period is added to or compounded with the principal. Thus, earlier interest also earns interest. Such rate of interest is called compound rate of interest. Final amount A at the end of n years can be calculated as:

$A = P \times (1+ r)$ n, where r is per cent rate of interest and n is number of years.

It may be noted that it is not uncommon to find compound rate of interest in actual life. Consider ordinary savings account with any bank in India. Banks offer 4% rates of interest per annum. But every 6 months, interest is calculated and credited to the account. Thus, unless withdrawn, during second half of the year, interest will be calculated on that previous amount of interest as well. So, we can say that savings accounts earn 4% compound rate of interest, compounding being done half yearly.

Interest and Money

In general, people associate interest with borrowing and lending of money. However, interest exists under barter system also. Consider a farmer borrowing some grains from a neighbour. She returns somewhat larger quantity later. Ordinarily we will regard that extra quantity of grain as interest. One can argue that the monetary value of borrowed quantity should be evaluated at harvest time when prices are low. If one compares values of two quantities at respective prices, then the notion

of interest gets somewhat blurred. However, in an economy where money is yet to be introduced, evaluation at different prices cannot take place as there would be no monetary prices in existence. Even in money using economies, farmers borrow grain at time of sowing and return a larger quantity at harvest time. The difference between quantities returned and borrowed can be called grain rate of interest.

An economist named Wicksell suggested the notion of natural rate of interest, which was very close to the grain rate of interest. He says, " In an economy where prices are not changing, a forester can cut down trees and lend the sales proceeds at some monetary rate of interest. Alternatively, the trees are allowed to grow for one more year, then harvested and sold. During the year, trees will get thicker (may be taller too), and the quantity of wood will be larger. This increase in yield in constant prices can be called natural rate of interest. If the money rate available is just as much as the likely increase in yield from the tree than Wicksell would say that rate of interest in society was equal to the natural rate of interest.

Hire-Purchase or Credit Transactions

Sometimes we do not make full payment at the time of buying some appliances. We pay over several months in instalments. In such a system, payment is much larger than that of the original price tag. The difference is quite akin to interest. Such a transaction is equivalent to borrowing money from the seller to pay her for the goods.

These days, dealers are arranging finance from some other companies or banks for the purchase of consumer durables. The customer buys, say, a car from one company, which receives payment from some bank. Thereafter, the customer continues to pay instalments to the bank over agreed period.

Discount

It is one practice under which the borrower is made of pay 'agreed' rate of interest in advance. For example, if agreed rate is 10% and someone borrows Rs.100, the lender has given her only Rs.90 and the borrower has to return Rs.100 at the end of one year. This kind of practice raises the effective rate of interest. Our borrower gets only Rs.90 and has to pay Rs.10 on it as interest. Thus, the real rate paid by him is not 10% but 11%.

Flat Rate of Interest

If one looks at advertisements of the finance companies these days, an interesting variation of interest rate comes to light. These companies usually advertise a very low rate to entice the customers. This is what they call the 'flat' rate of interest. The procedure is as follows:

Suppose someone borrows at a flat rate of 10% and agrees to return the amount – principal plus interest in 10 'equated' or equal annual instalments. The financier calculates interest for the entire principal for ten years, adds it to the principal and works out ten equal instalments. The catch is that with every instalment paid, a part of principal is also returned. So, the borrower continues to pay interest even on that part of the loan, which she has repaid. This way, the effective rate of interest may turn out to be at least twice as high as advertised.

Element of Risk and Uncertainty

In usual sales and purchase, the seller gets the full consideration or value of the goods when the transaction is effected. But, in lending of funds, the borrower gives a promise to pay back the money. Here, there are definite risks of two types: one is the income risk – the borrower may not be able to pay interest. Another risk

could be capital risk – even the principal lent out may not be returned. Such risks are related to amount and period of loan as well as trustworthiness of the borrower. Sometimes, the purpose for which money is borrowed also adds to uncertainty or risk of default. That is the reason for lenders to demand some assets (like gold, buildings, land etc.,) by way of 'security' for the loan. In case of default, the lender has a right to dispose off the proffered securities to recover his dues.

12.3.2 Capital, Investment and Interest

We say that interest is a reward for the services of capital. How does one change the supply of capital in response in rate of interest? This question opens up an interesting set of possibilities. We have to make a distinction between the stock of capital at a given moment of time and an attempt to supply more capital. Such attempt to change availability of capital is called 'investment' which is a flow concept. One can say that if more investment is taking place in an economy, it is building up more capital (stock). How does rate of interest affect this process of building more capital or augmenting supply of capital through additional investment? We can illustrate it with the help of an example :

Example : Relation between Investment and Rate of Interest

Consider a business proposal: You invest Rs.100,000 in some machine. The machine generates, say, an income stream of Rs. 10,000 per annum and has to be scrapped at the end of the period of ten years. The scrap may also be sold at Rs.10, 000.

So, we can say that present cost of machine = Rs.100, 000

Income stream 10 years = Rs.10,000 + Rs.10,000 + … = Rs.100,000

Scrap value at the end of tenth year = Rs.10,000.

Thus, total receipts = Rs. 110,000

The question is, should this investment be made, or, should this machine be installed?

Can we really add up income to be generated in future time periods straight away? Can we say that Rs.10,000 to be available at the end of say next year are as good as the same amount today? Our response will not be yes. Why? If I lend out say Rs. 10,000 today, it shall become some thing more than that at the end of 1st year. At end of second year the amount will be much more and so on. We are all familiar with compound interest formula. We can use it to find out present value of a future income.

Let us use the information given in the example above to work out the changes referred to in the previous paragraph. Suppose that an investment is made in the year 2000. So the first income is generated only in 2001. What will be the present value (in the year 2000) of the sum of Rs.10, 000 that are expected to be received during 2001? Using compound interest formula we can say:

Rs. 10000 = (1+ r) P where P is that sum which can become equal to Rs.10, 000 at the end of year if it earns an interest at the rate r per annum.

So the present worth of Rs.10, 000 of 2001 would have to be = 10000 / (1 + r) in 2000. The income generated in 2002 has to be discounted like this twice, that of 2003 thrice and so on. These discounted values of future incomes can be added up to get an overall present value of an insurance. Thus:

$$PV = \frac{10000}{(1+r)} + \frac{10000}{(1+r)^2} + \frac{10000}{(1+r)^3} + \ldots + \frac{10000}{(1+r)^{10}} + \frac{10000}{(1+r)^{10}}$$

Notice that two terms in the end are alike. This is because we have assumed that at the end of 10 years, the machine can be sold as scarp for Rs.10000. So, in tenth year, we get two amounts of Rs.10, 000 each, one as regular income and another as sales proceeds of the scrap.

We have to compare the present value with the present cost (Rs.100, 000) and decide whether the project is worth being taken up. Note that if we raise 'r' in the above expression, each of the terms becomes smaller, and therefore, present value becomes smaller. On the other hand, a fall in 'r' will make the present value larger. So given the present cost, the decision to make investment shall depend upon 'r', the rate of interest or discount. If, say, initially, PV = PC, then, we can infer that at going rate of interest, the project is just viable. However, a fall in 'r' will make it more attractive while a rise in 'r' will render it unprofitable'. Therefore, we can say that investment depends upon 'r' - the rate of interest.

A person takes up an investment in the expectation of some return from it. However, those returns may change overtime as more and more investment is made. To appreciate this assertion, consider the situation as given in Fig. 12.5.

Fig.12.5

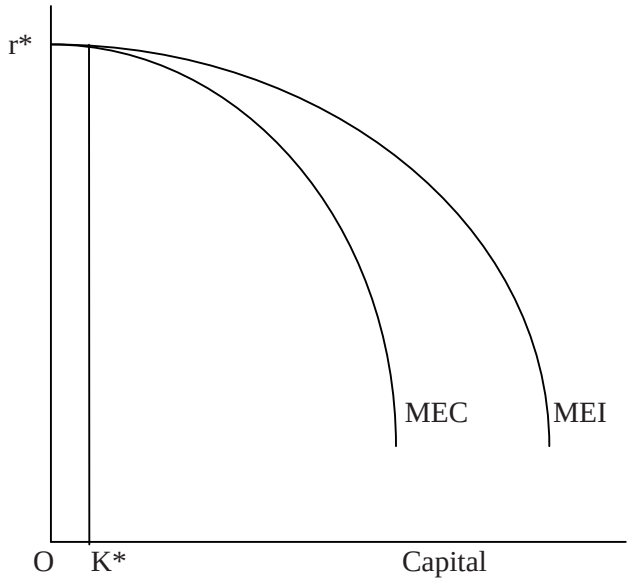


Fig.12.5: Initial stock of capital is K* and it earns a rate of return = r*. This rate attracts investors. More investment raises the capital stock beyond K*. This leads to a fall in the marginal efficiency, MEC, (marginal product of capital), given the other factors of production. Hence, the MEC tends to decline. But this decline can be attributed to the fact that investment has taken place. Thus we can say that marginal efficiency of investment tends to decline as investment rises.

When the capital stock is K*, it earns a return at the rate of r*. As this rate may be found attractive to investors, and they may take up more net investment. The process increases the capital stock. As they increase capital stock, enhanced capital earns lower and lower rate of return because, given the other factors of production, an increase in capital alone will lead to a fall in marginal product of capital. This is called marginal efficiency of capital. As investment takes place to take stock of capital beyond K*, MEC starts declining. Thus, we can say that the ability of investment to generate income flows tends to decline as capital stock increases. Marginal efficiency of investment (MEI) is defined as that rate of return, which would render, discounted present value of future yields exactly equal to present cost of investment

The investor faces a problem. She knows that capital is earning a certain rate. As she takes up more investment, capital stock rises and then the rate of return earned falls. So, she must equate the falling rate of return with the cost of capital in the market. This cost will be nothing but opportunity cost of investment. Hence, she equates market rate of interest with expected rate of return on capital. The investment is taken up only till that level when expected rate of return is at least as large as market rate of interest. If expected rate is r* and market rate is 'r' than, investment continues till $r^* \geq r$.

We can interpret the concepts developed in forgoing paragraphs in a slightly different manner. The marginal efficiency of investment is called internal rates of return (IRR) as well. Investment continues so long as IRR > market rate of interest.

We can also regard MEI functions as demand for capital. The supply function of capital is the cost of capital, which is just a mirror image of MEI function. The market equilibrium will occur when MEI equals cost of capital. Fig. 12.6 shows market equilibrium.

Fig.12.6

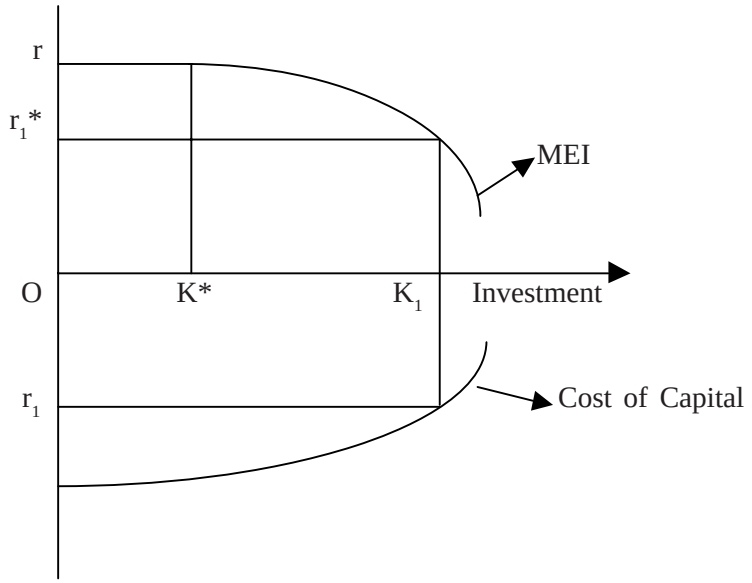


Fig.12.6: Investment raises capital stock from K* to K_1. Hence investment is OK_1 - OK*. Lower segment of the diagram shows cost of capital curve. r_1 is the market rate of interest while r^*_1 is IRR at this level of investment.

12.3.3 Real and Nominal Rates of Interest

We have to resolve one more problem: You lend some money to a friend at an agreed rate of interest. When the money comes back, say at the end of 5 years period, you get interest as well as principal. But, if prices have increased substantially during this period, how will you react to the changed purchasing power of your money? Let us make it more concrete. Say, you gave Rs.10, 000 to your friend. She agreed to return your money with a simple interest of 5% per annum at the end of 5 years. But, during the same period prices rise at, say, 12% per annum. You find that change of prices at the repayment time.

You get Rs.10000 plus Rs.2500 = Rs.12500

But prices double in the meantime. So despite the interest that you charged from your friend, you find that you cannot buy the bundle of goods, which had a cost of Rs.10000 initially with the total amount of Rs.12500. You may need something close to Rs.20000 now to be able to purchase that old bundle!

So this rate of 5% does not compensate you for the changes is purchasing power. Such a rate is called a nominal rate of interest. However, if you could change a rate of interest that exceeds the rate of price rise, only than you will be compensated for waiting for 5 years. Hence, a rate of interest that exceeds the rate of inflation will be real rate of interest. In the example above, if one desires a real rate of interest of 5% per annum, one must charge 17% per annum.

Check Your Progress 2

1) What is Simple Rate of Interest ?

.....

.....

2) Explain the idea of Compound Rate of Interest.

.....

.....

3) Explain the notion of Marginal Efficiency of Investment.

.....

.....

4) Explain the Concept of Present Value

.....

.....

5) What is rate of discount?

.....

.....

6) What is difference between real and nominal rates of interest?

.....

.....

12.3.4 Theories of Interest

There are a number of theories, which seek to determine the rate of interest. These theories try to explain the phenomenon of interest in terms of different set of variables.

a) **Loanable Funds Theory**

This theory relies on demand for borrowings and supply of loanable funds to determine the rate at which transaction will take place. It assumes that at any moment of time there will be some people who would spend less then their current income (savers) and others who plan to spend more than their income. The former will constitute the supplies of loanable funds while the latter constitutes the group,

which demands such funds. The rate at which demand for funds equals supply of funds will be the rate of interest. Such a situation is depicted in Fig. 12.7

Fig. 12.7

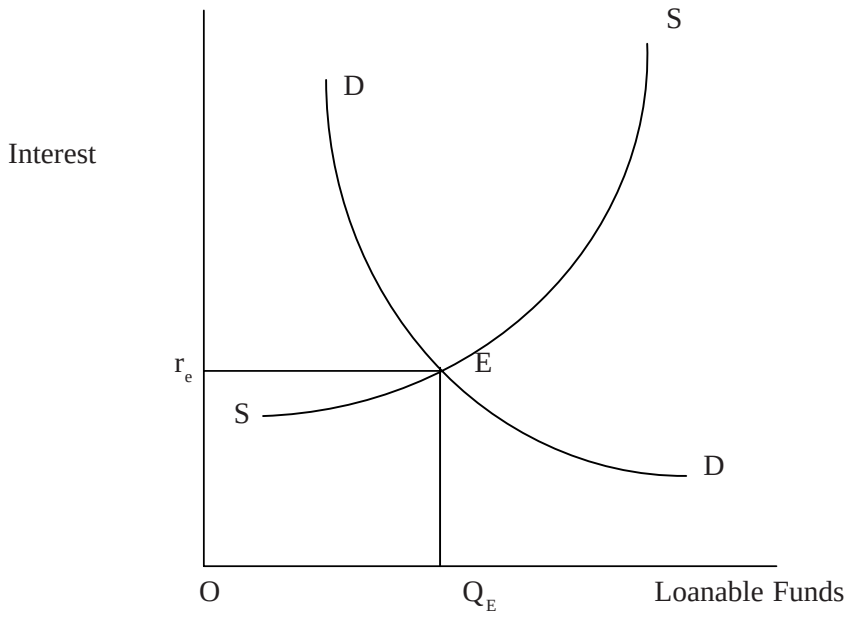


Fig. 12.7: Presents a simple demand and supply curve diagram you are so familiar by now. The curve DD is demand curve for the funds. This shows amounts the borrowers would like to borrow at different rates of interest. Likewise, the amounts all the savers in the society are willing to save and lend are shown by supply curve marked SS. The intersection of these two at point E gives us equilibrium rate of interest re and the quantity Q_E that will be borrowed and lent at that rate.

At r_e rate of interest, Q_E quantity of funds is borrowed (and lent). Note that demand for funds may arise on account of any three of the following:

(i) Investment demand, (ii) consumption demand and (iii) financial demand. It is more likely to be a composite of all the three demands.

Similarly, we can say that supply of funds may arise from net savings, de-hoarding of past savings and also from new creation of money.

b) **Liquidity-Preference Theory**

Keynes had developed this approach and he related demand for money and rate of interest to aggregate level of income in the society. In his formulation demand for liquid money would depend on transaction, precaution or speculation, given the level of income. But supply of money was a policy determined variable. The rate of interest was thus determined by interaction of a demand function with a given supply of money. However, in his approach, the rate of interest has nothing to do with determination of rate of remuneration of a factor of production. We have given a more detailed explanation of liquidity preference approach in the subsequent units.

c) **Time Preference Approach**

Irving Fisher developed this approach. His idea was that consumer tries to compare present consumption and future consumption. The rate at which future consumption

can substitute for present consumption (and vice-versa) will be marginal rate of substitution between present and future consumption. This is called the rate of time preference. It shall be equal to slope of indifference curve between present and future consumption.

Check Your Progress 3

1) Explain Loanable funds theory. (50 words)

.....

.....

2) Explain Time preference theory. (50 words)

.....

.....

12.4 REWARD FOR ENTREPRENEURSHIP: PROFIT

We regard entrepreneurship to be the fourth economic factor of production. Recall that an entrepreneur brings together land, labour and capital and thus facilitates production. Her role in production is clear. If other factors of production are not brought together, there may not be any production at all. In capitalist system, the possibility of profit becomes key determinant of whether an activity will be undertaken or not. Even under various non-capitalistic forms of organisation, profit may serve as a benchmark for efficiency of firm or efficiency of some innovation or technological change. Thus, in all situations, if a firm is making larger profit compared to some other similarily placed firm, it must be more efficient or must be using either better resources or better techniques. But decision like introduction of better techniques involves some risk as well. Hence, often attempts are made to relate profit to elements of uncertainty and risk. To understand her role, we can divide entrepreneurial functions into two parts:

a) Organisation and

b) Risk bearing.

a) **Organisation:** This consists of routine day-to-day activities associated with a business organisation and is called management. We find that these days, most companies are being managed by professional managers, who receive salaries and other benefits. Such an arrangement places a part of entrepreneurship at par with labour.

b) **Risk Bearing:** Every business activity runs some risk of failure in the market. This arises because of uncertainty of marketplace, natural causes, political factors etc. If a business fails, the entrepreneur looses substantial parts of investment. Thus, risk of loss is always present. However, some activities like introducing a new product, using a new technology etc., involve much greater risks and reward for these activities must be higher. Otherwise, these would not be undertaken. Hence it is said that profits are reward for risk bearing.

12.4.1 Accounting Profits and Economic Profits

An accountant defines the profit as the difference between total revenue earned during the year and cost (including depreciation) incurred during the same period. The cost comprises payments for raw materials, fuels / energy, wages and salaries, rents, insurance and interests. The depreciation is provided for taking care of wear

and tear of capital stock. So the net surplus earned during the year, after meeting the above costs, is called profit by the accountant.

However, such calculations do not seem to account for some implicit costs. Take for example the remuneration to the person when she is actually working for her business. Similarly, companies accumulate some funds of their own in course of time. Should interest on those funds be also calculated and added to the cost? Economic profit will take into account this kind of implicit cost as well. So economic profit will be less than accounting profit by the amount of such implicit costs.

12.4.2 Theories of Profits

Economists have, over the years, developed several theories regarding profits. For example, Joseph Schumpeter attributed profits to innovation. But Frank Knight associated them with uncertainty.

a) **Profits as Rewards for Innovation**

Schumpeter regards profit a phenomenon, which is related to a dynamic economy only. He identifies five types of changes that lead to economic development or make the society dynamic. These changes are:

i) Introduction of new products

ii) Introduction of new methods of production

iii) Discovery of new raw materials

iv) Discovery of new markets

v) Introduction of new forms of organisation

Innovations are actual application of some new body of knowledge to real business situation. An innovator need not be an inventor. But she uses some invention to change her production function or the relationship between inputs and output. Such innovation might be in form of new technique of production, may involve reaching out to new markets, involving all the activities pertaining to marketing etc.

Schumpeter is of the opinion that one who innovates is able to earn more profits, and thus gets more incentive to innovate further. She will soon attract followers or imitators. These people, very soon catch up with original innovator. As a consequence, she makes more efforts to stay ahead. Thus, innovation leads to profits and profits make it possible to innovate (acting as incentive)

b) **Uncertainty and Profit**

Frank Knight defined profit as the difference between selling price and costs. In such situation profit emerges as a residual. Selling price and costs depend on a host of factors. Some of those can be covered by 'risk'. Such risks can be anticipated and provisions can be incorporated into the cost structure. Most of predictable risks are 'insurable' as well. Hence, company can get an appropriate insurance policy to cover such risks. The premium paid for such policy is included in cost of production. This type of risk condition is completely predictable and discountable. Hence it would be as good or as bad as production under perfect certainty.

But Knight points to another dimension of uncertainty and says that producer is all the time anticipating consumer's wants and preferences in advance. She must do so, as she has to produce things that can satisfy those wants at a point of time

in future. This essentially happens because of time lag involved between anticipation of demand, production and offering goods to consumers. To some extent, future results of her operations to produce things to satisfy that demand are also uncertain. Further, even the manager doing routine organisation work is liable to make error of judgement. Here, she bears uncertainty and risk in the sense of having to protect factors of production against fluctuation in their income from an uncertain market. Thus, the income of entrepreneurs consists of two components, a salary or wage component, which is contractual in nature and another residual income that may fluctuate in response to change in market place. Some economists prefer to call only this second component as 'profit'.

Thus we find that one significant difference between other factor incomes and profit. Whereas wage, rent interest are all payments, which have been agreed to and settled in advance, profits cannot be put on a similar footing. Uncertainty leads to fluctuation in both costs and revenue. They may not balance. Thus, ultimately profits are the 'surplus' that remain after meeting the entire contractual payment obligation.

c) **Profits and Market Structure**

Some economists insist that profit as one generally understood is essentially a result of market imperfections. If perfect competition prevailed, every producer will use same technology, will have perfect knowledge about product, cost and market condition. Such a scenario leads to cost minimisation for all the production. They sell at going market price. All the cost and revenue determinants are perfectly certain. Hence, entreprenureship is just organisation or day-to-day supervision only. So, profits should drop down to bare minimum or 'normal' compensation for supervision etc.

However, if market is not perfect, firm can determine quantities or prices in such a manner that suits it best. It may involve breaching the condition of perfect information. Firms may device some innovation and keep it a secret from others. So long as that secret is maintained, the concerned firm continues to earn more than others do.

A. P. Lerner tried to measure the effect of monopoly power over profit. We know that equilibrium condition for a firm is equality between marginal cost and marginal revenue. When competition is perfect, price (average revenue) is also equal to marginal revenue. Prices tend to deviate from marginal revenue only when competition is no longer perfect. Hence, the difference between price and marginal revenue (= MC), that is, P - MR (or P-MC) will indicate firms control over market. It is expressed as a fraction of price. Thus, the degree of monopoly is (P - MC) / P. Higher this ratio, higher will be the rate of profits earned by a firm.

Check Your Progress 3

1) What are the main functions of an entrepreneur?

.....

.....

2) Why could profit be 'minimal' or 'normal' under perfect competition?

.....

.....

3) Why should profit as 'surplus' beyond under perfect competition?

.....

.....

12.5 LET US SUM UP

The present unit has three distinct sections. We considered remuneration for the services of the land, or the rent in Section 12.2. We considered Ricardian, Marshall's and modern theories of rent. We also explained the notion of quest-rent, which can be earned by any factor in short supply during the short run. Section 12.3 discussed the concept of interest. We considered some usually encountered notions and various explanations for the accrual of interest. We also tried to relate interest with capital and investment and concluded by paying attention to some important theories of interest. The reward for entrepreneurship, or, profit has been discussed in Section 12.4. Here, entrepreneurial function, effects of uncertainty and some important theories have been given a brief exposition.

12.6 KEY WORDS

Contractual Rent : Gross payment agreed to between owner and user of a resource.

Credit : Allowing someone to use goods/services without immediate payment.

Discount : Amount of money deducted from the face value of an asset.

Economic Profit : Reward for enterprise is calculated by deducting all implicit costs from accounting profits.

Implicit Costs : Cost of the factors, which a producer does not hire. These may be owned by her.

Innovation : Practical application of an invention - usually involves a new technique, material or form of market or organisation.

Normal Profits : Minimum profit to reward an entrepreneur. It is usually associated with perfect competition.

Quasi-Rent : Return to a factor over and above its average cost - a short-run phenomenon.

Transfer Earnings : The minimum payment needed to retain a factor in present employment. This is what a factor can get in next best (alternative) employment.

12.7 SOME USEFUL BOOKS

Stonier A.W. and Hague D.C.(1980), *A TextBook of Economic Theory*, Macmillan: London

Lipsey. R. G., *An Introduction to Positive Economics.* (6th edition), E.L.B.S. and Weidenfeld and Nicolson: London.

12.8 ANSWERS/HINTS FOR CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) Read Sub-section 12.2.1 and answer

2) Read Sub-section 12.2.1 and answer

3) Read Sub-section 12.2.1 and answer

4) Read Sub-section 12.2.3 and answer

Check Your Progress 2

1) Read Sub-section 12.3.1 and answer

2) Read Sub-section 12.3.1 and answer

3) Read Sub-section 12.3.2 and answer

4) Read Sub-section 12.3.2 and answer

5) Read Sub-section 12.3.1 d and answer

6) Read Sub-section 12.3.3 and answer

Check Your Progress 3

1) Read Sub-section 12.3.4 and answer

2) Read Sub-section 12.3.4 and answer

Check Your Progress 4

1) Read Sub-section 12.4.1 and answer

2) Read Sub-section 12.4.2 C and answer

3) Read Sub-section 12.4.2 C and answer

UNIT 13 CIRCULAR FLOW AND NATIONAL INCOME

Structure

13.0 OBJECTIVES

After going through the unit you would be able to explain

- meaning of the term 'circular flows';
- the distinction between money flows and real flows;
- derivation of national income from circular flows;
- the meaning of production, income generation and expenditure flows; and
- the relationship among various macro-economic aggregates.

13.1 INTRODUCTION

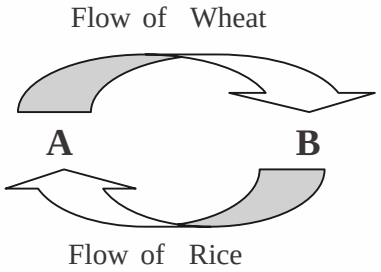
An economy operates with the help of economic categories like producers, consumers, government, capital sector and rest of the world. These categories perform various economic activities comprising production, consumption, income generation and addition to capital stock and economic transactions with the rest of the world. In the process of performing such economic activities, goods and services flow from one group of agents to another and vice-versa. Corresponding to each such flow, there takes place a counter monetary flow. For example, if one person gets 2 kg. of sugar from a firm, a commodity-flow from a firm to a household is taking place. This flow is matched by a monetary flow, from the household to the firm. Such flows, if aggregated at various ends, can be summed up as national income, gross domestic product etc.

The knowledge of these circular flows along with national income and various other related macro economic aggregates is essential for understanding macroeconomic theory which deals with the determination of levels of national income, employment and prices.

13.2 CONCEPT OF CIRCULAR FLOW

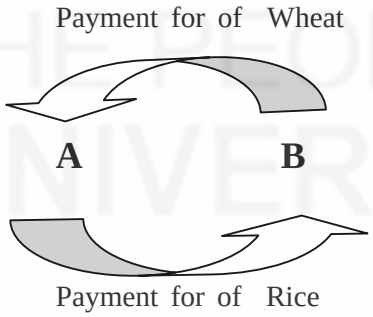
The concept of circular flow pertains to the flow of real transaction or money transaction from one economic agent to another. The flow is not one-sided; it is two-sided. Because of this feature it can be termed as circular flow. Suppose person A gives wheat to B and person B in turn gives rice to A, then this can be termed as circular flow which is shown below:

Fig. 13.1



In Fig. 13.1 the direction of the arrows shows the receiving agent. For example, B is receiving wheat from A and therefore, the arrow is pointing towards B. Similarly, A is receiving rice from B. Thus, the arrow is pointing towards A. In the above example, goods have been exchanged so the flows can be referred to as real flows. Instead of goods, if money was exchanged, the flows could have been money flows. Note from the example that when B received wheat from A money would be given by B to A. Similarly, A would have given money to B for the purchase of rice. These money flows can be shown as below:

Fig. 13.2



Comparing Figures 13.1 and 13.2, we would notice that real flows take clockwise movement, i.e., from left to right. On the other hand, money flows take anti-clock-wise movement, i.e., from right to left.

13.2.1 Difference between Money Flows and Real Flows

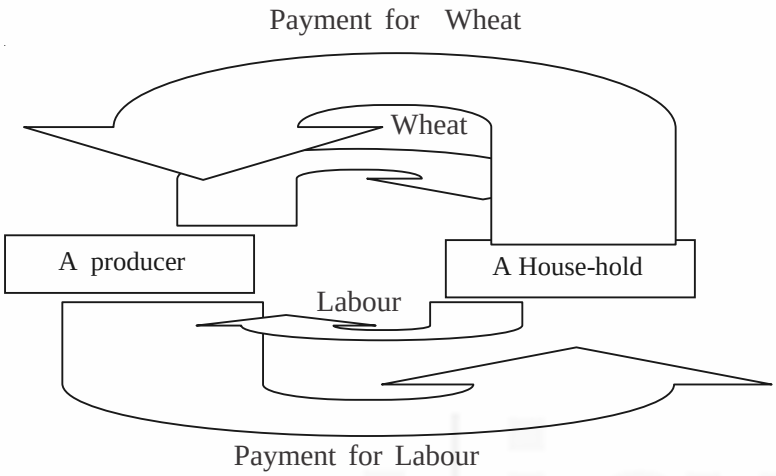
The distinction between money flows and real flows should be clearly understood. Real flows are the flows of goods from one transactor to another and vice-versa. Similarly, real flows can be flows of services from one transactor to another and vice-versa. Real flows are difficult to measure as they comprise bundles of goods or services, expressed in different units and it is impossible to aggregate these transactions or flows. It is precisely because of this reason that we measure money flows.

Money flows, as the name suggests, show the flow of money from one trasactor to another. Suppose transactor A supplies goods to transactor B. That is a real flow. The transactor B, in turn must have paid for these goods to transactor A, which is a money flow. Similarly, transactor B may have supplied labour services

or services of land to transactor A which is a real flow. Transactor A, in turn, must have paid for these factor services in the form of wages to transactor B that would be money flow.

The distinction between money and real flows and their interaction can be very well shown with the help of a diagram such as Fig. 13.3, where transactor A is represented as a producer and transactor B as a household.

Fig. 13.3



In Fig. 13.3, a producer supplies wheat to a household. The director of the arrow indicates who receives the goods. Similarly, the household supplies factor services to a producer as shown by the arrow. Note that clock-wise director of the arrow indicates real flows.

Corresponding to real flows we can also see money flows taking place in the opposite direction or in an anti-clock-wise direction. For instance, for the goods supplied by the producer to a consumer, the consumer has paid for these goods in money terms, which can be called consumption expenditure. Similarly, the producer has paid for these factor services. We can call these factor payments. Remember that anti-clock-wise arrows indicate money flows.

It is important to realise that a barter economy where goods/services are exchanged for goods/services will have only real flows. On the other hand, in an economy where goods/services are exchanged for money and then money is exchanged for goods we will have real as well as money flows. It is also possible that in a modern economy we may have only money flows taking place without any corresponding real flows. For instance, if a father gives pocket money to his son, money flow may take place from father to son. But son, in turn, has not supplied anything in return and thus the circular money flow is not complete.

Can we think of some cases where the circular money flows may complete circular movement?

13.2.2 Flows between Enterprises and Households

Various transactions taking place among transactors or economic agents can be better understood when put in the form of flows.

An enterprise is an economic agent, which employs factor services supplied by households. It creates goods and services, which may either be supplied to other firms in the form of raw materials, or produce consumer goods meant for the final

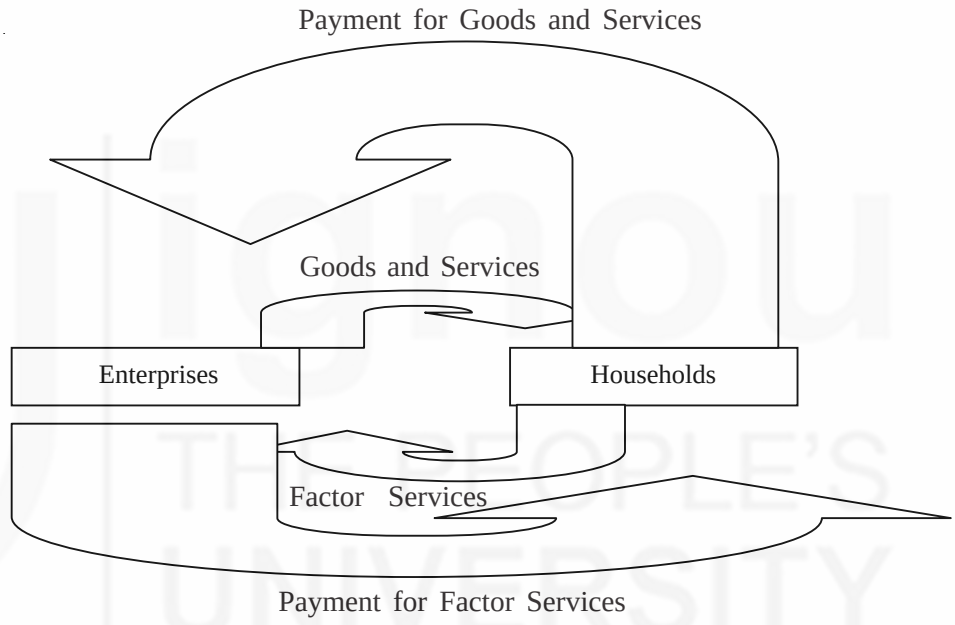
consumption. It may produce machines/plants to help in the creation of more goods and services.

Similarly, we can define a household, which, by definition, supplies the factor services such as land, labour capital and entrepreneurship to enterprises and consumes consumer goods and services produced by enterprises.

The distinction between households and producers is not always mutually exclusive. A person can be a household as well as a producer. To take an illustration, a teacher is a producer when she produces teaching services and will be a household when she buys or consumes the goods and services produced by other producers. Thus, the distinction is not personal, but functional in nature.

The flows between enterprises and households can be shown with the help of Fig. 13.4.

Fig. 13.4



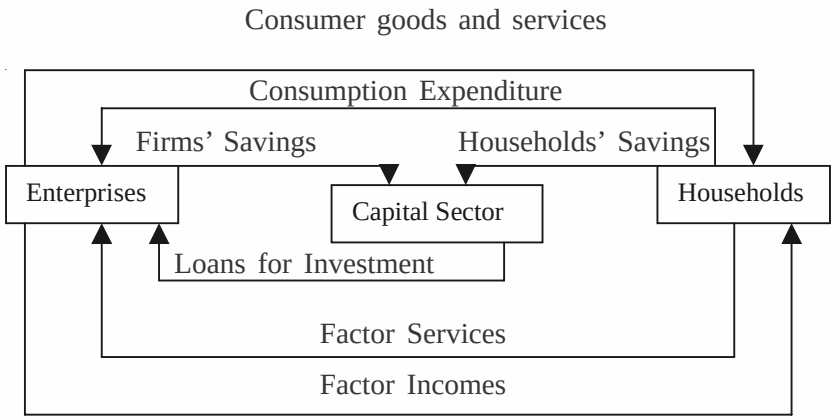
In this figure, both real and money flows are shown. The flow of consumer goods and services from enterprises to households and of factor services from households to enterprises constitute real flows. Similarly, flows taking place from consumers to producers in the form of consumption expenditure and from enterprises to households in the form of factor incomes relate to money flows. It would not be out of place to state that money flows are the counterparts of real flows. Note that Fig. 13.4 is not much different from Fig. 13.3. In Fig. 13.3 we had shown transaction between one firm and one household only, now all the producers and consumers have been added together to make two groups.

13.2.3 Flows between Enterprises, Households and Capital Sectors

So far we have discussed flows in a situation where there is no saving and investment. To introduce saving and investment we have to include capital sector along with enterprises and households.

Capital sector collects savings of various sectors and lends these to enterprises for investment. The introduction of capital sector along with enterprises and households is illustrated in Fig. 13.5.

Fig.13.5



In Fig. 13.5 , the flows between enterprises and households are shown as in Fig. 13.4. The additional flows shown here are between (i) households and the capital sector and (ii) enterprises and the capital sector. The factor income received by households need not be fully used for final consumption expenditure; a part of the incomes may be saved in banks or funds used for buying shares, or buying, say, an LIC policy which are all considered a part of capital sector. Thus, the arrow from households towards capital sector is indicative of the savings flowing from household sector to capital sector. These savings are collected by capital sector from households along with savings of the enterprises in the form of undistributed profits, accumulated for expansion purposes, and depreciation fund used for replacement investment purposes.

The savings of capital sector are employed to finance gross investment of the economy, which is shown, with the help of arrow from capital sector towards enterprises. The investment activity of the economy is undertaken by the enterprises, which produce capital goods for net accumulation of capital stock, or for replacing the worn-out capital. Savings in national income accounting are so defined asto be equal to investment.

13.2.4 Flows between Enterprises, Households, Capital and Government Sectors

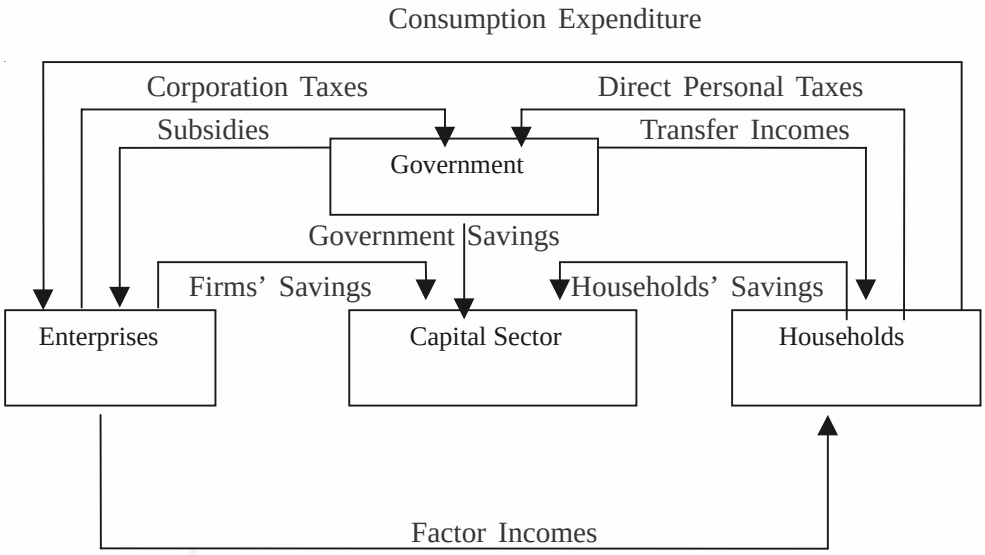
Flows between enterprises, households and capital sector were shown in Fig. 13.5 above. Let us now introduce the government sector.

The government sector can be viewed in two ways: First, the government can act as a producer or enterprise meaning thereby that it can contribute to total production activity along with the private sector. Second, the government can act as a redistributor of incomes, i.e., tax a particular sector of the economy and subsidise another either by giving cash help to the enterprises or offering transfer incomes in the form of old-age pensions or unemployment benefits etc., to households. The government production activity can be categorised as, 'General Government' where the government produces primarily services which are collectively consumed. These can take the form of police, or defence services, which are ordinarily not available for sale; instead they are consumed collectively and are financed from the tax revenue raised by the government. The government has undertaken departmental and non-departmental enterprises, which produce goods and services for sale in the market, are grouped in the category of enterprises. Therefore, General Government is only the activity of the government as a redistributor of income or producer of services, which is meant for collective consumption.

An incorporation of the government sector in flow chart is shown with the help

of Fig. 13.6. We are introducing a change in our presentation here: Now onwards, we will show only the monetary flows without the corresponding real flows. This is being done to keep the diagrams simple in appearance.

Fig. 13.6



In Fig. 13.6, the flows between enterprises, households and capital sector are same as shown in Fig.13.5. In addition, however, Fig. 13.6 has with an additional flow, namely, a part of the saving may come from government sector to the capital sector. These savings may be positive or negative. If the government expenditure on transfer incomes, subsidies or maintenance of collective services is more than the tax revenue raised in the form of direct personal tax, indirect taxes and corporation tax, savings are negative. On the other hand, if government expenditure were less than tax revenue, government saving contributed by the government sector would be positive.

The income received by households for the supply of factor services to the government sector need not be spent only on the purchase of consumer goods produced by enterprises. A part of the factor income may be passed on to the government in the form of direct personal taxes as indicated by the arrow towards government from households. On the other hand, the government can give transfer incomes to households as shown by arrow facing households from government. Similarly, the incomes received by enterprises from the sale of consumer goods to households or for meeting collective consumption of government, may get leaked to government in the form of indirect taxes and corporation taxes as shown by the arrow facing government originating from enterprises. The government may also use tax revenue to subsidise production of goods and services by enterprises. This is shown by the arrow facing enterprises starting from government.

13.2.5 Flows in an Open Economy

So far we have shown the flows in a closed economy, i.e., an economy that does not have any transactions in the form of exports and imports. Now, we introduce transactions of an economy with the other countries of the world (referred to as rest of the world) to get flows in an open economy (an economy which has economic transactions with the rest of the world).

When an economy is opened up, the following variables have to be incorporated in the flows of an economy:

1) A part of the output produced by enterprises of the economy may be retained for consumption or investment purposes within the economy and the rest may be exported to the rest of the world. The payment for such exports is made by rest of the world to enterprises of the domestic economy.

2) Consumption expenditure of households may be not only on goods and services produced within the economy but also on those imported from the rest of the world.

3) Households may earn factor incomes not only from domestic enterprises but also from normal residents of an economy who are temporarily, upto one year, stationed in other countries. Similarly the normal residents of the rest of the world may temporarily be stationed within the economy in question and therefore factor income earned by them is a part of the national income of the country of which they are the normal residents. This together gives us the variable called 'net factor income from abroad' which can be positive or negative. It is positive if factor income earned by the normal residents of an economy in the rest of the world is more than factor incomes earned by the normal residents of the rest of the world stationed in the economy.

4) Another factor to be considered is the fact that savings accumulated in the capital sector may not originate from household's enterprises or the government. A part of the savings may flow from the rest of the world which is termed as 'net capital inflow from the rest of the world,' which may be positive or negative. It is positive when borrowings from rest of the world are more than lending to rest of the world and negative when lending exceeds borrowing.

5) Another factor is that savings generated within the economy and from the rest of the world may be used not only for generating gross domestic capital formation (replacement investment plus net domestic capital formation) but also for the purposes of net investment aboard, which can be positive or negative. It is positive when investment made by the economy in the rest of the world is more than investment made by the rest of the world in the economy in question and negative in case of a reverse situation.

6) Lastly, just like there can be unilateral transfers (which do not have quid-pro-quo) within the economy there can be unilateral transfers from the rest of the world to the economy and vice-versa. The variable is termed as 'net current transfers from the rest of the world', which can be positive or negative. It is positive when current transfers from the rest of the world to the economy in question are more than current transfers by the economy to the rest of the world are and negative when the reverse is true.

In Fig. 13.6, the flows in a closed economy were shown and the corresponding flows in an open economy are shown in Fig. 13.7. The incorporation of variables arising because of opening up of the economy would bring in quite a difference to the flows of a closed economy shown in Fig. 13.6.

Enterprises receive money not only through consumption expenditure of households but also by net exports of goods and services of enterprises. Net exports are the difference between exports and imports. It can be positive or negative. It is positive when exports are more than imports and negative when reverse is true. In the figure, the arrow originating from the rest of the world to enterprises indicates exports whereas the arrow originating from enterprises to the rest of the world indicates imports.

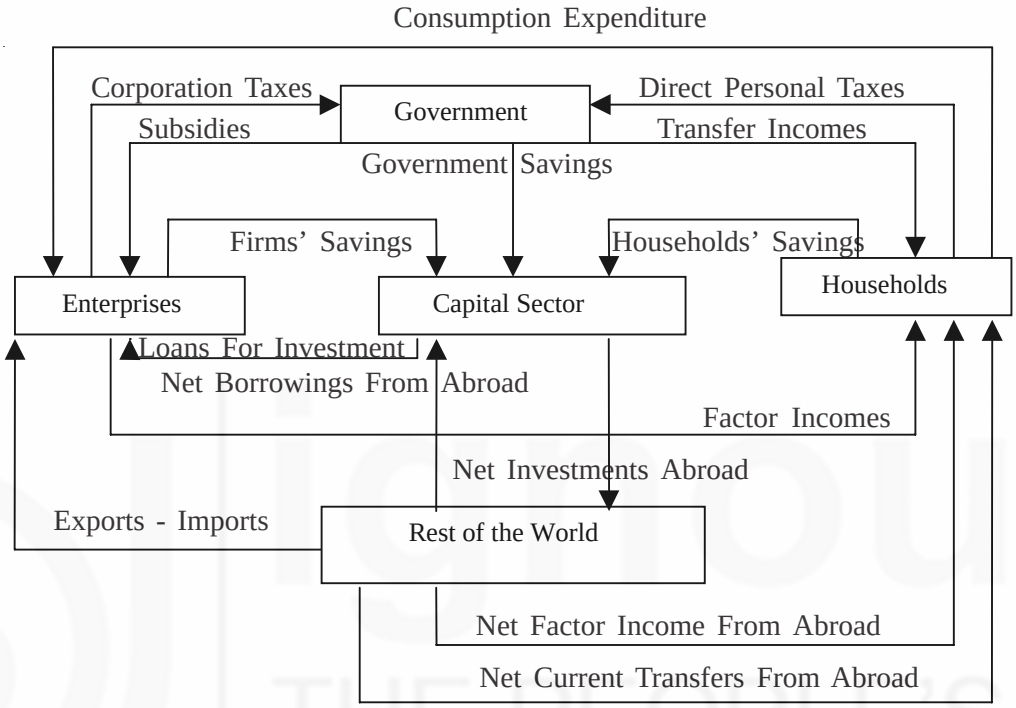
Similarly, 'net factor income from aboard' is shown with the help of an arrow pointing towards households from the rest of the world. The same is true of 'net

current transfers from aboard' where arrow is pointing towards households from rest of the world.

Net borrowings from the rest of the world are indicated by an arrow pointing towards capital sector starting from the rest of the world.

Finally, net foreign investment is shown with the help of an arrow pointing towards the rest of the world from capital sector.

Fig. 13.7



Thus, Fig. 13.7 presents a full picture of the flows taking place in an economy, which has enterprises, households, government, capital sector and the rest of the world sector as the main players. The situation gets more complicated if each of the sectors is sub-divided into smaller units. For instance, enterprises sector is to be divided into a number of enterprises; household sector into individual households; the capital sector into banks, LIC, GIC, share market etc., government into Central Government, State Governments, Local Governments and the rest of the world into a number of countries. The situation will get complicated because inter-unit transactions of a sector are also to be accounted for.

Check Your Progress 1

1) Distinguish between money flows and real flows with the help of suitable illustrations.

.....

.....

.....

2) State various economic transactions, which are used to study circular flow of an economy

.....

.....

.....

3) How does circular flow get complicated when capital sector is introduced along with enterprises and household sectors?

.....

.....

.....

.....

13.3 CIRCULAR FLOWS AND NATIONAL INCOME

The circular flows presented in Section 13.2 are essential for the purposes of visualising the working of an economy. By studying these flows it is possible to derive various macro-economic aggregates. Some of these aggregates are gross domestic product (GDP), net domestic product (NDP), gross national product (GNP), net national product (NNP) and national income (NY). In the following sub-sections let us try to derive these aggregates from circular flows. We would employ Fig. 13.7 to arrive at NY in three phases, viz., as flow of goods and services, as flow of factor incomes and as flow of final expenditures.

13.3.1 National Income as Flow of Goods and Services

Taking a fresh look at Fig. 13.7, we can try to view NY aggregate at the enterprises end. If we add up the money value of the flow of goods and services produced over a year without duplication, after deducting the production of capital goods meant for replacing worn out capital stock and adding with this net factor income from abroad, it would be possible to get a figure of NY of an economy. To spell out further, enterprises produce consumer goods (C) and take up net domestic capital formation (I). Add with this net factor income from abroad to get NY of an economy. Thus we can say Y = C + I + NFIA where Y is NY. In this definition it may be kept in mind that the value of goods and services have to be computed at factor cost (FC) and not at market price (MP), where value at MP = Value at FC + NIT. NIT is net indirect taxes i.e., indirect taxes minus subsidies. Moreover it is also to be seen that those goods and services which are produced by enterprises meant for the purposes of intermediate consumption (raw materials bought by one enterprises from another are not to be added along with goods and services meant either for final consumption on the part of households or for adding to the total capital stock of the economy. This needs to be done to avoid double counting. To give an example, if we take the total production of wheat along with total production of bread we would indulge in duplication since bread also includes wheat in the form of flour.

This way of calculating NY is known as production method, or product method. Production method as we will see subsequently is also known as value added method.

13.3.2 National Income as Flow of Factor Incomes

Again, going back to Fig. 13.7, let us view NY aggregate, at the households' end. Households supply factor services to enterprises to produce goods and services. These factor services can be supplied by four factors of production, viz., labour, land, capital and enterprise are to be remunerated in the form of wages, rent, interest and profits, respectively.

Thus, by adding up wages, rent, interest and profits along with net factor income from abroad we get NY of an economy. Or, Y = W + R + In + P + NFIA, where Y is national income which is equal to the sum of wages (W), rent (R), interest (In),

profits (P) and net factor income from abroad (NFIA). What we have done is to add up all the factor incomes received by households for supplying factor services to enterprises.

By definition, national income viewed, as flow of final goods and services is identical to the one viewed as flow of factor income generated in the process of production.

Sometimes factor incomes instead of being categorised into W, R, In and P are put differently. In this new way of categorisation of factor incomes we have the groups, viz., compensation of employees (CE), operating surplus (OS) and mixed income of self-employed (MY) so that Y = CE + OS + MY + NFIA, where CE is compensation given to labour for rendering labour services, OS is factor income generated by the ownership and to distinguish between CE and OS and NFIA is net factor income from aboard (already defined). While CE and OS are easy to understand, MY requires a little explanation. MY arises in the case of self-employed. For example, if we ask a petty shopkeeper, running a shop at her own place of residence using her own capital, herself working as labourer or manager and herself undertaking the risk of running the business out of the income of such an enterprise, how much is rent, interest, wages and profits? It would not be possible for her to categorise her income under various heads. Such factor incomes instead can be put under the label mixed income of self-employed.

13.3.3 National Income as Flow of Final Expenditures

Making use of Fig. 13.7 national income can also be viewed as sum of final expenditures of various transactors of an economy. In other words, this time we look at not the production of final goods and services but how they are disposed of. The various heads of final expenditures can originate from either households in the form of private final consumption expenditure (C_h) or from government in the form of public final consumption expenditure (C_g) or from firms in the form of purchase of net capital goods (NDKF) and change in inventories (K) or from the rest of the world in the form of purchase of net exports (NE).

Change in inventories is defined as stocks of finished goods or raw materials/semi-finished products at the end of the year minus the stocks of these goods at the beginning of the year. Change in stocks is positive if closing stocks are more than opening stocks and negative in case of reverse situation.

We are now in a position to arrive at NY viewed as flow of final expenditures as equal to: $Y = C_h + C_g + NDKF + NE$. Since NY is defined as NNP at FC, it is important to make final adjustment in the above equation by deducting net indirect taxes (NIT); since C_h, C_g, NDKF and NE are normally presented at market price, in order to convert these figures at factor cost, NIT has to be deducted.

In the final reckoning the equation of national income as flow of final expenditure will be $Y = C_h + C_g + NDKF + NE - NIT$.

13.3.4 National Income viewed as Production, Income and Expenditure Flows

We are in a position to state now that in Sections 13.3.1,13.3.2 and 13.3.3 national income aggregate is arrived at considering circular flows from production, income and expenditure respectively. Since production of goods and services requires factor services and factor incomes are generated and such incomes are disposed of for financing final consumption expenditure or saved. Savings, in turn, are used to finance capital formation activity within the economy or for financing net foreign investment.

These three ways of measuring will give us the same magnitude of national income, provided full statistical data required are available. In reality, the requisite information may not be available because of which we are forced to employ a combination of these three methods to measure national income of an economy.

The first step to measure national income of an economy is to divide the economy into a number of industrial sectors like agriculture, mining, logging, manufacturing, construction, real estate, government services, transport services, commercial services etc. Then depending on the availability of data we decide which method to employ. For example, in agriculture and manufacturing sectors production figures may be more readily available and thus we find out the contribution of these sectors by employing production or value added method; for construction sector expenditure figures may be more easily available and its contribution to national income can be estimated by using expenditure method. Finally, for services sector, figures of incomes generated during a year are more easily available which necessitates the employment of income method.

Check Your Progress 2

1) Show how Production Flow, Income Flow and Expenditure Flow in an economy are related to each other.

.....

.....

.....

.....

2) State the main components of National Income as flow of :

i) Currently produced goods and services.

.....

.....

.....

ii) Currently generated factor incomes

.....

.....

.....

3) Why does India employ a combination of production, income and expenditure methods to measure its national income?

.....

.....

.....

13.4 NATIONAL INCOME AGGREGATES

National income is a macro-economic aggregate, which is indicative of economic progress of an economy. There are a number of other related concepts of equally importance and one should clearly understand the inter-relationship among various macro-economic aggregates. These would be extensively employed in units of Block-7.

13.4.1 National Income and Various Related Concepts

Some of the important related concepts of national income are as follows:

1) **Gross National Product at market price (GNP_{mp})** : It is the sum of the values of currently produced goods and services without duplication, over a year, by the normal residents of an economy, gross of depreciation, where goods and services are valued at the market prices.

2) **Gross National Product at Factor Cost (GNP_{fc})** : It is the sum of the value of currently produced goods and services, over a year, by the normal residents of an economy, gross of depreciation, when goods and services are valued at factor cost (market price minus net indirect taxes).

3) **Net National Product at Market Price (NNP_{mp})** : It is the sum of value of currently produced goods and services without duplication, over a year, by the normal residents of an economy, net of depreciation, where goods and services are valued at market price.

4) **Net National Product at Factor Cost (NNP_{fc})** : It is the sum of value of currently produced goods and services without duplication, over a year, by the normal residents of an economy, net of depreciation, where goods and services are valued at factor cost (market price minus net indirect taxes).

5) **National Income (NY)** : It is the same as NNP_{fc}.

6) **Gross Domestic Product at Market Price (GDP_{mp})** : It is the sum of value of currently produced goods and services without duplication, over a year, within the domestic territory of an economy, gross of depreciation, valued at market price.

7) **Gross Domestic Product at Factor Cost (GDP_{fc})** : It is the sum of the value of currently produced goods and services without duplication, over a year, within the domestic territory of an economy, gross of depreciation, where goods and services are valued at factor cost (market price minus net indirect taxes).

8) **Net Domestic Product at Market Price (NDP_{mp})** : It is the sum of value of currently produced goods and services without duplication, over a year, within the domestic territory of an economy, net of depreciation, where goods and services are valued at market price.

9) **Net Domestic Product at Factor Cost (NDP_{fc})** : It is the sum of value of currently produced goods and services without duplication, over a year, within the domestic territory of an economy, net of depreciation, where goods and services are valued at factor cost (market price minus net indirect taxes).

10) **Net National Disposable Income (NNDY)** : It is the factor and transfer incomes earned or enjoyed by the normal residents of an economy, over a year, inclusive of net indirect taxes. It is identical to NNP_{mp} + net current transfers from rest of the world.

11) **Income from Domestic Product Accruing to Private Sector (Z)** : It is the factor income enjoyed by households and private corporate sector in an economy, over a year.

12) **Private Income (PY)** : It is the factor income and current transfers within the economy along with net current transfers from the rest of the world enjoyed by the normal residents of an economy, over a year.

13) **Personal Income (Personal Y)** : It is factor income and current transfers within the economy along with net current transfers from the rest of the world enjoyed by households of normal residents of an economy, over a year.

14) **Personal Disposable Income (PDY)** : It is factor income and current transfers within the economy along with net current transfers from the rest of the world, net of personal direct taxes and other administrative payments, at the disposal of households of normal residents of an economy, over a year.

15) **Personal Consumption Expenditure (C_h)** : It is personal disposable income minus personal savings (i.e., savings of households), over a year.

13.4.2 Interrelationships among various Macro-economic Aggregates

In Section 13.4.1 various national income and related concepts are introduced. In this section, it is the interrelationships among these aggregates, which will be introduced.

GNP_{mp} – Net Indirect Taxes (NIT) = GNP_{fc}

GNP_{fc} – Depreciation (D) = NNP_{fc}

NNP_{fc} – net factor income from abroad = NDP_{fc}

NDP_{fc} + NIT + NFIA + net current transfers from RoW (NCT from RoW) = NNDY

NNDY – X – NCT from RoW – NIT = NDP_{fc}

NDP_{fc} – Income from domestic product accruing to Government administrative departments – savings of non-departmental enterprises = income from domestic product accruing to private sector (Z).

Z + NFIA + national debt interest + transfer payments by government administrative departments + Net Current Transfers from RoW = Private Income

Pvt. Y – undistributed profits of private corporate sector – corporation tax = Personal Y.

Personal Y – direct personal taxes – miscellaneous receipts of government administrative departments = PDY

PDY – personal consumption expenditure = Household savings.

Household savings + private corporate savings + government savings + depreciation = Gross domestic savings

$$\frac{\text{Gross domestic saving}}{GDP_{mp}} \times 100 = \text{Rate of gross domestic saving}$$

Gross domestic capital formation = depreciation + net domestic fixed capital formation + change in stocks.

$$\frac{\text{Gross domestic capital formation}}{GDP_{mp}} \times 100 = \text{Rate of gross domestic capital formation}$$

Gross domestic savings – Depreciation = net domestic savings

$$\frac{\text{Net domestic saving}}{\text{NDP}_{mp}} \times 100 = \text{Rate of net domestic saving}$$

Gross domestic capital formation – Depreciation = net domestic capital formation

$$\frac{\text{Net domestic capital formation}}{\text{NDP}_{mp}} \times 100 = \text{Rate of net domestic capital formation}$$

Rate of gross domestic capital formation – Rate of gross domestic savings =

Rate of net foreign capital inflow = Rate of net domestic capital formation – Rate of net domestic savings.

Check Your Progress 3

1) Starting from Personal Consumption Expenditure arrive at GDP_{fc}.

.....

.....

.....

2) State the relationship between NNDY and PDY.

.....

.....

.....

3) What are the factors, which help in the distinction between

i) GDP and NDP

ii) GNP and GDP

iii) NY and NNDY

iv) GDP_{mp} and GDP_{fc}

v) NNP_{mp} and NNDY

vi) Personal Y and PDY

.....

.....

.....

.....

.....

13.5 LET US SUM UP

In this unit, we have given you an idea of the concept of circular flows and how national income of an economy can be derived by studying the working of circular flows.

The concept of circular flow relates to the flow of real transactions or money

transactions from one group of transactors to another. Flow of real transactions give us real flows and the flow of money from one group of transactors to another give us money flows.

Real or money flows can be studied between enterprises and households. Their study can be extended in an economy, which has enterprises, households and capital sector as the transactors. Similarly, we can extend this study further by incorporating Government sector and the rest of the world sector. Once enterprises, households, capital sector, government sector and the rest of the world sector are introduced we study the flows of an open economy.

National income can be studied in its three ways viz., as flow of goods and services or as flow of factor incomes or as flow of final expenditures: National income looked at in either of the three ways gives us the same total. Finally, in the last section of this unit we have gone into the discussion of national income and various related concepts and also introduced the inter-relationships among the related concepts.

The main concepts introduced are GNP_{mp}, NNP_{mp}, NNP_{fc}, NNDY, income from domestic product accruing to private sector, private income, personal income, personal disposable income, personal savings, rate of gross and net domestic capital formation, rate of gross and net domestic savings and rate of net foreign capital inflow. An attempt is also made to discuss the relationship among these concepts.

13.6 KEY WORDS

Capital Good : It is a good, which helps, in further production of consumer goods or intermediate goods or machines.

Capital Sector : It is a group of transactors, which mop up savings of various sectors and use the saving for creation of capital goods (or investment).

Change in Inventories : Inventories are stocks of finished goods/semi-finished goods/intermediate goods. Change in inventories is total inventories at the end of the year minus total inventories at the beginning of the year for an economy.

Circular Flow : It is a flow of goods or services or money from one (set of) transactor to another (set).

Depreciation : It is the loss in the value of capital asset because of normal wear and tear and expected obsolescence.

Compensation of employees : Remuneration given by enterprises to employees for rendering labour services.

Departmental Enterprises : Part of government sector. It consists of those enterprises, which run as a part of government departments producing goods/services like railways, post and telegraph etc.

Direct Personal Taxes : These are the taxes imposed on households in the form of income tax or wealth tax. Those on whom they are imposed pay them.

Economic Agents : They are groups of transactors, which indulge in

economic activities like production/income generation/ addition to capital stock. Economic agents can be classified into producers, households, capital sector, rest of the world and government.

Economic Transactions : These are the transactions consisting of production, income generation and addition to capital stock.

Enterprises : They are the transactors, which employ factors of production to generate a flow of goods and services in the economy.

Factor Services : These are the services rendered by factors of production like land, labour, capital and enterprise.

Final Consumption Expenditure : This is an expenditure incurred by households, enterprises and rest of the world to purchase final consumer goods, capital goods and net exports respectively.

Factor Cost : It is the total cost incurred to employ factors of production to give rise to a flow of goods and services in an economy. It is equal to value of market price minus net indirect taxes.

GDP : It is the value of goods and services produced in an economy over a year, without duplication but gross of depreciation. It is the goods and services produced within the domestic territory of an economy,

GNP : It is the value of goods and services produced in an economy over a year, without duplication but gross of depreciation. It is the goods and services produced by the normal residents of an economy.

Government Final Consumption Expenditure : It is the expenditure incurred by government on the purchase of intermediate goods plus compensation of government employees. This expenditure is incurred to meet the collective consumption of the economy.

General Government : It is the sector, which produces goods and services that are not sold at a price. Such goods are meant to meet collective consumption requirements of an economy. The expenses of these goods are met by tax and non-tax revenue of the government.

Households : It is a sector, which supplies factor services to producers or enterprises. The factor incomes received by households are used to meet their final consumption requirements and the balance is used for savings, which are passed on to the capital sector.

Income from Domestic Product accruing to private sector : It is the factor income enjoyed by households and private sector over a year in an economy.

Intermediate Consumption : It is the purchase of raw materials and services by a firm from other firms to produce goods or services.

Investment : It is the creation of capital goods in an economy over a year. It can be for replacement of worn out capital or for addition to total capital stock of an economy.

Macro Economics : It is that branch of Economics, which deals with aggregates of an economy like employment, inflation, balance of payment and national income.

Money Transactions : These are the transactions between one transactor and another or between one group of transactors and another in terms of money to money without being backed by real transactions.

Money Flows : These are the flows, which are normally a consequence of real flows between transactors.

Market Price : It is the price at which a commodity or service is actually purchased by a households or a firm.

Mixed Income of Self-employed : It is the factor income generated by unincorporated enterprises where it is not possible to distinguish between compensation of employees and operating surplus.

Net Domestic Capital Formation : It is that part of total production of capital goods and inventories which are meant to add to total capital stock of an economy over a year in an economy.

Net Current Transfers from the Rest of the World : It is the difference between unrequited transfers from the rest of the world, over a year, and such transfers from the economy to the rest of the world.

Net Exports : It is the difference between total value of exports and imports over a year.

Net factor Income from abroad : It is the difference between factor incomes earned by the normal residents of an economy stationed abroad temporarily and the factor incomes earned by normal residents of the rest of the world stationed in the economy temporarily.

NDP : It is the value of goods and services produced in an economy, over a year, without duplication, net of depreciation. This concept is related to the concept of domestic territory.

National Income : It is the same aggregate as NNP_{fc}.

Net Indirect Taxes : It is the difference between indirect taxes and subsidies.

Net National Disposable Income : It is the total income at the disposal of a nation by way of factor income as well as transfer incomes from the rest of the world. It is identical to NP at market price plus net current transfers form abroad.

Non-departmental Enterprises : These are the government enterprises for which autonomous corporations are set up. The goods or

services produced by these enterprises are sold for a price. They are the profit making enterprises set up in the public sector.

Normal Residents : They are the households or institutions, which have their centre of interest in the economy but some of which may temporarily be stationed aboard.

Open Economy : It is an economy, which has economic transactions with the rest of the world.

Operating Surplus : It is the factor income generated by ownership and management of property. It consists of rent, interest and profits.

Personal Income : It is the factor income and transfer incomes enjoyed by the households of an economy over a year.

Private Income : It is the factor income and transfer incomes enjoyed by households and private sector of an economy over a year.

Personal Disposable Income : It is the factor income and transfer incomes left at the disposal of the households after paying direct personal taxes and miscellaneous receipts of government administrative departments from personal income.

Private Final Consumption Expenditure : It is the expenditure incurred by the households of an economy, over a year, on the purchase of goods and services meant for final consumption.

Quid-Pro-quo : Whenever you get something from a transactor in return for something it is called quid-pro-quo.

Real Transactions : Those transactions, which are related to exchange of goods or services between two transactors or two groups of transactors.

Real Flows : These are the flows of goods or services from one set of transactors to another.

Rest of the World Sector : This sector deals with economic transactions of an economy with the rest of the world.

Replacement Investment : It is that part of currently produced capital goods, which are meant to replace the capital stock arising out of normal wear and tear, and expected obsolescence.

Rate of Gross Domestic Capital Formation : It is defined as gross domestic capital formation divided by GDP_{mp} multiplied by 100.

Rate of Net Domestic Capital Formation : It is net domestic capital formation divided by NDP at mp multiplied by 100.

Rate of Gross Domestic Savings : It is equal to gross domestic savings divided by GDP at mp multiplied by 100.

Rate of Net Domestic Savings : It is equal to net domestic savings divided by NDP at mp multiplied by 100.

Rate of Net Foreign Capital Inflow : It is the difference between rate of gross domestic capital formation and rate of gross domestic savings.

13.7 SOME USEFUL BOOKS

Studenski, Paul (1972), *Income of Nations*, (Chap 11-12), Macmillan, New Delhi

Beckerman, Wilfred (1980), *National Income Analysis*, (Chap 1-3), Wiedenfeld and Nicolson: London

C.S.O., *National Accounts Statistics* (relevant pages), Ministry of Statistics and Programme Implementation, Govt. of India, New Delhi.

13.8 ANSWERS/ HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) Money flows are the flows, which take place between one transactor and another or between one group of transactors to another. For example producers produce goods and services and pass these on to households for which payments are made by them to producers. In return households supply factor services to producers and producers make factor payment to households. Real flows, on the other hand, are the flows of goods and services from producers to households and household supply factor services to producers, which constitute real flows.

2) The economic transactions can be grouped into production income generation, addition to capital stock and the rest of the world transactions. The economic transactors can be divided into enterprises, households, government, capital sector and the rest of the world sector.

3) When capital sector is introduced along with enterprise and household sectors the circular flow gets complicated. The complication arises from the fact that they buy consumer goods and services produced by enterprises need not use the whole of factor income, which households get from enterprises, and the leakage takes place in the form of households savings, which get transferred to capital sector.

Similarly, the enterprises may also save a part of their revenue, which they get by selling consumer goods and services in the form of depreciation fund and undistributed profits.

These savings of households and enterprises which capital sector accumulates are lent to enterprises for investment or creation of capital goods (fixed capital goods or change in inventories). Thus, it is the savings and the investment, which need to be introduced in the circular flow besides other flows, which we have when only enterprises and households operate.

Check Your Progress 2

1) Production flow is the goods and services produced in an economy over a year. Production of goods and services requires factor inputs, which are supplied by households.

Employment of factors of production leads to generation of factor incomes or income flows. The incomes received by households for supplying factor inputs to producers can be used to either buy consumer goods and services produced by enterprises or save. Whatever is saved is, in turn, used for the purposes of addition to capital stock (or investment of an economy).

Thus, production flow leads to income flow and income flow leads to flow of expenditures or flow of addition to capital stock. The process continues once again by consumption expenditure and capital expenditures going to enterprises and once again the process get initiated.

2) i) Currently produced goods and services can be divided into consumer goods and services and capital or investment goods. Goods or services produced can also be of the nature of intermediate goods but they are not included in other two categories because otherwise there would be duplication in the computation of GDP of an economy.

ii) Factor incomes generated over a year can be classified into compensation of employees, operating surplus and mixed income of self-employed. Operating surplus is constituted of rent, interest and profits. Mixed income of self-employed is that category of factor incomes where it is not possible to distinguish between compensation of employees and operating surplus. Net factor income from abroad has also to be added to domestically generated factor incomes to arrive at national income of an economy.

iii) Expenditures currently generated can be divided into (a) Private final consumption expenditure; (b) Government final consumption expenditure; (c) Gross domestic capital formation; (d) Net exports to RoW.

From the sum of (a), (b), (c) and (d) we have to deduct net indirect taxes, depreciation and add with this net factor income from abroad to arrive at national income of an economy.

3) Production, income and expenditure are three ways of measuring national income of an economy.

National income measured by any of the three methods gives us the same total.

In the case of India we use combination of three methods to measure national income of India. For agriculture etc., we employ production method, for services sector income method and for construction sectors it is the expenditure method, which is employed.

For construction sector, expenditure method is employed because expenditure data are more readily available.

Check Your Progress 3

1) Personal consumption expenditure + personal savings = Personal disposable income.

Personal disposable income + direct personal taxes + miscellaneous receipts of government administration departments = Personal Income.

Personal income + undistributed profits of private corporations + corporation tax = private income.

Private income – net factor income from abroad – net current transfers from abroad – national debt interest – transfers from government = Income from domestic production accruing to private sector.

Income from domestic product accruing to private sector + savings of non-departmental enterprises + income from domestic product accruing to government administrative departments = NDP at fc.

NDP at fc + Depreciation = GDP at fc.

2) NNDY – net indirect taxes – income from domestic product accruing to government administrative departments – savings of non-departmental enterprises + national debt interest + current transfers from government – undistributed profits – corporation taxes – direct personal taxes – miscellaneous receipts of government administrative departments = Personal disposal income.

3) i) Depreciation

   ii) net factor income from abroad

   iii) net indirect taxes + net current transfers from abroad

   iv) net indirect taxes

   v) net current transfers from abroad

   vi) direct or personal taxes + miscellaneous receipts of government administrative departments.

(Rs. crore)

Total value of output	545981
Intermediate cost	417244
Gross value added, including bank charges	128737
Imputed bank charges	-11901
Gross value added	116836
Consumption of fixed capital	-24184
Net value added	92652

Source: NAS, 1999

14.3.4 Method Employed in Construction Sector

In this sector, expenditure-cum-commodity flow method is used. Construction activity is classified into two groups, namely, (a) kutcha construction, and (b) pucca construction.

Expenditure method is used in kutcha construction, which is labour-intensive. In it hardly any input of construction material having economic market value is used. It uses such freely available inputs as leaves, reeds and mud. Such construction is usually seen in rural areas where poor farmers and labourers construct their huts, small houses and cattle sheds mostly with their own labour and freely available inputs. Data for this sector are obtained through surveys by NSSO.

The commodity flow approach is adopted in pucca construction. Pucca construction is mostly undertaken in urban areas and involves costly inputs and modern techniques of construction.

The inputs used in pucca construction are bricks, cement, steel, wood, electric fittings and sanitary fittings. Estimated values for these items from are collected from ASI, government agencies and also directly from dealers. After deducting the value of intermediate consumption and consumption of fixed capital from the value of output, net value added in each group is obtained. In India, the value added from construction during 1996-97, at constant (1993-94) prices, is as follows:

(Rs. crore)

Value of output	123544
(a) New construction	100399
(b) Repairs and maintenance	23145
Value of intermediate consumption	76162
Gross value added	47382
Consumption of fixed capital	-2308
Net value added	45074

Source: NAS, 1999

14.3.5 Difficulties in Measurement of National Income of India

The difficulties faced in measuring national income of India can be divided into two categories:

1) conceptual difficulties, and

2) statistical difficulties.

Conceptual difficulties. All the countries, whether developed or developing, face difficulties relating to the concepts used in national income accounting. Even economists differ on what constitutes national income, i.e., what to include in national income. Examples of such difficulties are:

- absence of an agreed definition of national income
- difference between final goods and intermediate goods
- transfer payments
- services rendered without remuneration etc.
- consumers' durable goods etc.

Statistical difficulties

Inadequate and unreliable data

Data available are neither sufficient nor correct and sufficiently detailed. For example, in India, it is possible to get information, though inadequate, from producing units in organised sector. But it is difficult to get data from agricultural and related activities like household, crafts and indigenous credit (functioning of village money lenders). Moreover production is not only unorganised but scattered also.

Non-monetised sector

A significant part of the product, especially product of agriculture and cottage industries in rural India is bartered. Valuation of such goods, which have been exchanged through barter system, becomes very difficult since they do not reach the market.

Production of goods for self-consumption

The small farmers who constitute a sizeable number in India produce goods mainly for their own use. The value of such goods cannot be computed because they do not come to market for sale. Hence a good deal of guesswork in such cases become inevitable.

Illiteracy and ignorance

Majority of the people in India are illiterate, uneducated and ignorant. They do not maintain account of their income and expenditure. These people do not even know, much less give data, about their income and expenditure. They are ignorant of cost accounting, which is being progressively used by most of the developed countries.

Lack of occupational classification

In India most of the people earn their living from more than one occupation. For example, Indian farmers work primarily in agriculture. But when free, they also work in cottage and small industries. Should they be treated agricultural workers or industrial workers? When people have different sources of income, it becomes difficult to know the main source and consequently a large part of income gets excluded from national income.

Valuation of a new good at constant prices

When a new commodity is produced for the first time, it is easy to know its current price but difficult to get its constant price. For example, in India it is difficult to work out the value of a colour T.V. at constant prices of 1970-71 because it was not being produced in that year.

Consumption of fixed capital

By deducting depreciation cost from gross national product (GNP), we get the value of net national product. But it is quite difficult to measure the correct value of consumption of fixed capital (e.g., machines, buildings, etc.) during the year. So some sort of arbitrariness gets involved while estimating depreciated amounts.

Check Your Progress 2

1) Identify the sectors in which production method is employed to find out their contribution to National Income of India.

.....

.....

.....

.....

2) Enumerate the steps involved in finding out net value added of agricultural sector in India.

.....

.....

.....

.....

3) What are the main conceptual difficulties faced in computation of National Income of India?

.....

.....

.....

.....

14.4 LET US SUM UP

In this unit we have given you an account of three methods to measure national income of an economy. These three methods, namely, production (or value added), income and expenditure have been explained in detail along with various points which should be kept in mind before employing each of them.

It has also been shown how reconciliation of the three measures of national income is achieved.

A detailed discussion of the methods to measure national income of India is also attempted after giving a historical account of national income measurement. The sub-sectors of Indian economy as employed by the Central Statistical Organization of India are presented. The details of the methods employed in agriculture, registered manufacturing sector and construction sector are given by way of illustration.

Finally, difficulties, both conceptual and statistical, faced in measurement of national income of India are explained.

14.5 KEY WORDS

Compensation of Employees : It is the factor income generated in an economy in the process of production by rendering labour services. It is also referred to as income from work that includes wages and salaries, employers' contribution to social security, commissions and real wages.

Change in Stocks : It is the difference between total stocks or inventories (including unsold finished products, semi-finished products and raw materials) at the end of the year and total stocks at the beginning of the year.

C.S.O. : Stands for Central Statistical Organization, which is responsible for the publication of 'National Accounts Statistics' every year providing estimates of national income and various other macro-economic aggregates.

Double Counting : When the value of output of a product is included along with the other products in which it has entered as raw material, double counting occurs. It should be avoided in estimation of GDP of an economy.

Depreciation Provision : It is the provision of funds in an economy to account for normal wear and tear and expected obsolescence of national capital stock.

Domestic Territory : It includes geographical or political boundary of an economy besides ships in the international waters and embassies or consulates in foreign countries.

Direct Taxes : The taxes which are imposed on persons or institutions and they are to be paid by them alone. In other words, these are the taxes the incidence of which cannot be transferred to others.

Death Duties : The taxes, which are paid by those who inherit the property from the parents or relations or friends etc.

Durable Goods : These are the goods, which can be used again and again like tables, machines etc. To be more precise they are referred to as durable-use goods.

Economic Growth : It is the growth of real per capita income in an economy.

Economic Development : It is economic growth of an economy coupled with all those factors, which sustain economic growth over a sufficiently long period of time.

Economic Welfare : It is that part of total welfare of an economy, which

is associated with production and consumption of goods and services.

Expenditure Method : It is sum of all expenditures by various economic agents on the goods and services produced by an economy.

Finished Product : It is that part of total production of an economy which is meant for either final consumption on the part of households, government exports, or for capital accumulation purposes.

Factor Incomes : These are incomes distributed to suppliers of factor services (labour, land, capital and enterprise) for work done over a year in an economy.

Gross Value Added at mp : It is the market value of output of an enterprise minus intermediate consumption.

Mixed Income of Self-employed : It is factor income (compensation of employees plus operating surplus) generated by self-employed.

Gross Value Added at factor cost : It is gross value added at market price minus net indirect taxes paid by a producing unit.

Income Method : It is sum of factor incomes currently generated, which is used along with net factor income from abroad to arrive at national income of an economy.

Intermediate Inputs : Inputs employed by a firm in the process of production, which is bought from other firms or producing units.

Industrial Sectors : Sectors, which are identified for the purposes of measuring national income of an economy.

Imputed Rent : It is the value of rent worked out for owner-occupied dwellings.

Intermediate Consumption : It is the value of inputs used by a producing unit.

Net Factor Income from abroad : The factor income (compensation of employees and property and entrepreneurial income) received from abroad by normal residents of an economy minus such factor incomes paid to normal residents of foreign countries.

Net Value Added at Factor Cost : It is gross value added at factor cost minus depreciation allowance.

Net Value Added at Market Price : It is net value added at factor cost plus net indirect taxes.

Non-Durable Goods : These are the goods, which can be used only once.

Net Exports : It is value of exports minus value of imports of an economy.

Output : It can be physical units of a commodity produced

by a producing unit or it can be physical units of a commodity multiplied by its price over a year.

Owner-Occupied Dwellings: These are the houses, which are occupied by their owners.

Own-Account Production : It is the production of goods or services by self-employed households or institutions.

Operating Surplus : It is the factor incomes consisting of rent, interest and profits arising out of ownership and management of capital.

Primary Sector : This is a sector, which consists of agriculture, forestry and logging, fishing, mining and quarrying.

Secondary Sector : This is a sector, which consists of manufacturing (registered and unregistered) and electricity, gas and water supply.

Tertiary Sector : This is a sector which consists of trade, hotels and restaurants, transport, storage and communication, financing, insurance, real estate and business services and community, social and personal services.

Transfer Incomes : These are the incomes, which do not arise on account of factor services; they are incomes transferred from one sector to another or transferred from one unit to another within a sector. They are not included in national income of an economy.

14.6 SOME USEFUL BOOKS

C.S.O, *National Accounts Statistics*, (Latest), Ministry of Planning, Government of India, New Delhi

Hicks J.R., M. Mukherjee and S.K. Ghosh, 1984, *The Framework of the Indian Economy - An Introduction to Economics*, OUP, Delhi (Chapters 11, 12, 13)

C.S.O, *National Accounts Statistics* - Sources and Methods, Ministry of Planning, Government of India, New Delhi, April, 1989.

14.7 ANSWERS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) a) Gross Investment - depreciation (or consumption of fixed capital) = Net Investment.

   b) Net Factor Income from abroad = Net compensation of employees from abroad + net property and entrepreneurial income from abroad.

   Net exports = Value of exports - value of imports.

   c) Gross national product - depreciation = Net national product.

   d) Gross domestic product - depreciation = Net domestic product

2) Various constituents of net factor income from abroad are:

   1) compensation of employees from abroad.

   2) property and entrepreneurial income from abroad.

   3) net retained earnings of resident companies abroad. Net factor income from abroad can be negative when factor income from abroad is less than factor income paid to the rest of the world.

3) a) GDP at MP = Sum of gross value added at market price of all the producing units of an economy.

   b) GDP at mp = Compensation of employees + operating surplus
   + mixed income of self-employed + net indirect taxes
   + net factor income from abroad + depreciation.

   c) GDP at mp = final consumption expenditure of households
   + final consumption expenditure of government
   + gross domestic capital formation + exports - imports.

Check Your Progress 2

1) Agriculture and allied activities of fishing, mining and quarrying and logging.

2) a) Identify the crops.
   b) For each crop find out area under cultivation.
   (c) Multiply (b) by yield per hectare.
   d) Multiply (c) by average price.
   e) From (d) subtract the intermediate cost of agriculture and allied activities.
   f) With (e) add the value of operation of irrigation systems.
   g) From (f) deduct consumption of fixed capital.
   h) The figure arrived at (g) will be the net value added of agricultural sectors.

3) Conceptual difficulties relate to :
   a) absence of an agreed definition of national income;
   b) difference between final goods and intermediate goods;
   (c) transfer payments;
   d) services rendered without remuneration etc.;
   e) consumer durables etc.

   Statistical difficulties consist of :
   a) inadequate and unreliable data,
   b) non-monetised sector,
   c) production for self-consumption,
   d) illiteracy and ignorance,
   e) lack of occupational classification,
   f) value of a new good at constant prices,
   g) consumption of fixed capital estimation.

ignou
THE PEOPLE'S
UNIVERSITY

Indira Gandhi
National Open University
School of Social Sciences

EEC-11
Fundamentals of Economics

Block

6

MACROECONOMICS AGGREGATES

Expert Committee

Mr. Kalyanjit RoyChaudhary
St. Stephens College
University of Delhi
Delhi

Mr. B.S. Bagla
PGDAV College
University of Delhi
Delhi

Mr. R.S. Malhan
St. Stephens College
University of Delhi
Delhi

Mr. R.S. Bharadwaj
Shivaji College
University of Delhi
Delhi

Dr. Gopinath Pradhan
IGNOU

Dr. Narayan Prasad
IGNOU

Dr. Madhu Bala
IGNOU

Dr. Kaustuva Barik
IGNOU

Mr. Saugato Sen
IGNOU

Prof. S.K. Singh (retd.)
IGNOU

Course Editor	Course Coordinator	Block Coordinator
Mr. Bhawani Sankar Bagla	Dr. Gopinath Pradhan	Dr. Gopinath Pradhan

Block Preparation Team

Unit No.	Unit Writer	Unit Editor (IGNOU Faculty)
13	Sh. R.C. Malhan	Dr. G. Pradhan
14	Sh. R.C. Malhan	Dr. G. Pradhan

Production	Cover Design	Word Processing
Mr. Arvind Kumar	Ms. Arvinder Chawla	Mrs. Rekha Mishra
Mr. Manjit Singh		Ms. Daisy Lal

March, 2003

ISBN-81-266-0764-5

Further information on the Indira Gandhi National Open University courses may be obtained from the university's office at Maidan Garhi, New Delhi-110 068.

Printed and published on behalf of the Indira Gandhi National Open University, New Delhi, by Professor Kapil Kumar, Director, School of Social Sciences.

Laser typeset by : HD Computer Craft, EA1/75, Main Market Inderpuri, New Delhi-110012. Ph: 25831437 ; 3108 1117

Printed at :

BLOCK 6 MACROECONOMICS AGGREGATES

Introduction

After going through blocks, 1 to 5, which deal with micro-economic insights of economic theory, the learner is exposed to macro economic concepts through Block-6. Basically, an attempt is made to show that at the aggregate level, the income of the nation can be seen to have generated by interrelationships between various economic activities. **Unit 13** depicts the national income as a result of circular flow of aggregate income and expenditure. The remaining unit of the block, i.e., **Unit 14** accounts for the method followed for measuring the national income.

UNIT 14 NATIONAL INCOME MEASUREMENT

Structure

14.0 OBJECTIVES

After going through this unit you will be able to understand

- three methods of measuring national income;
- constituents of national income in its three phases, viz., Production, Income and Expenditure;
- classification of Indian economy into various sectors for estimation of national income;
- methods of estimation of national income used for each of the above sectors; and finally
- problems and difficulties, which arise in course of estimation of national income in India.

14.1 INTRODUCTION

National income is an important macro-economic aggregate. With certain qualifications, it can be taken as an indicator of economic growth, economic development and economic welfare. In view of its importance, a wrong or a biased estimation of national income can have far reaching consequences. Therefore, we will devote a major part of this unit discussing the methods to measure national income of an economy.

14.2 NATIONAL INCOME

National income can be defined as the monetary aggregate of the current achievements of an economy, without any double counting. It is estimated, usually for one year. Therefore, it is net monetary value of all the goods and services produced by the constituents of a national economy over a period of one year.

The concepts of flow of production, generation of income and expenditure have

already been introduced in Unit 13. It has also been shown how the flow of production gives rise to generation of income, which in turn, leads to flow of expenditure. Flow of expenditure again gives rise to flow of production and thus the processes of production, income generation and expenditure go on period after period, showing thereby that these processes take place continuously in an economy. The measurement of production, income and expenditure flows with respect to normal residents of an economy gives us three methods of measuring national income of an economy.

14.2.1 Methods to Measure National Income

There are three methods to measure national income of an economy. These are:

1) production method or value added method,

2) income method, and

3) expenditure method.

Each of these methods corresponds to a flow taking place in the economy. These three methods are, in fact, three ways of looking at the same variable, national income. The statistical data and tools to measure national income by each of the three methods may be different but conceptually each of these will give the same result. If the different methods do not give us the same estimate of national income, it will be primarily because of lack of relevant statistical data required to measure national income.

14.2.2 Production Method

Basically, three steps are involved in applying the production method to compute the national income of an economy. These steps are:

1) to identify the producing enterprises and to classify them into industrial sectors according to their activities

2) to estimate net value added at factor cost of each producing enterprise within the domestic territory of an economy and to add up net value added by all the sectors to arrive at net domestic product at factor cost

3) to estimate net factor income from abroad, which has to be added to net domestic product at factor cost to arrive at net national product at factor cost/ national income of an economy.

Classification of Industrial sectors

Broadly speaking, the industrial sectors are divided into three categories:

a) Primary Sector,

b) Secondary Sector, and

c) Tertiary or Services Sector.

Primary Sector: It includes agriculture and allied activities of forestry, fishing, mining and quarrying. This sector produces commodities by exploiting natural resources like coal, iron ore and other minerals. In India the primary sector is divided into: (i) agriculture, (ii) forestry and logging, (iii) fishing, and (iv) mining and quarrying.

Secondary Sector: It consists of the manufacturing sector. In India, the secondary sector is divided into: (i) registered manufacturing, (ii) unregistered manufacturing, (iii) construction, (iv) electricity, gas and water supply.

Tertiary Sector: It consists of the services. In India this sector comprises: (i) railways, (ii) transport by other means and storage, (iii) communications, (iv) trade, hotels and restaurants, (v) banking and insurance, (vi) real estate, ownership of dwellings and business services, (vii) public administration and defence and (viii) other services.

Estimation of Net Value Added

After the producing sectors of an economy are identified, the next step is to find out net value added of each of these sectors. The term value added refers to addition of value by a producing unit to raw materials and services (known as intermediate inputs) used in production. What a producer produces is termed as the output. Value added is the difference between the value of output and the cost of intermediate inputs.

Let us illustrate the concept of value added with the help of an example. Suppose there is a producing unit, X, which sold goods worth Rs. 48, 000 in a year and added to stocks of goods worth Rs. 2,000, such that the value of output of X is Rs. 50,000 (Rs. 48, 000 + Rs. 2000 = Rs. 50, 000). Suppose X bought raw materials and services worth Rs. 43, 000 from other producing units of the economy. In this case, value added of X is the value of output less value of raw materials and service used up in the process of production. Thus Rs. 50, 000 - Rs. 7, 000 = Rs. 43, 000 is value added by the producing unit X over the year in question. Gross domestic product of an economy is the sum of value added by all the producing units within the domestic territory of an economy. This concept of value added can be shown with the help of another example.

Suppose an economy has only three producing units, viz., (a) an enterprise engaged in cutting trees, (b) another engaged in converting trees cut by the first into logs and (c) a third engaged in converting logs, produced by the second enterprise into tables. Further, suppose that the first enterprise engaged in cutting trees does not require any raw materials for undertaking its activity. The value of output and value added of these three enterprises over a year are shown Table 14.1

Table 14.1: Value Added of Enterprises

(Rs. 000)

Enterprise	Product Produced	Value of output	Value of intermediate Consumption	Value added
1	Wood	50	nil	50
2	Logs	80	50	30
3	Tables	150	80	70
	Total	280	130	150

It is clear from Table 14.1 that value of wood produced (or trees cut) is Rs. 50, 000 and the same wood when converted into logs by the second enterprise is worth Rs. 80, 000. These logs further processed by the third enterprise into tables which worth Rs. 150, 000. The total value of output by all the three enterprises is equal

to Rs. 2, 80,000 (i.e., Rs. 50, 000 + Rs. 80, 000 + Rs. 150, 000). This cannot be called gross domestic product of an economy because of counting, since wood is included in logs and logs in tables. There are two ways of avoiding double counting: (a) Take only the final products and do not consider the production of enterprises, which produce raw materials or semi-finished products. But this method is difficult to apply since it is not easy to decide between a raw material and a finished product. For example, wood, if used to make logs, is a raw material but would be a finished product, if used as a fuel by household. Therefore, this method is not employed for estimating GDP of an economy. In the example given above, the value of production of the third enterprise, i.e., Rs. 1, 50,000 is GDP of an economy; (b) the other way of avoiding double counting is to consider value added by each producing unit. Value added makes the output of a producing unit free of the value of intermediate consumption.

Value added by each of the three producing units in Table 14.1 is Rs. 50, 000, Rs. 30, 000 and Rs. 70, 000 respectively, which add up to Rs. 1, 50,000. This is the GDP of the economy estimated by first and second methods. It may be worthwhile to note that the second method is easier to apply than that of the first.

It is also important to note that :

i) sum of gross value added at market price of all producing units gives us gross domestic product at market price;

ii) sum of gross value added at factor cost of all producing units gives us gross domestic product at factor cost;

iii) net domestic product at factor cost is estimated by adding net value added at factor cost of all the producing units of an economy; and finally,

iv) net domestic product at market price is computed by summing up net value added at market price of all the producing units of an economy.

Net Factor Income from Abroad

The concept of net factor income from abroad has to be added to net domestic product at factor cost for getting the national income. It consists of

1) net compensation of employees,

2) net income from property and entrepreneurship, and

3) net retained earnings of resident companies abroad.

Net Compensation of Employees from Abroad

Net compensation of employees receivable from abroad is equal to the difference between compensation of employees received by resident employees who are living or employed abroad temporarily and compensation of foreign nationals working temporarily in the domestic economy. The clause temporary resident applies to those employees who stay abroad for less than one year. In case they stay for one year or more in a foreign country they would be treated as normal residents of that country and their income would be a part of the national income of the employer country. In such a situation, whatever remittances they send to their country would be treated as current transfers from abroad and will not form a part of the national income. Net compensation of employees, as it is defined, can be a positive or a negative value.

Net Income from Property and Entrepreneurship from Abroad

Net income from property and Entrepreneurship from abroad is the difference between the income received by way of interest, rent, dividend and profit by the resident producers of a country and payments of similar type made to the rest of the world. This also includes net interest received by the government on foreign loans.

Net Retained Earnings of Resident Companies Abroad

Retained earning refers to the undistributed profit of the companies. Resident companies abroad (i.e., companies belonging to one country and working in the domestic territory of some other country) retain a part of their profits for further investment abroad. Likewise, foreign companies and their branches retain a part of their profits in the countries of their operation. The difference between retained earning of the foreign companies located in a country and retained earning of resident companies located abroad is equal to net retained earnings from abroad.

Thus, net factor income from abroad is equal to net compensation of employees from abroad plus net property and entrepreneurship income from abroad plus net retained earning of resident companies abroad.

Let us state the definition of net factor income from abroad, given in a formal way, by Central Statistical Organization (CSO) of India. It is stated as, "Income attributable to factor services rendered by the normal residents of the country to the rest of the world less factor services rendered to them by the rest of the world. It also includes retained earnings of foreign-controlled rupee companies and branches of foreign companies in the domestic territory. Residents include both individuals and institutions. Tourists or commercial travelers of a given country traveling abroad are treated as residents of their home countries. The official diplomatic and consular representatives of a given country including members of official missions and members of the armed forces stationed abroad are to be considered extra territorial by the country in which they are located and as residents of the given country. The factor incomes generated by such residents are domestic product of the resident country. Factor incomes of locally recruited staff of foreign diplomatic military establishments are included in factor incomes from abroad."

The concept of net factor income from abroad is employed to arrive at gross national product at market price, gross national product at factor cost, net national product at market price and net national product at factor cost (national income) of an economy. Thus, we have the following concepts:

a) Gross domestic product at market price + net factor income from abroad = gross national product at market price.

b) Gross domestic product at factor cost + net factor income from abroad = gross national product at factor cost.

c) Net domestic product at market price + net factor income from abroad = net national product at market price.

d) Net domestic product at factor cost + net factor income from abroad = net national product at factor cost (national income)

Normally, domestic product (gross or net, at market price or factor cost) is first computed and adjustments are made by including net factor income from abroad to arrive at national product (gross or net at market price or factor cost).

In estimating national income of an economy by the production method, the following points should be kept in mind.

1) Production for self-consumption has to be included in national income. In order to find out the value of production for self-consumption, the physical production has to be multiplied by market price to get imputed value of production for consumption.

2) Imputed rent in case of owner-occupied houses has to be included in national income. Imputation of rent has to be done such that it will be equal to the services rendered by owner-occupied dwellings.

3) Own-account production by the government, private enterprises and households has to be found out and added to other items for calculating national income of an economy.

4) Sale and purchase of second-hand goods by sectors do not constitute a part of current production and, therefore, are not included in national income in the year in which their sale or purchase takes place. But if sale and purchase of second hand goods take place through a broker, then the services rendered by these brokers are a part of national income of an economy. The services rendered by these brokers are expressed as equal to the commission and brokerage earned by them.

14.2.3 Income Method

Net value added at factor cost of a producing unit is identical to the sum of factor incomes - compensation of employees and operating surplus or mixed income of the self-employed entrepreneurs. So, we can sum up compensation of employees and operating surplus or mixed income of the self-employed to estimate national increase. If factor incomes generated by all the producing units within the domestic territory of an economy are added up, net domestic product at factor cost is arrived at. Further, if net factor income from abroad is added to net domestic product at factor cost of an economy, net national product at factor cost or national income of an economy can be estimated.

Given net national product at factor cost if the depreciation provision is added to it, gross national product at factor cost is arrived at. If net indirect taxes are added to gross national product at factor cost, gross national product at market price can be calculated.

The main steps involved in estimating national income by the income method are:

1) to identify the producing enterprises, which use services of the factors of production

2) to classify various types of factor payments

3) to estimate various components of factor payments

4) to estimate net factor income from abroad, which has to be added to net domestic product at factor cost to arrive at net national product at factor cost or national income of an economy.

The classification of producing units that is adopted by the production method of estimating national income can be used for the income method also.

The factor payments are generally classified into the following categories:

a) Compensation of employees

b) Rent

c) Interest

d) Profits

e) Mixed income of the self-employed

Moreover, factor payments can be classified into: (a) compensation of employees, (b) operating surplus, (c) mixed income of the self-employed. There are a few points to be kept in mind while estimating national income by income method.

1) A distinction has to be made between factor and income transfer income. While factor incomes are earned by factors of production, transfer incomes are enjoyed by various economic agents without supplying factor services. It is only factor incomes that constitute national income. Accordingly, transfer incomes are excluded from national income of an economy.

2) The services of owner-occupied dwellings are equal to imputed rent of the dwelling. Imputed rent adjusted for maintenance expenditure of dwellings is included in national income by production method.

3) Income earned by the act of smuggling or gambling as well as windfall gains like lotteries are not included in the estimation of national income.

4) National Income of an economy includes direct taxes like income tax and corporate tax. It may be useful to remember that compensation of employees includes income tax to be paid by them and are included in national income before deduction of corporate tax. Death duties, gift tax, wealth tax, etc., are supposed to be paid from the wealth or past savings of those persons who pay these taxes and not out of current income. Therefore, such taxes are not included in the estimation of national income.

5) Sale and purchase of second-hand goods are not included in national income of an economy. The sale proceeds of second-hand goods received by a person do not relate to any service rendered and, therefore, do not constitute a part of national income.

14.2.4 Expenditure Method

Income generated in the process of production is received by factors of production. Such income can be divided into two parts viz., (a) income from work and (b) income from ownership of capital and entrepreneurship. Incomes from work are enjoyed by the workers while those of ownership of entrepreneurship are enjoyed by their owners. The income earned by factors of production is either saving. Savings generated, in turn, are used for adding to the capital stock or what is called investment. If the final consumption and gross investment expenditure of all economic agents including the rest of the world are added up, this gives us the gross domestic product at market price for an economy. From the GDP at market price, we deduct depreciation provision and net indirect taxes to get net domestic product at factor cost. Add net factor income from abroad to net domestic product at factor cost to get net national product at factor cost (or national income) of an economy. Various components of final expenditures constituting gross domestic product at market price are:

1) private final consumption expenditure,

2) government's final consumption expenditure,

3) gross domestic fixed capital formation,

4) change in stock, and

5) net export of goods and services.

Private Final Consumption Expenditure

Private final consumption expenditure is defined as the expenditure on current account of resident and non-resident households in the domestic market and on profit-making bodies serving households. The expenditure, here, relates to outlays on new durable as well as non-durable goods (except land) and on services net of sales (sales less purchases) of second-hand goods, scrap and wastes. This definition is based on the concept of expenditure within the domestic territory of the country and is not in line with the concept of a national product. It is important to keep in mind that it is not possible to take account of the direct purchases made by the resident households from abroad and deduct the purchases of non-resident households in the domestic market to get the final expenditure of resident households only. Therefore, the final private consumption expenditure also includes the purchases of goods from abroad or goods, which have been imported from abroad. Moreover, the figure of final private consumption expenditure includes the imputed gross rent of owner-occupied dwellings, consumption of own-account production and payment by households of wages and salaries in kind valued at cost, e.g., provision for food, shelter and clothing to the employees, wherever they exist.

We require two types of statistical data for the estimation of final private expenditure:

(a) total volume of sales in the market, and (b) retail prices at which goods and services are purchased by the households. The volume of final sales is to be multiplied by the retail prices of goods and services.

Production for self-consumption is a part of production and hence an income and is also a part of final consumption expenditure. Accordingly, the volume of production for self-consumption has to be multiplied with the prices existing in the markets near the producing unit. Similarly, the imputed rent of owner-occupied dwellings is included in production and income and also in private final consumption expenditure in the domestic market.

Government Final Consumption Expenditure

Government final consumption expenditure is defined as the current expenditure on goods and services used up in providing services of government administrative departments less the sales by them. Here, we are considering the services rendered by general government which consists of all departments, offices, organizations and other bodies, which are agencies or instruments of the Centre, state or local public authorities, financed by budgets or extra budget funds. Government enterprises, public corporations and departmental enterprises are excluded from it.

The value of government final consumption expenditure is equal to the value of the services produced (such as public health, cultural services, defence, and law and order) by the government for collective use by the public. These services are valued at their cost to the government, since they are not normally sold to its citizens. The cost of these services is the sum of value of: (a) intermediate consumption, (b) compensation of employees (wages, salaries in cash and in kind),

(c) the direct purchases of goods and services made by the government abroad for their embassies and consulates located abroad and, (d) less sale of goods and services produced by the general government. Examples of sales by general government are nominal money charged by government hospitals from individuals availing themselves of the hospital facilities or government publications sold by the government to the general public.

Gross Domestic Fixed Capital Formation

Gross fixed capital formation consists of the outlays of industries, producers of government services and producers of private non-profit services to households, on addition of new durable goods to the stocks of fixed assets less net sales of similar second-hand and scrapped goods. The outlays of government services on durable goods for military use are excluded from gross fixed capital formation. In it, outlays on the improvement of land, on the development and extension of timber tracts, plantations etc., are included, provided they take more than one year to become productive. Outlays by households on residential constructions are also included in gross fixed capital formation. Gross fixed capital formation is inclusive of the consumption of fixed capital. Net fixed capital formation is defined as gross fixed capital formation less the consumption of fixed capital.

Gross domestic fixed capital formation is the gross fixed capital formation with reference to the domestic territory of the country. It consists of acquisition of fixed assets by resident industries and the producers of government services and of private non-profit services to households. Gross fixed capital formation in India is divided into public, private and households sectors by the type of institutions. Gross fixed capital formation is divided into two groups, viz., construction and, machinery and equipment.

The expenditure on construction is calculated either by finding out the total money spent on new construction or by adding up the value of inputs used in construction. The expenditure on material inputs consists of expenditure on cement, steel, bricks, wood, fixtures and fittings. Besides, factor income payments to labour and capital are to be added with the expenditure on inputs to find out the value of construction. The value of inputs used in construction is calculated by multiplying the volume of material inputs with the prices paid by the builders at the site of construction. This way of calculating expenditure on construction is known as the commodity flow method. The expenditure on compensation of employees, rent, interest and profits is to be added to the value of inputs to work out the value of new construction.

Major alterations of the old buildings or addition to rooms of existing old buildings and work-in-progress at the site of construction and own-account production of fixed assets by all the producing sectors are also taken to be a part of the expenditure on new construction.

Similarly, final expenditure on machinery and equipment is found out by multiplying the volume of final sales and the retail prices prevailing in the market. Own-account production of machinery and equipment by producers is also to be included in the final expenditure on machinery and equipment.

Change in Stocks

Stocks consists largely of materials and supplies, work-in-progress (except in construction projects) and finished products in the possession of industries. Standing timber and crops are not included in stocks, but livestock raised for slaughter, logs

and harvested crops are. Change in stocks is the difference between markets or book values of the stocks in the beginning and at the end of the year.

Stocks are classified according to the kind of economic activity of the owner or according to type of the items concerned. Classification also provides for (a) stocks of newly produced and imported goods classified according to industries which normally produce these goods, (b) stocks of second hand goods classified according to the type of goods, and (c) stocks of scrap and wastes, classified according to the type of material.

The physical change in stocks has to be multiplied by the market prices to get the value of change in stocks.

Net Export of Goods and Services

Net export of goods and services is the difference between value of export and import of goods and services over a year. Accordingly, net export can be positive or negative, positive when exports are more than imports, and negative when reverse is the case.

In India, export of goods and services is defined as all transfers of the ownership of goods from residents of the country to non-residents and services provided by resident producers of the country to non-residents. If exports are recorded by the custom authorities, they may consist of the outward movement of merchandise across the customs frontiers of country and of other goods across the boundaries of her domestic territory, including the direct purchases in the country of extra-territorial organizations and non-resident persons. Exports are recorded f.o.b. (i.e., free on board). Imports of goods and services are all transfers of the ownership of goods from non-residents to residents and services provided by non-resident producers to residents of the country. The customs authorities record imports of goods as inward movement of merchandise across the custom frontiers of a country and of other goods across the boundaries of her domestic territory, including the direct purchases in the country of extra-territorial organizations. Imports are recorded c.i.f. (i.e., cost, insurance and freight). They also include the charges of resident producers for transport and insurance services in respect of these imports. Since the imports of merchandise into a country are valued c.i.f., the exports of services of the country should also include the charges in respect of the imports for the transport and insurance services provided by resident producers of the given country.

Gross domestic product at market price seen through expenditure method is thus equal to private final consumption expenditure plus government's final consumption expenditure plus gross fixed capital formation plus change in stocks less net exports of goods and services (exports of goods and services minus imports of goods and services). Net domestic product at market price is equal to gross product at market price minus consumption of fixed capital or depreciation provision. Accordingly, net domestic product at factor cost is equal to net domestic product at market price minus net indirect taxes (indirect taxes minus subsidies). Net national product at factor cost (or national income) is equal to net domestic product at factor cost plus net factor income from abroad.

It may seem confusing asto why imports should be deducted and exports added along with other components to arrive at gross domestic product. The reason for including exports is very easy to understand. While production is done within the domestic territory, a part of it is purchased by foreign countries. These purchases by foreigners are in addition to purchases by households, government or capital

sector. Therefore, these should be included in expenditure on domestic products. Imports of goods and services are deducted from the sum of private final consumption expenditure, government final consumption expenditure and gross domestic capital formation since all these components include a part of the expenditure on goods and services imported from abroad. This is done to make these figures free of imports. So it would be wrong to say that national product is equal to domestic product plus net exports of goods and services. The relationship between national product and domestic product arises because of net factor income from abroad. More specifically, it can be stated that national product (gross or net, at market price or factor cost) is equal to domestic product (gross or net, at market price or factor cost) plus net factor income from abroad and domestic product includes net exports of goods and services.

There are a few points to be kept in mind while estimating national income by the expenditure method.

1) Expenditure on all intermediate goods and services is to be excluded. This is done primarily to avoid double counting. A good or service is said to be intermediate if a producing unit sells it to another producing unit for resale or for further processing. On the other hand, if the good or service is sold for final consumption or for capital formation or for exports, it is a final good or service. In practice, it is not always easy to distinguish between an intermediate and a final good because the same good can be intermediate or final depending on its use.

2) All government expenditure on transfer payments is excluded from national income. Some examples of such transfer payments are unemployment benefits, old age pensions and scholarships given to students for education purposes. Those who receive these transfers are not expected to render any service in exchange.

3) All expenditures on second-hand goods are excluded since they are not from the currently produced goods and services. Similarly, expenditure on the purchase of old shares or bonds or debentures from other people or new shares, bonds or debentures from producing units or government are excluded since they are not payments for a good or service currently produced. There is only transfer of property from one person to another.

14.2.5 Reconciliation of the Three Measures of National Income

Production, income generation and expenditure are three flows, which take place in every economy. Corresponding to these, we have three methods to measure the flow of national income of an economy. In other words, national income flow can be viewed from the angle of production, income generation or expenditure. Thus, the value of national income is the same viewed from either production or income or expenditure angles. This equality holds good conceptually and is based on the assumption that complete statistical data are available for estimation of national income. Generally speaking, statistical data on expenditure are not adequately available. So there may be some discrepancy in the figures estimated by the three methods of national income. National income measured by the three methods independently helps us to crosscheck the reliability of the estimates. In actual practice, national income is rarely estimated by using only a single method production, income or expenditure alone. The economy is divided into various industrial sectors and for different sectors, different methods are employed depending on the availability and reliability of data. For commodity-producing sectors, normally product (or value added) method is employed and for services-producing sectors, income method is used. In contrast to these two, for the construction sector a combination of expenditure and commodity flow methods is employed. A reconciliation of the three methods of measuring national income is shown with the help of Table 14.2

Table 14.2 : Reconciliation of the Three Methods of Measuring National Income

Production Method		**Income Method**		**Expenditure Method**
Net value added at factor cost of primary sector		Compensation of employees		Private final consumptions Government final consumption expenditure
+		+		+
				Gross domestic fixed capital formation
				+
Net value added at factor cost of secondary sector		Operating surplus		Change in stock
+	=	+	=	+
Net value added at factor cost of tertiary sector		Mixed income of the self-employed		Net exports of goods and services
		+		Minus net indirect taxes Minus consumption of fixed capital
+				+
Net factor income from the rest of the world		Net factor income from the rest of the world		Net factor income from the rest of the world

Check Your Progress 1

1) Distinguish between the following:

a) Gross and net investment

b) Net factor income from abroad and net exports

c) Gross and net national product

d) Gross and net domestic product.

.....

.....

.....

.....

.....

2) Identify the various constituents of net factor income from abroad. Can net factor income from abroad be a negative figure?

.....

.....

.....

.....

3) What are the constituents of GDP at market price using

a) Production approach

b) Income approach

c) Expenditure approach.

.....

.....

.....

.....

.....

14.3 NATIONAL INCOME MEASUREMENT IN INDIA

In India national income is measured by employing a combination of product, income and expenditure methods. In fact for different sectors of Indian economy the choice of method depends on the availability of statistical data. For example, a sector producing tangible goods like agriculture or manufacturing, production data may be more easily available and hence production method is employed to find out their contribution to national income of India. On the other hand, for a sector like government where more services are produced, it is the income figures, which are more easily available, and therefore, income method is used. For a sector like construction it is the expenditure figures, which are readily available to employ expenditure method.

Historical Background: Pre-Independence Estimate of National Income

The pre-independence estimates of national income of India were a result of individual efforts by prominent economists. Dadabhai Naoroji was the first to estimate the national income of India. According to his estimate, national income during 1867-68, was Rs. 340 crore and per capita income was Rs. 20. Subsequently, William Digby gave the estimate of national income of India in 1897-98 as Rs. 390 crore and per capita income of Rs. 17. According to Findlay Shirras, national income in 1911 was Rs. 1942 crore and per capita income was Rs. 80. Dr. V.K.R.V. Rao, the well-known expert on national income, estimated national income of India during 1925-29 at Rs. 2301 crore and per capita income at Rs. 78. R.C. Desai's estimates of national income and per capita income for the year 1930-31 were Rs. 2809 crore and Rs. 72 respectively. All these estimates, though valuable, suffered from lack of adequate data and were not strictly comparable because of different methods and coverage employed. In spite of such limitations, all these estimates do reflect the poverty and economic backwardness of India before independence.

Post-Independence Estimates of National Income

It was only after attaining independence in 1947 that official estimates of national

income began to be prepared. The first attempt in this direction was the appointment of the National Income Committee on August 4, 1949 under the Chairmanship of Prof. P.C. Mahalanobis, with Prof. D.R. Gadgil and Prof. V.K.R.V. Rao as its members. This committee submitted its first report in 1951 and the final report in 1954. In its first report, the committee gave official estimates of national income for 1948-49 besides giving the conceptual framework of national income computation. In the final report, the committee published revised estimates for 1948-49 and fresh estimates for 1949-50 and 1950-51 both at current and constant prices (with 1948-49 as the base year). The committee also suggested methods and procedures to be followed for the preparation of estimates of national income.

Following the recommendation of the National Income Committee, the National Sample Survey (NSS) was set up in 1950 to collect required data for national income estimation on a regular basis.

Central Statistical Organization (CSO)

In 1954, the Government of India set up an official agency by the name of Central Statistical Organization (CSO) for estimation of national income. CSO published the first issue of white paper on national income in the year 1956 and since then it has been publishing regularly annual official estimates for national income under the title 'National Accounts Statistics'.

CSO functions under the Department of Statistics, Ministry of Planning, Government of India. Estimates of national income, both at current and constant prices, are prepared along with various related macro-aggregates over a number of years. Besides, CSO keeps publishing "National Accounts Statistics - Sources and Methods", from time to time to give the detailed methodology and sources of statistical data for estimation of sector-wise national income. The latest publication on source and methodology by CSO was dated February 1988.

Since its inception, CSO has used five different years, namely, 1948-49, 1960-61, 1970-71, 1980-81 and 1993-94 as base years for the computation of national income at constant prices. Accordingly, there have been four series of national income estimate, viz., (a) conventional series, (b) revised series, (c) new series and (d) revised series on National Accounts Statistics.

14.3.1 Sub-Sectors of Indian Economy

CSO has classified Indian economy into the following 13 sub-sectors divided into 9 groups. Besides, there is an additional fourteenth sub-sector called external transactions. The sectors are grouped as follows:

1) Agriculture, forestry and fishing

   a) agriculture

   b) forestry and logging

   c) fishing

2) Mining and quarrying

3) Manufacturing

   a) registered

   b) unregistered

4) Electricity, gas and water supply

5) Construction

6) Trade, hotels and restaurants

   a) trade

   b) hotels and restaurants

7) Transport, storage and communication

   a) railways

   b) transport by other means

   c) storage

   d) communication

8) Financing, insurance, real estate and business services

   a) banking and insurance

   b) real estate, ownership of dwellings and business services

9) Community, social and personal services

   a) public administration and defence

   b) other services

Methods Used in Measuring National Income

The production, income and expenditure methods are three alternative ways to measure national income of an economy. These methods have already been explained in great detail earlier. Conceptually speaking, if complete statistical information is available, each of the three methods should ultimately lead to the same estimate of national income. But in practice, it is difficult to apply any one method to find out the contribution of each sector to national income. It is possible to combine different methods to measure the contributions of different sectors of an economy. The choice of method depends on the availability of sufficient and reliable statistical data. Central Statistical Organization uses four methods (production, income, expenditure and commodity flow) for arriving at domestic product estimate in India. The economic activity covered by production method has been increasing over a period of time because of the availability of production statistics for more sectors.

Production method, at present, is employed in the commodity production sectors. These sectors are:

1) agriculture,

2) forestry and logging,

3) fishing,

4) mining and quarrying, and

5) registered manufacturing.

Income method is used specially in the services sectors. The sectors covered by it are:

1) unregistered manufacturing,

2) electricity, gas and water supply,

3) transport, storage and communication,

4) trade, hotels and restaurants,

5) banking and insurance,

6) real estate, ownership of dwellings and business services,

7) public administration and defence, and

8) other services.

A combination of expenditure and commodity flow methods is adopted in the construction sector. While expenditure method is used in rural construction, commodity flow approach is used in urban construction.

CSO gathers relevant statistical data for applying the appropriate method to each of the sub-sectors to find out its contribution to gross domestic product at market price. By deducting depreciation and net indirect taxes wherever necessary, the contribution of each sector to net domestic product at factor cost is worked out. Thus, the sum total of net value added at factor cost of all the sectors gives us the value of net domestic product at factor cost. By adding net factor income from abroad to net domestic product at factor cost, net national product at factor cost or national income of India is measured

14.3.2 Method Employed in Agricultural Sector

Agriculture and Allied Activities: The contribution to domestic product from agriculture, livestock and allied activities except the government's irrigation system is estimated by using production method. The process involves estimation of the gross value of products and by-products, ancillary activities of raw materials, services and consumption of fixed capital in the process of production to obtain net value added. For irrigation, an income approach is used. The total factor incomes generated as a result of providing the irrigation service is measured. Separate estimates of the gross value of output for agricultural crops and livestock production are prepared while gross/net value added estimates are worked out for the activity as a whole.

The random sampling method is used to find out the production of each principal crop. In this method, some fields are selected at random in every district where a particular crop is grown. That crop is cut over one hectare of land. The output from it is multiplied by the average wholesale price prevailing in the district market in that season. This gives the value of output of one hectare of a crop. The value of output is multiplied with the number of hectares on which the crop is grown. Similar exercises are undertaken in each district and at the state level. Random sample crop-cutting surveys are conducted annually by state governments for 36 principal crops.

For minor crops, data on the value of output are collected from various sources. The value of output of livestock like milk and milk products, meat and meat products, eggs and poultry is obtained by census method.

For working out gross value added from agriculture and allied activities, deductions are made from the value of output for current cost of inputs in: (a) seed, (b) manure (chemical fertilizers and organic manure), (c) current repairs and maintenance of fixed assets and operational costs, (d) livestock feed, (e) irrigation charges (payable to government), (f) market charges, (g) electricity, (h) pesticides and insecticides and (i) diesel oil.

The consumption of fixed capital is estimated for the following assets, which are used in agricultural production:

a) agricultural implements, machinery and transport equipment,

b) farmhouses and cattle sheds,

c) orchards and plantations,

d) bunding and other land improvements,

e) wells and other irrigation sources, and

f) meat stalls.

From the estimates of gross value added in agricultural sector, the estimates of consumption of fixed capital are deducted to obtain net value added. Value added from agriculture and allied activities for India for the year 1997-98 at constant (1993-94) prices is shown below:

(Rs. crore)

Value of output of agriculture	228544
Value of output of livestock	78683
Value of output of agriculture and livestock	307227
Inputs	- 59011
Gross value added of agriculture and allied activities	248216
Operation of irrigation system	5932
Gross value added in agriculture	254148
Consumption of fixed capital	- 13073
Net value added	241075

Source: NAS, 1999 Source: NAS, 1999

14.3.3 Method Employed in Registered Manufacturing Sector

This sector includes enterprises registered under the Indian Factories Act, 1948. The Act covers the large-scale manufacturing units. Data relating to these enterprises are collected annually both on census and sample basis by National Sample Survey Organisation (NSSO). The survey is known as Annual Survey of Industries (ASI). The census method is used in case of factories. The value of output is obtained from ASI. By deducting from it the value of intermediate consumption (cost of raw materials, fuel and electricity, services, etc.) and consumption of fixed capital, net value added is obtained which is shown for India for the year 1996-97, at constant (1993-94) prices as follows:

(Rs. crore)

Total value of output	545981
Intermediate cost	417244
Gross value added, including bank charges	128737
Imputed bank charges	-11901
Gross value added	116836
Consumption of fixed capital	-24184
Net value added	92652

Source: NAS, 1999

14.3.4 Method Employed in Construction Sector

In this sector, expenditure-cum-commodity flow method is used. Construction activity is classified into two groups, namely, (a) kutcha construction, and (b) pucca construction.

Expenditure method is used in kutcha construction, which is labour-intensive. In it hardly any input of construction material having economic market value is used. It uses such freely available inputs as leaves, reeds and mud. Such construction is usually seen in rural areas where poor farmers and labourers construct their huts, small houses and cattle sheds mostly with their own labour and freely available inputs. Data for this sector are obtained through surveys by NSSO.

The commodity flow approach is adopted in pucca construction. Pucca construction is mostly undertaken in urban areas and involves costly inputs and modern techniques of construction.

The inputs used in pucca construction are bricks, cement, steel, wood, electric fittings and sanitary fittings. Estimated values for these items from are collected from ASI, government agencies and also directly from dealers. After deducting the value of intermediate consumption and consumption of fixed capital from the value of output, net value added in each group is obtained. In India, the value added from construction during 1996-97, at constant (1993-94) prices, is as follows:

(Rs. crore)

Value of output	123544
(a) New construction	100399
(b) Repairs and maintenance	23145
Value of intermediate consumption	76162
Gross value added	47382
Consumption of fixed capital	-2308
Net value added	45074

Source: NAS, 1999

14.3.5 Difficulties in Measurement of National Income of India

The difficulties faced in measuring national income of India can be divided into two categories:

1) conceptual difficulties, and

2) statistical difficulties.

Conceptual difficulties. All the countries, whether developed or developing, face difficulties relating to the concepts used in national income accounting. Even economists differ on what constitutes national income, i.e., what to include in national income. Examples of such difficulties are:

- absence of an agreed definition of national income
- difference between final goods and intermediate goods
- transfer payments
- services rendered without remuneration etc.
- consumers' durable goods etc.

Statistical difficulties

Inadequate and unreliable data

Data available are neither sufficient nor correct and sufficiently detailed. For example, in India, it is possible to get information, though inadequate, from producing units in organised sector. But it is difficult to get data from agricultural and related activities like household, crafts and indigenous credit (functioning of village money lenders). Moreover production is not only unorganised but scattered also.

Non-monetised sector

A significant part of the product, especially product of agriculture and cottage industries in rural India is bartered. Valuation of such goods, which have been exchanged through barter system, becomes very difficult since they do not reach the market.

Production of goods for self-consumption

The small farmers who constitute a sizeable number in India produce goods mainly for their own use. The value of such goods cannot be computed because they do not come to market for sale. Hence a good deal of guesswork in such cases become inevitable.

Illiteracy and ignorance

Majority of the people in India are illiterate, uneducated and ignorant. They do not maintain account of their income and expenditure. These people do not even know, much less give data, about their income and expenditure. They are ignorant of cost accounting, which is being progressively used by most of the developed countries.

Lack of occupational classification

In India most of the people earn their living from more than one occupation. For example, Indian farmers work primarily in agriculture. But when free, they also work in cottage and small industries. Should they be treated agricultural workers or industrial workers? When people have different sources of income, it becomes difficult to know the main source and consequently a large part of income gets excluded from national income.

Valuation of a new good at constant prices

When a new commodity is produced for the first time, it is easy to know its current price but difficult to get its constant price. For example, in India it is difficult to work out the value of a colour T.V. at constant prices of 1970-71 because it was not being produced in that year.

Consumption of fixed capital

By deducting depreciation cost from gross national product (GNP), we get the value of net national product. But it is quite difficult to measure the correct value of consumption of fixed capital (e.g., machines, buildings, etc.) during the year. So some sort of arbitrariness gets involved while estimating depreciated amounts.

Check Your Progress 2

1) Identify the sectors in which production method is employed to find out their contribution to National Income of India.

.....

.....

.....

.....

2) Enumerate the steps involved in finding out net value added of agricultural sector in India.

.....

.....

.....

.....

3) What are the main conceptual difficulties faced in computation of National Income of India?

.....

.....

.....

.....

14.4 LET US SUM UP

In this unit we have given you an account of three methods to measure national income of an economy. These three methods, namely, production (or value added), income and expenditure have been explained in detail along with various points which should be kept in mind before employing each of them.

It has also been shown how reconciliation of the three measures of national income is achieved.

A detailed discussion of the methods to measure national income of India is also attempted after giving a historical account of national income measurement. The sub-sectors of Indian economy as employed by the Central Statistical Organization of India are presented. The details of the methods employed in agriculture, registered manufacturing sector and construction sector are given by way of illustration.

Finally, difficulties, both conceptual and statistical, faced in measurement of national income of India are explained.

14.5 KEY WORDS

Compensation of Employees : It is the factor income generated in an economy in the process of production by rendering labour services. It is also referred to as income from work that includes wages and salaries, employers' contribution to social security, commissions and real wages.

Change in Stocks : It is the difference between total stocks or inventories (including unsold finished products, semi-finished products and raw materials) at the end of the year and total stocks at the beginning of the year.

C.S.O. : Stands for Central Statistical Organization, which is responsible for the publication of 'National Accounts Statistics' every year providing estimates of national income and various other macro-economic aggregates.

Double Counting : When the value of output of a product is included along with the other products in which it has entered as raw material, double counting occurs. It should be avoided in estimation of GDP of an economy.

Depreciation Provision : It is the provision of funds in an economy to account for normal wear and tear and expected obsolescence of national capital stock.

Domestic Territory : It includes geographical or political boundary of an economy besides ships in the international waters and embassies or consulates in foreign countries.

Direct Taxes : The taxes which are imposed on persons or institutions and they are to be paid by them alone. In other words, these are the taxes the incidence of which cannot be transferred to others.

Death Duties : The taxes, which are paid by those who inherit the property from the parents or relations or friends etc.

Durable Goods : These are the goods, which can be used again and again like tables, machines etc. To be more precise they are referred to as durable-use goods.

Economic Growth : It is the growth of real per capita income in an economy.

Economic Development : It is economic growth of an economy coupled with all those factors, which sustain economic growth over a sufficiently long period of time.

Economic Welfare : It is that part of total welfare of an economy, which

is associated with production and consumption of goods and services.

Expenditure Method : It is sum of all expenditures by various economic agents on the goods and services produced by an economy.

Finished Product : It is that part of total production of an economy which is meant for either final consumption on the part of households, government exports, or for capital accumulation purposes.

Factor Incomes : These are incomes distributed to suppliers of factor services (labour, land, capital and enterprise) for work done over a year in an economy.

Gross Value Added at mp : It is the market value of output of an enterprise minus intermediate consumption.

Mixed Income of Self-employed : It is factor income (compensation of employees plus operating surplus) generated by self-employed.

Gross Value Added at factor cost : It is gross value added at market price minus net indirect taxes paid by a producing unit.

Income Method : It is sum of factor incomes currently generated, which is used along with net factor income from abroad to arrive at national income of an economy.

Intermediate Inputs : Inputs employed by a firm in the process of production, which is bought from other firms or producing units.

Industrial Sectors : Sectors, which are identified for the purposes of measuring national income of an economy.

Imputed Rent : It is the value of rent worked out for owner-occupied dwellings.

Intermediate Consumption : It is the value of inputs used by a producing unit.

Net Factor Income from abroad : The factor income (compensation of employees and property and entrepreneurial income) received from abroad by normal residents of an economy minus such factor incomes paid to normal residents of foreign countries.

Net Value Added at Factor Cost : It is gross value added at factor cost minus depreciation allowance.

Net Value Added at Market Price : It is net value added at factor cost plus net indirect taxes.

Non-Durable Goods : These are the goods, which can be used only once.

Net Exports : It is value of exports minus value of imports of an economy.

Output : It can be physical units of a commodity produced

by a producing unit or it can be physical units of a commodity multiplied by its price over a year.

Owner-Occupied Dwellings: These are the houses, which are occupied by their owners.

Own-Account Production : It is the production of goods or services by self-employed households or institutions.

Operating Surplus : It is the factor incomes consisting of rent, interest and profits arising out of ownership and management of capital.

Primary Sector : This is a sector, which consists of agriculture, forestry and logging, fishing, mining and quarrying.

Secondary Sector : This is a sector, which consists of manufacturing (registered and unregistered) and electricity, gas and water supply.

Tertiary Sector : This is a sector which consists of trade, hotels and restaurants, transport, storage and communication, financing, insurance, real estate and business services and community, social and personal services.

Transfer Incomes : These are the incomes, which do not arise on account of factor services; they are incomes transferred from one sector to another or transferred from one unit to another within a sector. They are not included in national income of an economy.

14.6 SOME USEFUL BOOKS

C.S.O, *National Accounts Statistics*, (Latest), Ministry of Planning, Government of India, New Delhi

Hicks J.R., M. Mukherjee and S.K. Ghosh, 1984, *The Framework of the Indian Economy - An Introduction to Economics*, OUP, Delhi (Chapters 11, 12, 13)

C.S.O, *National Accounts Statistics* - Sources and Methods, Ministry of Planning, Government of India, New Delhi, April, 1989.

14.7 ANSWERS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) a) Gross Investment - depreciation (or consumption of fixed capital) = Net Investment.

   b) Net Factor Income from abroad = Net compensation of employees from abroad + net property and entrepreneurial income from abroad.

   Net exports = Value of exports - value of imports.

   c) Gross national product - depreciation = Net national product.

   d) Gross domestic product - depreciation = Net domestic product

2) Various constituents of net factor income from abroad are:

   1) compensation of employees from abroad.

   2) property and entrepreneurial income from abroad.

   3) net retained earnings of resident companies abroad. Net factor income from abroad can be negative when factor income from abroad is less than factor income paid to the rest of the world.

3) a) GDP at MP = Sum of gross value added at market price of all the producing units of an economy.

   b) GDP at mp = Compensation of employees + operating surplus
   + mixed income of self-employed + net indirect taxes
   + net factor income from abroad + depreciation.

   c) GDP at mp = final consumption expenditure of households
   + final consumption expenditure of government
   + gross domestic capital formation + exports - imports.

Check Your Progress 2

1) Agriculture and allied activities of fishing, mining and quarrying and logging.

2) a) Identify the crops.
   b) For each crop find out area under cultivation.
   (c) Multiply (b) by yield per hectare.
   d) Multiply (c) by average price.
   e) From (d) subtract the intermediate cost of agriculture and allied activities.
   f) With (e) add the value of operation of irrigation systems.
   g) From (f) deduct consumption of fixed capital.
   h) The figure arrived at (g) will be the net value added of agricultural sectors.

3) Conceptual difficulties relate to :
   a) absence of an agreed definition of national income;
   b) difference between final goods and intermediate goods;
   (c) transfer payments;
   d) services rendered without remuneration etc.;
   e) consumer durables etc.

   Statistical difficulties consist of :
   a) inadequate and unreliable data,
   b) non-monetised sector,
   c) production for self-consumption,
   d) illiteracy and ignorance,
   e) lack of occupational classification,
   f) value of a new good at constant prices,
   g) consumption of fixed capital estimation.

ignou
THE PEOPLE'S
UNIVERSITY

Indira Gandhi
National Open University
School of Social Sciences

EEC-11
Fundamentals of Economics

Block

6

MACROECONOMICS AGGREGATES

Expert Committee

Mr. Kalyanjit RoyChaudhary
St. Stephens College
University of Delhi
Delhi

Mr. B.S. Bagla
PGDAV College
University of Delhi
Delhi

Mr. R.S. Malhan
St. Stephens College
University of Delhi
Delhi

Mr. R.S. Bharadwaj
Shivaji College
University of Delhi
Delhi

Dr. Gopinath Pradhan
IGNOU

Dr. Narayan Prasad
IGNOU

Dr. Madhu Bala
IGNOU

Dr. Kaustuva Barik
IGNOU

Mr. Saugato Sen
IGNOU

Prof. S.K. Singh (retd.)
IGNOU

Course Editor	Course Coordinator	Block Coordinator
Mr. Bhawani Sankar Bagla	Dr. Gopinath Pradhan	Dr. Gopinath Pradhan

Block Preparation Team

Unit No.	Unit Writer	Unit Editor (IGNOU Faculty)
13	Sh. R.C. Malhan	Dr. G. Pradhan
14	Sh. R.C. Malhan	Dr. G. Pradhan

Production	Cover Design	Word Processing
Mr. Arvind Kumar	Ms. Arvinder Chawla	Mrs. Rekha Mishra
Mr. Manjit Singh		Ms. Daisy Lal

March, 2003

ISBN-81-266-0764-5

Further information on the Indira Gandhi National Open University courses may be obtained from the university's office at Maidan Garhi, New Delhi-110 068.

Printed and published on behalf of the Indira Gandhi National Open University, New Delhi, by Professor Kapil Kumar, Director, School of Social Sciences.

Laser typeset by : HD Computer Craft, EA1/75, Main Market Inderpuri, New Delhi-110012. Ph: 25831437 ; 3108 1117

Printed at :

BLOCK 6 MACROECONOMICS AGGREGATES

Introduction

After going through blocks, 1 to 5, which deal with micro-economic insights of economic theory, the learner is exposed to macro economic concepts through Block-6. Basically, an attempt is made to show that at the aggregate level, the income of the nation can be seen to have generated by interrelationships between various economic activities. **Unit 13** depicts the national income as a result of circular flow of aggregate income and expenditure. The remaining unit of the block, i.e., **Unit 14** accounts for the method followed for measuring the national income.

UNIT 15 EQUILIBRIUM IN THE GOODS MARKET

Structure

15.0 OBJECTIVES

We discuss in this unit questions such as: (a) What is goods market? (b) Why is national income, what it is, at any time, rather than larger or smaller? (c) What causes national income to rise or fall? Having gone through this unit, you will get answers to these questions and will be able to get acquainted with:

- goods market;
- equilibrium in goods market;
- how consumption and savings are inter-related with income;
- how consumption and autonomous investment play a role in raising the level of national income; and
- why goods market, left to itself, may not achieve the full employment level of national income and still be in equilibrium, requiring the Government to step in?

These are the subject matters of macroeconomics, which is defined as the study of performance of the economy at the aggregated level measured by gross national product and general price level. You will also learn that goods market by itself will not indicate equilibrium of the economy unless it is accompanied by equilibrium in money market, which will be the subject of discussion of Unit 16.

15.1 INTRODUCTION

In the previous block the concept of national income and its measurement was discussed. It was shown that national income could be measured in three different methods-production, income and expenditure. These give you only an understanding of the concept of national income looked at from different angles. In other words, we were concerned with knowing only how to measure national income generated during a given year. We did not discuss the questions mentioned in 15.0. Answering these questions requires a theory of income determination and the nature of goods market and its equilibrium. In building up such a theory, we shall take up J.M. Keynes' analysis of income determination in its simplest form as the starting point. Goods market equilibrium along with that of money market, which will be discussed in Unit 16, will help us to understand equilibrium of the economy.

15.2 GOODS MARKET

National income and other related variables are known as real variables because they are concerned with flow of commodities/services and factor services provided by primary inputs of production such as land, labour capital and enterprise. These are to be distinguished from flows of money. It is because of this reason goods market is also known as real market as distinguished from money market. For an economy to be in equilibrium, it is essential for real market and money market together to be in equilibrium.

The factors, which basically influence both real market and money market, are interest rate and level of national income. It is quite possible that levels of interest rate and national income required to keep real market in equilibrium may not be the same as required to keep money market in equilibrium. Thus, through the process of adjustment, we work out those unique levels of interest rate and national income, if at all they exist, such that both real and money markets are in equilibrium. An equilibrium is said to be stable if a shift from that point generates forces to bring it back to equilibrium. This type of process of equilibrium for the economy will be discussed in Unit 17.

Important components of goods market are flow of goods/services, generation of factor incomes, consumption on the part of households or consumers, consumption of government, taxes imposed by government, transfer incomes given by government, exports and imports of an economy. Some of the variables mentioned above are influenced by national income and some by interest rate. Of course price level is an equally important variable along with exchange rate. In this unit, we are dealing with short-run equilibrium, following J.M. Keynes, and assume that price level and exchange rate are given, or do not change during the period of our analysis. So, we try to relate the variables in goods market to the national income alone. If we were to talk of long run equilibrium of an economy, price level and exchange rate along with interest rate and national income will become variable. But in this unit, we would not discuss the long-run equilibrium of an economy.

15.3 DEFINITIONS/ASSUMPTIONS

Before proceeding further, let us have a look at assumptions and definitions, which are necessary in an analysis of equilibrium level of national income. It is easier to study complex realities step by step rather than all at once. We, therefore, proceed in a series of simple steps. We would take up the study of forces that determine the level of national income under the simple assumptions that both price level and rate of interest are given, and the economy has only two kinds of units, viz., households and firms. This we term as a simple model of income determination. We shall, next introduce the government and, later, rest of the world sector.

Definitions

In order to develop the simple theory of income determination we shall be frequently using the terms such as national income, equilibrium income, autonomous and induced expenditure and circular flow of income. Some of these terms, like national income and circular flow have already been discussed in previous block. In this unit we deal with national income at constant prices only.

i) Equilibrium in National Income

This concept of equilibrium relates to a situation of state of rest, which is usually brought about by balance between opposing forces like that of demand and supply. National income is said to be in equilibrium when there is no tendency for it either to increase or decrease. The actual level achieved by it at equilibrium is said to be national income.

ii) Autonomous and Induced Expenditure

Any expenditure that is constant, irrespective of the level of income, is called autonomous expenditure. On the other hand, that expenditure is affected by the level of income and varies directly with it the variable under consideration in the theory is called induced expenditure.

iii) Functions

The relationship between an independent variable and dependant variable/variables is expressed in the form of a function. For example, if consumption level is dependent on level of income, it is expressed as $C = f(Y)$, where C is level of consumption and Y is level of income. This above expression is given in general form. It can be specified in various ways depending upon the nature of relationship. For example, if $C = a + bY$, it is said to be a linear relationship where C and Y are referred to variables same as explained above and 'a' is the level of consumption when income is zero and 'b' is the rate at which the consumption rises with the increase in level of income.

Assumptions

i) Potential income over short period is constant

The productive capacity of an economy, which consists of both natural and man-made resources, changes slowly over time, e.g., from year to year. Therefore, it is not unrealistic if one assumes that potential income, given the productive capacity of the economy, remains constant over short period. Such potential income is known as full employment level of income. With this assumption one can isolate the forces that determine national income over a short period.

ii) Unemployed supplies of all factors of production

This assumption implies that output can be increased by using currently unemployed factors of production at a constant cost. This also means we have excess capacities in the economy.

iii) Price level is held constant

This assumption is an offshoot of (ii), which means that there are excess capacities. Firms would like to sell more than what they are selling at existing prices and workers would be willing to work at the existing wages. Under these conditions both product and factor markets adjust quantities rather than prices in the face of short run fluctuations in demand. In terms of aggregate supply curve, i.e., supplies of individual firms together.

The aggregate supply curve is flat up to full employment level of income. Output can vary upto this level at constant prices.

iv) **Short run analysis**

We are concerned only with short-run equilibrium and need not consider what happens beyond this level

15.4 DETERMINATION OF EQUILIBRIUM LEVEL OF INCOME

Our economy consists of two kinds of spending units, viz., households and firms. The economy produces two kinds of commodities and services, i.e., consumer goods and services and investment goods. Consumer goods and services are sold by firms to the households and investment goods are sold by their producers to other firms. Aggregate spending or demand is, therefore, determined by consumption spending and private investment spending. For the time being, we assume that there is no government and no foreign trade in the economy. If production, or aggregate supply equals aggregate demand we would be able to get equilibrium level of income.

We have to specify the consumption behaviour of households and investment behaviour of business firms. We will take aggregate supply as being given. At a point where aggregate supply equals aggregate demand, an equilibrium level of income will be determined. Before we go to discuss the whole process of income determination it is necessary that we introduce the concept of 45° line which will be employed quite often during the course of discussion of income determination. At the same time we must be clear about the concepts of injections and leakages in the process of income generation.

15.4.1 Concepts of Injections and Leakages in Income Flow

For understanding the concept of equilibrium level of income it is essential to understand how does a change in autonomous variable lead to the process of change in income through injections and leakages before it gets settled at an equilibrium level.

15.4.1.1 Injections

Suppose there is an increase in the level of investment or an increase in the level of government expenditure. This would imply that more incomes would get distributed to factors of production or to business firms in the first round. These additional incomes subsequently in the next round will create a situation where a part of the increased income will be spent on the purchase of consumer goods and services and the remaining part will be saved. To the extent additional income gets used for financing consumption expenditure, this would lead to additional incomes which in the next round will lead to further increase in income and the process will go on. This process of income generation will continue until an equilibrium level of income is determined. Note that investment and government expenditure act as injections in the income flow.

15.4.1.2 Leakages

Next pertinent question to ask is, how far the income change can continue with a given dose of injection? The answer to this is given by leakages. As injections give rise to change in income, leakages also start taking place. To continue with the example given in 15.4.1.1, an increased in income in the first round gives rise to increased consumption expenditure. But the entire increment is not used for financing consumption expenditure. A part of it is saved. To the extent that increase in savings takes place, the process of rise in income gets arrested. So, saving acts as a leakage in income

increase process. Similarly a part of increase in income, initiated by a dose of investment increase or government expenditure increase, goes to government back in the form of an increased in tax revenue which again acts as a leakage in income increase process.

Thus we have seen that savings and taxes are leakages in income increase process. How much increase in consumption takes place in an economy depends upon marginal propensity to consume ($\Delta C/\Delta Y$) where ΔC represents change in consumption and ΔY change in income. Similarly how much increase in savings take place in an economy depends upon marginal propensity to save ($\Delta S/\Delta Y$) where ΔS is change in savings and ΔY is change in income. The increase in tax revenue will depend upon income tax rate (t) such that tax revenue will be equal to tY, where t is proportion income tax and Y is level of income.

15.4.1.3 Injections and Leakages taken Together

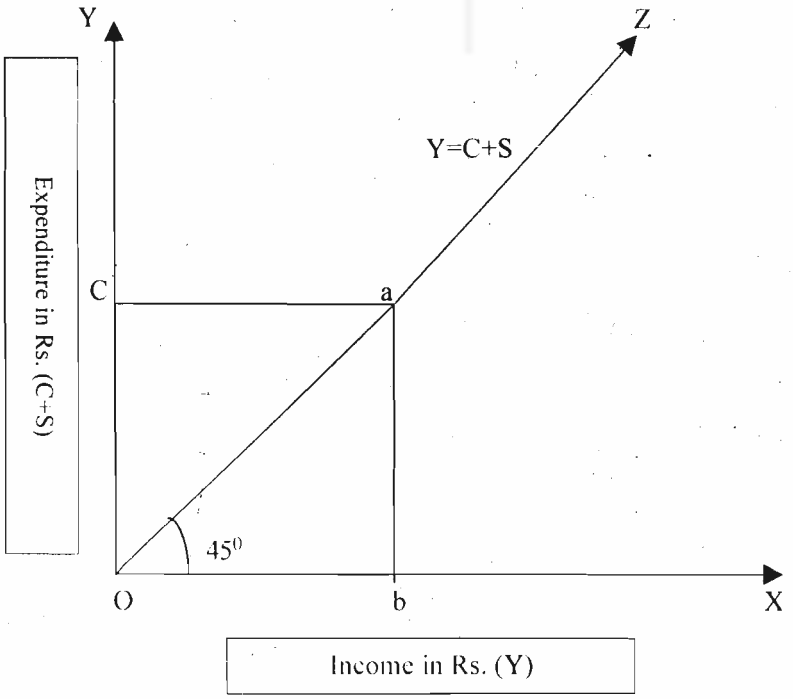
Let us take injections and leakages together in income change process. Initially an injection of Rs. 10 crore in the form of an increase in investment or government expenditure initiates the process of income increase. The process proceeds in a number of rounds. Consequently, savings/ tax revenues also keep rising. Such a process of increase in income will continue until the savings plus tax revenue together become equal to Rs. 10 crore. It is at that point equilibrium level of income gets determined.

Thus, injections = leakages becomes the necessary condition for equilibrium. We would go into further details to explain this process, because, at the point where injections = leakages will also be a point where aggregate demand of the economy will be equal to aggregate supply.

15.4.2 The 45° Line and its Importance

Throughout this unit we will use diagrams which will have a 45° line. Therefore, it is essential to explain the significance of this line. Let us first draw a 45° line.

Fig 15.1



In this diagram on X-axis we measure income and on Y-axis expenditure. A line passing through the origin O is drawn with an angle XOZ equal to 45°. The basic feature of 45° line is that if we take any point on this line, say, point Q the perpendicular drawn on X-axis from it will be equal to perpendicular drawn on Y-axis. So ab = ob. The same will be true from any such point or points on this line.

The implication of this line is that an income equal to ob, also represents income ab because ab = ob. Another way of saying the same thing is that if we have an income equal to ob, this will be used for consumption and for saving purposes, which would constitute the expenditure out of the income ob. Thus, income ob is definitionally identical to expenditure ab. Therefore, 45° line is also referred to as income = expenditure line. If Y represents income, C consumption expenditure and S savings, we can call this Y = C+S line. How much is consumed and saved will be shown by lines C and S. Once we are able to identify the level of consumption at a given level of income, savings can be found out by deducting C from Y.

15.4.3 Consumption Function

The households' consumption expenditure is largely a function of their disposable income defined as the difference between total income and direct tax liability. Using C for households consumption and Y for the income, the functional relationship between C and Y can be expressed as C = f (Y).

The functional relationship between C and Y is termed as consumption function. According to Keynes, consumption function depicts a fundamental psychological law that as Y increases, C also increases but not as much as the increase in Y. That is, C is directly related to Y but does not increase as fast as incomes. The exact relationship between C and Y can be explained with the help of average propensity to consume (APC) and marginal propensity to consume (MPC). Total consumption expenditure divided by total income is called APC. Notationally, APC = C /Y, where APC = average propensity to consume, C = total consumption expenditure and Y = total income.

The relationship between change in consumption expenditure due to change in income is termed as MPC. It is derived as the rate of change in consumption expenditure to the given change in income. Notationally, MPC = $\Delta C/\Delta Y$, where MPC is the marginal propensity to consume, ΔC is change in consumption expenditure and ΔY is change in income.

The concepts of consumption function, APC and MPC are explained in Table 15.1. Columns 1 and 2 of the table give the consumption function. It is seen that at zero level of income, households incur a consumption expenditure of Rs. 800 crore for their subsistence. This consumption expenditure is financed through dissaving.

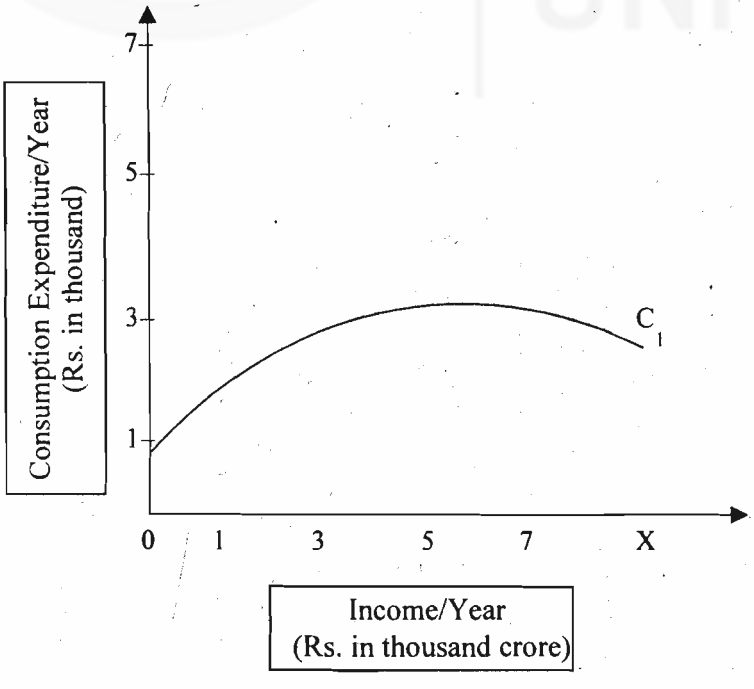
As households' income increases to Rs. 1000 crore, the consumption expenditure goes up to Rs. 1500 crore resulting in a dissaving of Rs. 500 crore. But when the income increases to Rs. 2000 crore, the households break even. At this level of income, consumption expenditure equals income and saving is zero and so is dissaving. With further increases in income positive saving emerges after the consumption expenditures are met.

Table 15.1: Household's Consumption Schedule

Income (Y)/Year (Rs. crore)	Consumption expenditure (Rs. crore)	APC = C/Y	MPC = ΔC/ ΔY
1	2	3	4
0	800	-	-
1000	1500	1.50	0.70
2000	2000	1.00	0.50
3000	2400	0.80	0.40
4000	2700	0.67	0.30
5000	2900	0.58	0.20
6000	3000	0.50	0.10
7000	3050	0.43	0.05

From Table 15.1 it is also noticed that when the households income increases from 0 to Rs. 1000 crore, the consumption expenditure increases from Rs. 800 crore to Rs. 1500 crore. Subsequent increases in income from Rs. 1000 crore to Rs. 2000 crore, raise the consumption expenditure from Rs. 1500 crore to Rs. 2000 crore. Thus, the first increase in income by Rs. 1000 crore raises consumption expenditure by only Rs. 700 crore and the next increase in income by another Rs. 1000 crore, increases consumption expenditure by Rs. 500 crore. This behaviour of consumption expenditure is based on Keynesian psychological law of consumption, which postulates that with the increase in income, consumption increases but not as fast as the income does

Fig. 15.2



The information contained in Columns 1 and 2 of Table 15.1 can be traced graphically to obtain the consumption curve CC_1 in Fig. 15.2. The X-axis measures income per

year in crore of rupees. The CC_1 curve starts at point C on the Y-axis at the consumption level of Rs. 800 crore and slopes upwards to the right. It is concave from below.

In Table 15.1 Column 3 shows APC that is obtained dividing Column 2 by Column 1. As seen in the table, APC is more than 1 so long as total income falls short of the consumption expenditure. It is equal to 1 when income equals consumption expenditure and less than 1 when income exceeds consumption expenditure. The fall in APC is explained by Keynes' 'psychological law of consumption'.

MPC is worked out in Column 4 of Table 15.1 by dividing the change in consumption expenditure by the change in income. For example, when income increases from 0 to Rs. 1000 crore, i.e. $\Delta Y = 1000$, the consumption expenditure increases from Rs. 800 crore to Rs. 1500, i.e. $\Delta C = 700$. Therefore, MPC = 700/1000 = 0.7. In our illustration, MPC is also falling.

Saving Function

A corollary to the concept of propensity to consume is the concept of propensity to save as income is either spent on consumption or saved, yielding the equality Y = C+S. Therefore, at each level of income, saving can be worked out by deducting consumption expenditure from income, i.e., S = Y - C. This is done in Column 5 of Table 15.2. At lower levels of income, saving is negative, but as the income increases, saving becomes positive and increases. It can be noted from the table that at zero level of income, saving is (-) 800 crore. As income increases to Rs. 2000 crore saving becomes zero. This is the break-even point. As income exceeds Rs. 2000 crore, saving becomes positive. The relationship between saving and income at various levels can also be analysed with the help of average propensity to save (APS) and marginal propensity to save (MPS). APS and MPS have been worked out in Columns 6 and 7 respectively of Table 15.2. It is observed that both APS and MPS increase with the increase in income.

Table 15.2 : Households' Consumption and Saving Schedule

Income (Y)/ Year (Rs. crore)	Consp. Exp.(C)/Yr. (Rs. crore)	APC C/Y	MPC ΔC/ ΔY	Saving(s) (Rs. crore)	APS S/Y	MPS Δ S/ ΔY
1	2	3	4	5	6	7
0	800	-	-	-800	-	-
1000	1500	1.50	0.70	-500	-0.50	0.30
2000	2000	1.00	0.50	0	0.00	0.50
3000	2400	0.80	0.40	600	0.20	0.60
4000	2700	0.67	0.30	1300	0.33	0.70
5000	2900	0.58	0.20	2100	0.42	0.80
6000	3000	0.50	0.10	3000	0.50	0.90
7000	3050	0.43	0.05	3950	0.57	0.95

Note that APC and APS add up to 1. This is because

$C + S = Y$

or $\frac{C}{Y} + \frac{S}{Y} = \frac{Y}{Y}$ (dividing both sides by Y)

or, APC + APS = 1.

Similarly, MPC and MPS add up to 1. To prove this start with

$C + S = Y$, so that

$\Delta C + \Delta S = \Delta Y$

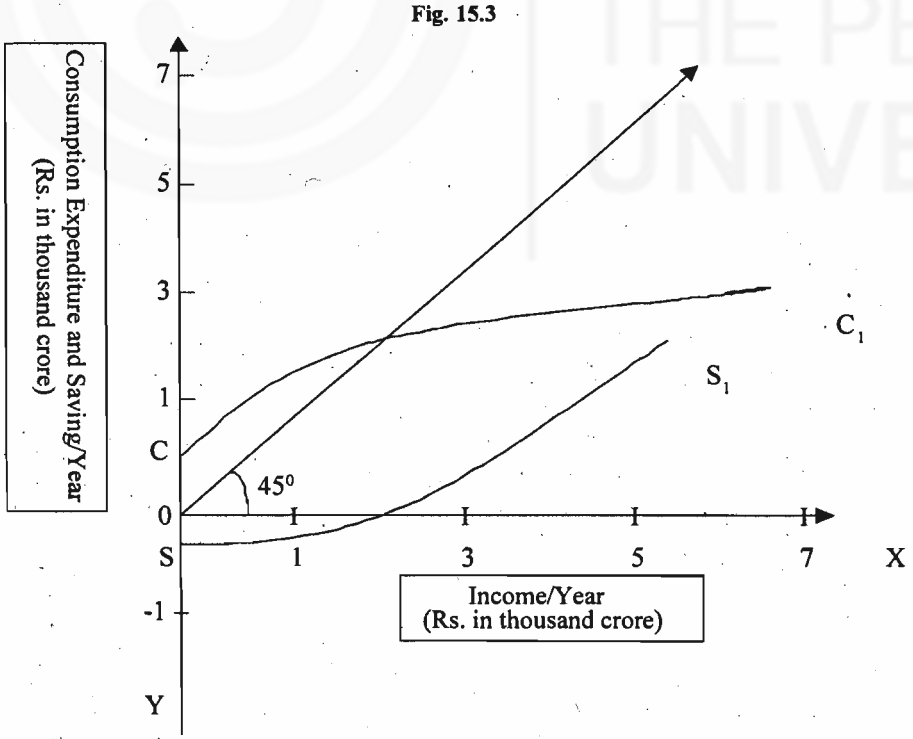
i.e., $\frac{\Delta C}{\Delta Y} + \frac{\Delta S}{\Delta Y} = \frac{\Delta Y}{\Delta Y}$ (dividing both sides by ΔY)

or, MPC + MPS = 1

To represent the saving function graphically and to show its relationship with the consumption function, let us return to consumption function diagram and introduce the 45^0 line. With income on the X-axis and consumption and saving on the Y-axis, the 45^0 line starting from the origin is equidistant from the two axes. With the same scale of measurement used on the X-axis and Y-axis, the 45^0 line, therefore, depicts the equality between consumption and saving. In other words, at every point on the 45^0 line, C + S = Y. Therefore, the 45^0 line can also be labelled as the C + S = Y line. With the help of consumption function (CC_1) and the 45^0 line (C + S = Y), the saving function (SS_1) is drawn in Fig. 15.3 as follows:

The saving curve starts from point S on the OY axis, intersects the X-axis at the income level of Rs. 2000 crore and slopes upward to the right. Since S = Y - C, the vertical distance between SS_1 curve and the X-axis is the same as the vertical distance between CC_1 curve and the 45^0 line at each level of income. The saving curve (SS_1) is concave from above.

Fig. 15.3



For purposes of illustration, very often the consumption and saving functions are drawn as straight lines, which implies that MPC and MPS are constant. This means that for

subsequent increases in income by the same amount, the rates of changes in consumption and expenditure remain constant. This is illustrated in Table 15.3.

For every increase in income by Rs. 1000 crore, the consumption expenditure increases by Rs. 800 crore and the saving changes by Rs. 200 crore. Consumption and saving curves drawn on the basis of information in Table 15.3 are straight lines. This is done in Fig. 15.4.

Consumption and saving functions are not just statistical realities. They are planned, expected or ex-ante relationship between income and consumption, and income and saving. Like all planned or intended relations, they reveal the expected amounts of consumption and saving at different levels of incomes.

Fig. 15.4

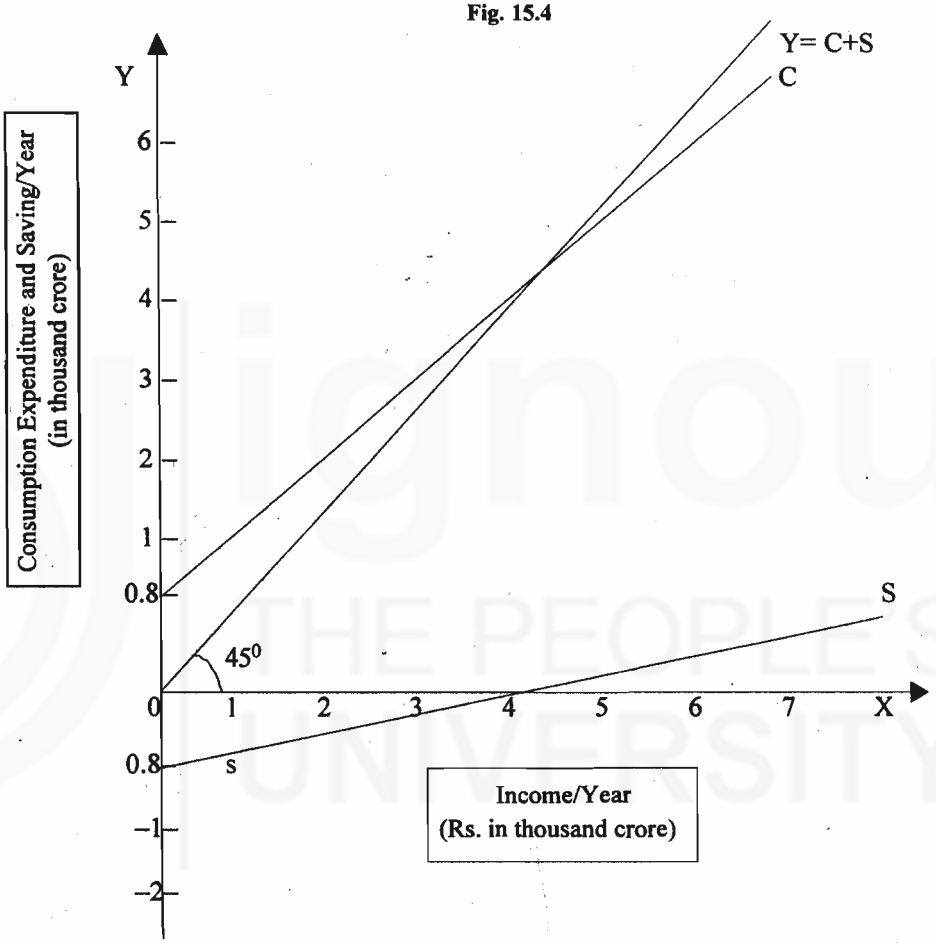


Table 15.3 : Households' Consumption and Saving Schedule

Income (Y)/ Year (Rs. crore)	Consumption Expenditure (C)/Yr (Rs. crore)	APC C/Y	MPC ΔC/ΔY	Saving (S)/Yr (Rs. crore)	APS S/Y	MPS ΔS/ΔY
1	2	3	4	5	6	7
0	800			(-)800		
1000	1600	1.60	0.8	(-)600	(-)0.60	-0.2
2000	2400	1.20	0.8	(-)400	(-)0.20	-0.2
3000	3200	1.07	0.8	(-)200	(-)0.07	-0.2
4000	4000	1.00	0.8	0	0	0
5000	4800	0.96	0.8	(+)20	.04	0.2
6000	5600	0.93	0.8	(+)400	.07	0.2
7000	6400	0.91	0.8	(+)600	.09	0.2

Consumption and saving functions depend, in addition to income, on a number of factors like distribution of income, wealth in the hands of the people, demonstration effects of higher consumption, prices prevailing in the economy and their expected trend, tax policy of the government and family ties. Since rich persons save more than poor do, a redistribution of income in favour of the poor may reduce saving. Similarly, an increase in consumption expenditure to catch up with one's rich neighbours, may also affect saving adversely. Changes in population, prices and direct taxes imposed by the government also affect the level of income, consumption expenditure and saving. Strong family ties and the desire to provide enough for children, desire to amass wealth to gain respectability and so on also motivate the people save more.

Since all the factors determining consumption and saving are of a diverse nature, their total impact on consumption and saving functions may be difficult to be ascertained correctly unless some restrictive assumptions about their behaviour are stipulated.

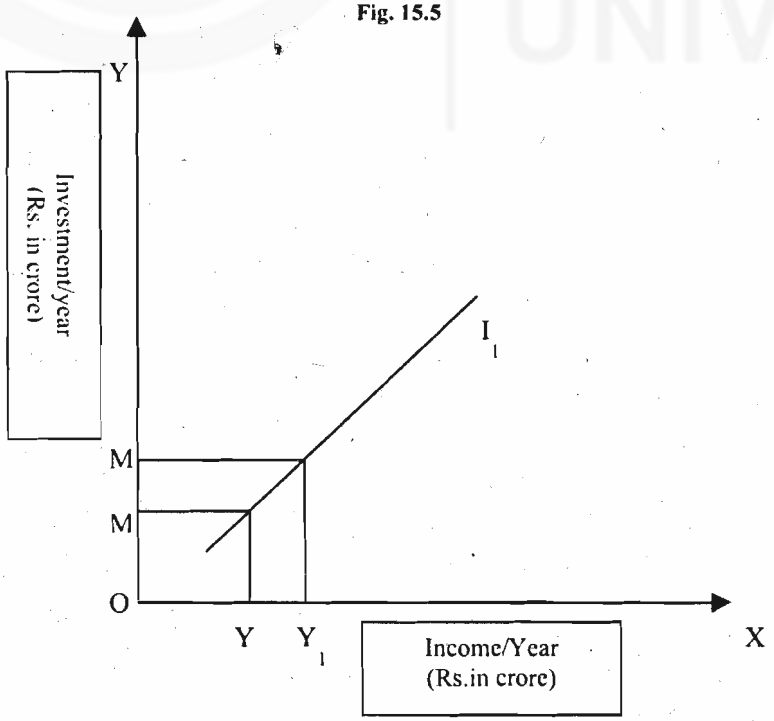
15.4.4 Investment Function

A second important component of aggregate demand is the private investment demand. You would recall that, in economics, investment is defined as expenditure on new man-made real assets such as machines, equipment, buildings, inventories, etc., that help in production. Investment expenditure is very important, for, it helps in creating the capacity to produce. It is through investment that an economy can grow. That is why the underdeveloped countries have been focussing their attention on increasing the level of investment expenditure.

Induced Investment and Autonomous Investment

Investment expenditure is of two types, viz., (a) induced investment and (b) autonomous, or independent investment. Induced investment is a function of income. As income increases, people are likely to spend more on consumer goods. This calls for an increase in the production of such commodities, for which investment has to be made. An increase in income thus induces investment. The concept of induced investment is illustrated graphically in Fig. 15.5.

Fig. 15.5

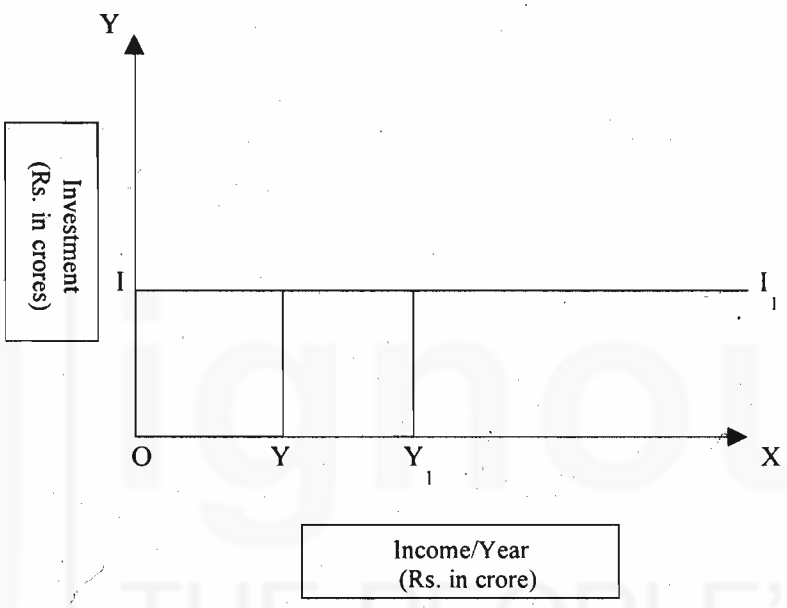


With income shown on the X-axis and investment on the Y-axis, induced investment

curve I_1, slopes upward from left to right. It shows that as income increases from OY to OY_1, the investment expenditure goes up from OM to OM_1 . In other words, an increase in income of the order YY_1 causes an induced investment of MM_1.

Autonomous investment is independently determined by a large number of exogenous factors like population growth, technological changes and discoveries of new resources. Levels of income and consumption expenditure do not have any impact on this type of investment. It can be graphically depicted by an I_1 curve drawn parallel to the X-axis in Fig. 15.6. It shows that investment remains autonomously fixed at a level of OI, irrespective of the level of income, whether OY or OY_1 in the economy.

Fig. 15. 6



Factors Determining Investment

Whether investment is of induced or autonomous variety, it is guided by the profit motive. Private entrepreneurs, while thinking of installing new machines or factories, keep profit uppermost in their minds. They compare expected returns from an investment with its cost to determine profitability. The expected returns from an investment are measured with the help of the concept of marginal efficiency of capital, while the rate of interest determines the cost of investment.

An entrepreneur buys a capital asset because it is expected to yield a flow of returns over its lifetime. Returns depend on the expected productivity of the asset and the prices at which the commodity produced would be sold on the one hand and the costs of production involved on the other. By subtracting the expected costs of production from the expected sale proceeds, expected annual returns from the asset can be worked out. Let these expected annual returns be denoted by R_1, R_2, R_3....., R_n, where R denotes return or yield and subscripts 1,2,3....., n represent the years over which the asset would be yielding these returns.

To ascertain whether investment expenditure on the asset would be profitable or not, an entrepreneur has to compare the expected returns from the asset with the cost of financing investment. The investment finance may be raised from the market or provided by the entrepreneur out of her own savings. In either case, there is a rate of interest involved, which constitutes the cost of financing investment.

To compare the expected returns with the rate of interest, Keynes uses the concept of marginal efficiency of capital (MEC). Defined as the rate of discount, MEC would make the present value of the series of annuities given by the returns expected from the capital asset during its life just equal to supply price. Supply price is the cost of the asset. It denotes the price at which the asset is acquired. Thus, MEC is the discount rate, which equalises the present value of the expected returns from the use of the asset during its lifetime with its supply price.

To calculate MEC the help of following discount rate formula is resorted to:

$$C = \frac{R_1}{(1+e)^1} + \frac{R_2}{(1+e)^2} + \frac{R_3}{(1+e)^3} + ... + \frac{R_n}{(1+e)^n}$$

where C = supply price of the asset, i.e., the cost at which the asset is purchased,

$R_1, R_2,...,R_n$ = expected returns from the asset during its lifetime,

and e = the rate of discount or MEC that equates the present value of the prospective yields with C.

An application of the formula can be illustrated with the help of a simple arithmetic example. Let us suppose that the supply price of an asset is Rs. 4200 and expected return from its use is Rs. 2420 per annum during its life of two years. Applying MEC formula:

$$4200 = \frac{2420}{(1+e)^1} + \frac{2420}{(1+e)^2}$$

Solving the above equation for e, we get a value of 1/10. Thus, MEC is 10%. This value of MEC has to be compared with the rate of interest to decide about the investment. Obviously, if the rate of interest happens to be less than 10%, it would be profitable to go in for the investment, and if the rate of interest is more than 10%, it would be uneconomical to purchase the asset. At the rate of interest of 10%, it would be just worthwhile to go in for the asset.

Another method of judging the profitability of an asset is to discount the expected returns from its use with the rate of interest. This would give us the present value of the expected returns, which may denoted by V. Thus,

$$V = \frac{R_1}{(1+r)^1} + \frac{R_2}{(1+r)^2} + \frac{R_3}{(1+r)^3} ..... + \frac{R_n}{(1+r)^n}$$

where V = Present value of the expected returns from the asset during its lifetime,

$R_1, R_2,...,R_n$ = expected returns from the asset during its lifetime, and

r = the rate of interest.

To illustrate the use of this method, let us use one simple arithmetic example of an asset yielding a return of Rs. 2420 per annum for two years and discount the same with the rate of interest of 10%.

$$V = \frac{2420}{(1+10/100)^1} + \frac{2420}{(1+10/100)^1} = \frac{2420 \times 10}{11} + \frac{2420 \times 10 \times 10}{11 \times 11}$$

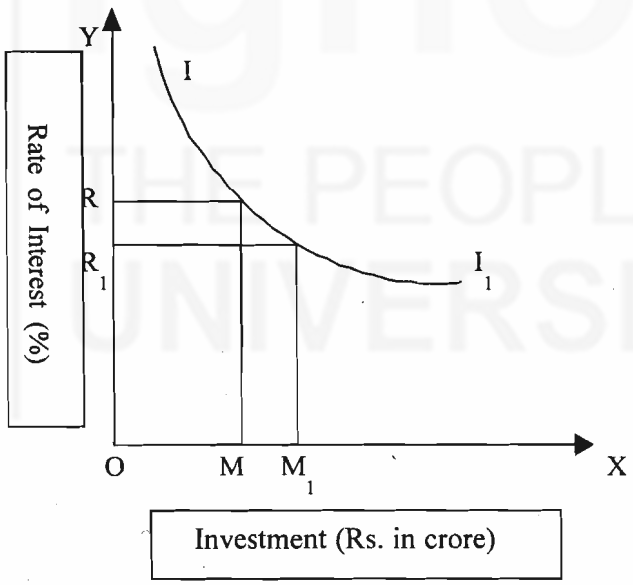
$$= 2200 + 2000 = 4200$$

The present value of expected returns from the asset during its life being Rs. 4200, its installation would be profitable only if it costs less than or equal to Rs. 4200. If the asset were priced at more than Rs. 4200, its installation would not be profitable to the entrepreneur.

From the above, it can be inferred that a profit-maximising entrepreneur goes on investing till MEC equals the rate of interest or present value of the expected returns discounted with the rate of interest is equal of supply price of the asset.

We have said above that MEC is a declining function of investment. That is, it falls as investment increases. This happens because of two factors: First, as the volume of investment increases, the expected yield tends to fall because of the law of diminishing returns. In addition, we may have to take into account the possible fall in the price of the commodity, which the firm produces, on a larger scale with more investment. The diminishing marginal product multiplied by a lower price means a fall in the revenue. From the falling revenue, we have to deduct possibly, the increasing operational costs, which lower the expected returns from the asset. Second, possible increase in the price of the asset as a result of an increase in the demand for it. These two factors lead to declining expected returns and rising supply price of the asset and reduce MEC. This implies that more investment lowers MEC and can, therefore, be justified at a lower rate of interest. What is true of one firm is also true of others. Consequently, the aggregate investment demand curve is downward sloping as shown in Fig. 15.7.

Fig. 15.7



With investment and rate of interest measured on X-axis and Y-axis respectively, II_1 is the aggregate investment curve. It is negatively sloped showing that as rate of interest falls, amount of investment increases. For example, a fall in the rate of interest from OR to OR_1 increases the investment from OM to OM_1.

15.4.5 Consumption Function , Investment Function and 45° Line

15.4.5.1 Aggregate Demand

Aggregate Demand (AD) can be obtained by summing up its components. Ignoring the component of net foreign demand and government demand in order to simplify the analysis, the method of obtaining aggregate demand is explained in Table 15.4

Table 15.4 : Aggregate Demand Schedule

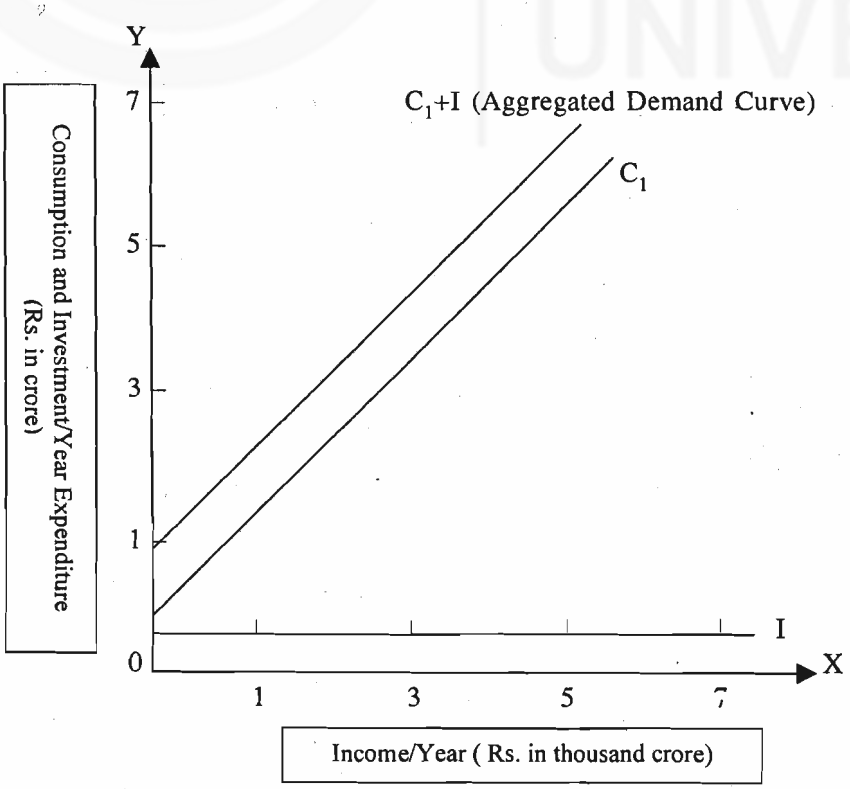
(Rs. crore)

Income/year	Consumption expenditure/year	Investment/Year	Aggregate Demand/year
1	2	3	4
0	800	200	1000
1000	1600	200	1800
2000	2400	200	2600
3000	3200	200	3400
4000	4000	200	4200
5000	4800	200	5000
6000	5600	200	5800
7000	6400	200	6600

It is assumed that investment is fixed at Rs. 200 crore per year and it is autonomous. Adding autonomous investment to consumption expenditure at each level of income, the aggregate demand is derived in Column 4.

Diagrammatically, aggregate demand curve, $C_1 + I_1$ in Fig. 15.8 is derived by adding autonomous investment to the consumption curve CC_1. Since autonomous investment is fixed, the $C_1 + I_1$ curve is parallel to the CC_1 curve.

Fig. 15.8

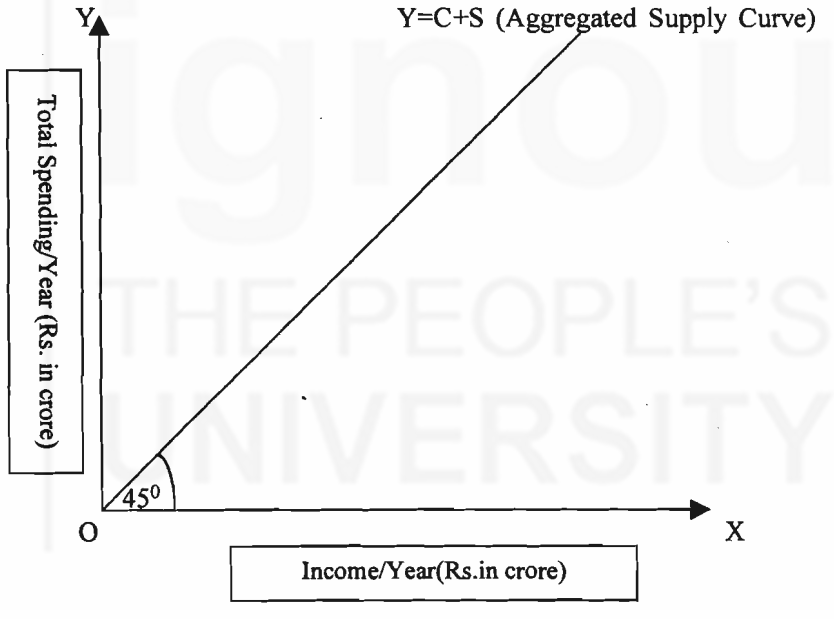


15.4.5.2 Aggregate Supply

Aggregate supply refers to total production of goods and services in an economy. In other words, it refers to country's national product or national income. In producing national output, various costs like wages, rent, interest, depreciation and net indirect taxes are incurred and the producers have to cover them through sale proceeds. Thus, aggregate supply refers to the total spending in the economy, which must cover costs of production.

Measuring income on the X-axis and total spending on the Y-axis, AS curve is depicted by the 45° line from the origin. This line, being equidistant from the two axes on which identical scale of measurement has been used, shows that the total spending covering consumption and investment must equal the income, at each level. If aggregate supply is zero, so is the national income, and if it is Rs. 200 crore, the national income and the total spending are also Rs. 200 crore. Since Y = C + S, the curve is also called Y=C+S line. The aggregate supply curve is shown in Fig.15.9.

Fig. 15.9



15.4.4.3 Income Determination

Having discussed the meaning and constituents of AD and AS, we can explain the process of income determination with the help of a schedule as well as graphs. Table 15.5 lists aggregate demand and aggregate supply in Columns 5 and 6 respectively.

Aggregate demand and aggregate supply are equal at the income level of Rs. 5000 crore. At this level of income, consumption expenditure is Rs. 4800 crore and saving is Rs. 200 crore. The autonomous investment shown in Column 4 is also Rs. 200 crore, which raises the aggregate demand to Rs. 5000 crore. At this level of income, not only is AD = AS, but also S = I . There is no change in the stocks of goods and as such, the equilibrium income has no tendency to vary. Thus, Rs. 5000 crore is equilibrium income.

Table 15.5 : Aggregate Demand and Aggregate Supply Schedule

(Rs. crores)

Income (Y)/ Year	Consumption Expenditure C/Yr	Saving (S)/Yr	Invest ment (I)/Yr	AD/ Yr	AS/ Yr	Stock Varia- tion	Behaviour of income
1	2	3	4	5	6	7	8
0	800	-800	200	1000	0	-1000	Increases
1000	1600	-600	200	1800	1000	- 800	Increases
2000	2400	-400	200	2600	2000	-600	Increases
3000	3200	-200	200	3400	3000	-400	Increases
4000	4000	0	200	4200	4000	-200	Increases
5000	4800	200	200	5000	5000	0	Inequilibrium
6000	5600	400	200	5800	6000	200	Decrease
7000	6400	600	200	6600	7000	400	Decreases

At an income level less than that of Rs. 500 crore, i.e., equilibrium income of Rs. 5000 crore, AD exceeds AS. For example, at the income level Rs. 3000 crore, AD is Rs. 3400 crore while AS is Rs. 3000 crore. The excess of AD over AS means that the entrepreneurs hope to get more for their output than what it costs them. Consequently, the income increases till AD equals AS. It may be seen from Table 15.4 , that at less than full employment level of income (say, Rs. 3000 crore), intended investment (Rs. 200 crore) is more than the intended saving (Rs. 200 crore), which has the effect of pushing up the income till S and I are equalised.

At an income higher than that of equilibrium, AD falls short of AS. For example, at an income level of Rs. 6000 crore, AD amounts to Rs. 5800 crore whereas aggregate supply is Rs. 6000 crore. This shortfall in demand increases the stocks of unsold goods. As a consequence, output and income get reduced till AD = AS. In terms of S and I, at an income level of Rs. 6000 crore, intended saving is more than intended investment (S = Rs. 400 crore and I = Rs. 200 crore). Aggregate demand does not purchase all the goods produced. As a result, unsold stocks of goods pile up, which has the effect of pushing down the income till S and I are equal.

Graphically, the income determination is explained in Fig. 15.10. In the figure income is determined at point E, where AD = AS. It is equal to Rs. 5000 crore. At a lower level of income, say, Rs. 4000 crore, AD exceeds AS and pushes up the income. On the other hand, at a higher level of income, say, at Rs. 6000 crore, AD falls short of AS, which has the effect of reducing the income. The tendency of income to rise or fall comes to rest at point E, where AD equals AS.

15.4.6 Saving and Investment Equality and Equilibrium in Goods Market

In Fig.15.11, an equilibrium position is shown with the help of saving and investment curves. It is determined at point E, where S = I. At a lower level of income, say, Rs. 4000 crore, intended investment exceeds intended saving, which pushes up the income and just the reverse happens at an income higher than the equilibrium level. The tendency of the income to vary comes to rest at point E where S = I. Thus, the condition of equilibrium is that intended saving must equal intended investment.

Aggregate demand at the equilibrium level of income is called effective demand. It equals Rs. 5000 crore in our example above.

Fig. 15.10

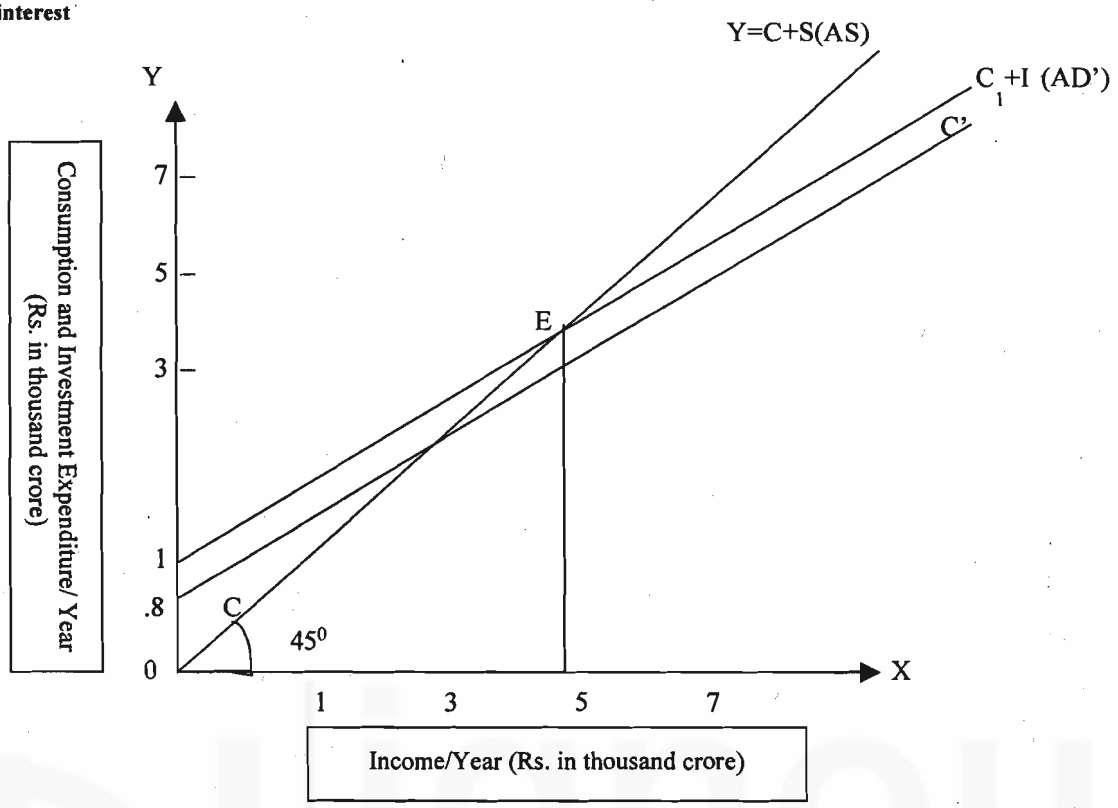
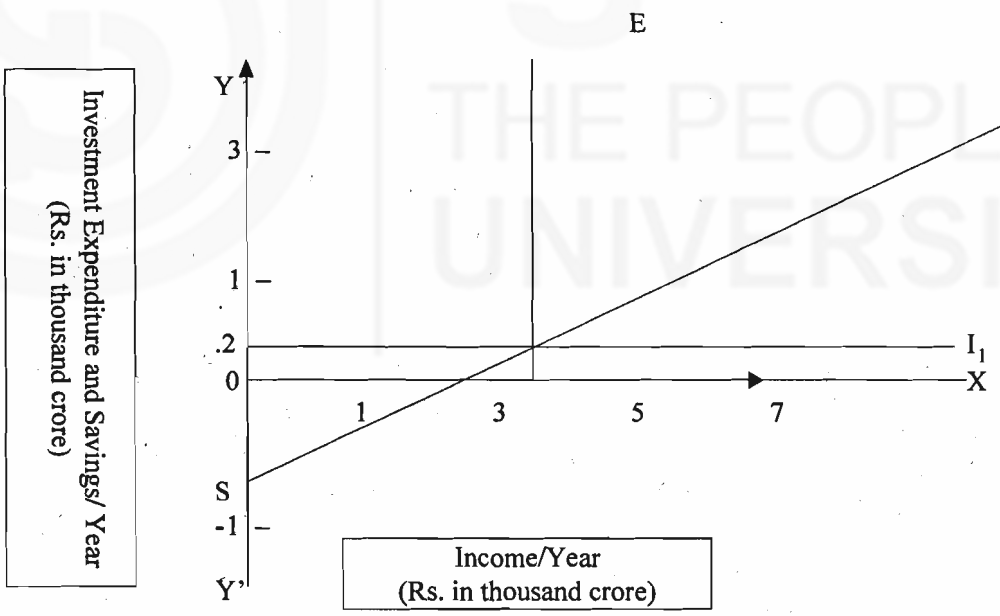


Fig. 15.11



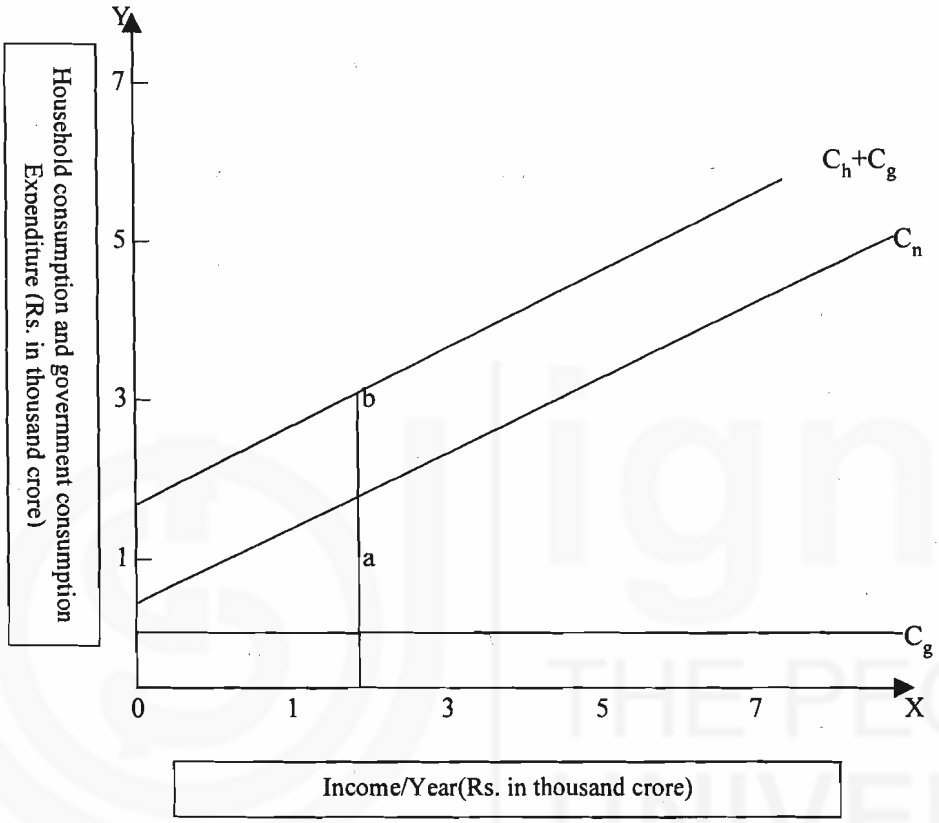
15.4.7 Introduction of Government in Goods Market Equilibrium

Government Demand: A third component of aggregate demand emanates from the government. In a modern welfare state, the government buys large amounts of consumer and capital goods while discharging its functions. It buys furniture, items of stationery, computers, vehicles, essential commodities for public distribution, arms and ammunitions and a lot of other goods to cater to the needs of the society. Very often, government purchases are not linked with any specific economic factors. They are determined by a number of socio-political and strategic considerations. Government

expenditure is also influenced by the fiscal conditions of stability in the country's output and employment. It is because of the variety of factors at work that the government demand cannot be predictably linked with any one economic variable.

The government demand can be incorporated in Fig. 15.11 by drawing a line parallel to C_h line and the gap between the new consumption and C_h line will indicate the level of government consumption at each level of income. It is shown diagrammatically in Fig. 15.12.

Fig.15.12



In this figure C is drawn the same way as in Section 15.4.5.1 and now a new C_h+C_g line is drawn which includes household consumption expenditure as well as government consumption expenditure. For example, at income level Rs. 2000 the household consumption is shown equal to a Rs. 2000 and the government consumption is equal to ab. Because of these constituents, total consumption level is Rs. 2000b (=Rs. 2000a + ab). Similarly we can draw various other points at different levels of income to draw $C_h + C_g$ line. The gap between $C_h + C_g$ and C_h lines is the level of government expenditure. The line $C_h + C_g$ parallel to C_h line implies that the level of government expenditure is constant and does not change with change in level of income.

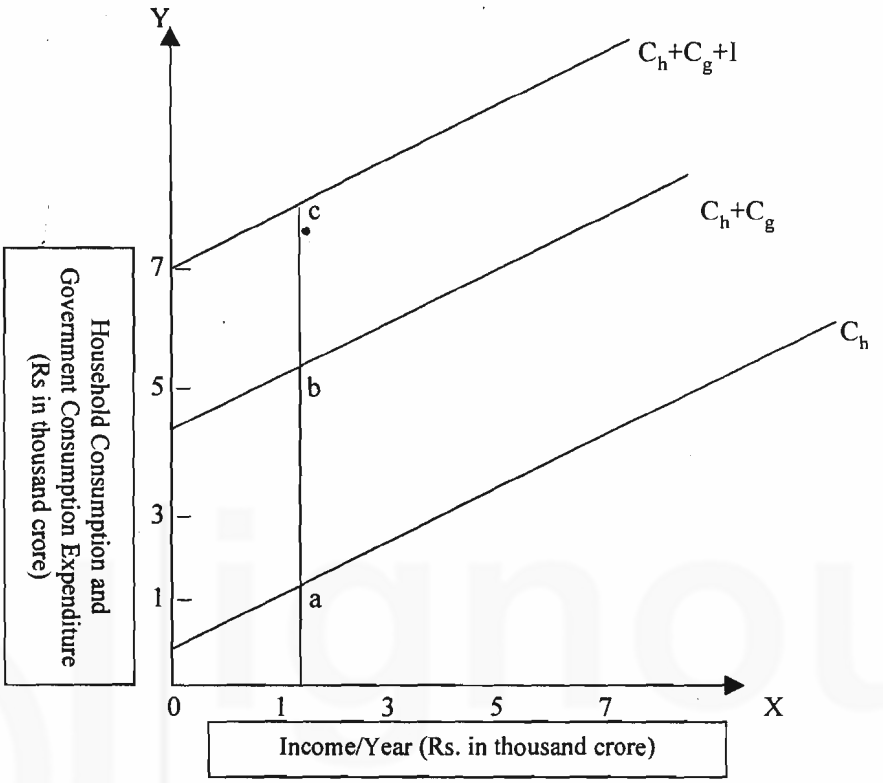
The aggregate demand in the economy will be equal to $C_h + C_g + I$, which is shown in Fig. 15.13.

In the figure it is shown that when income level is Rs. 2000 household consumption expenditure is Rs. 2000a, government consumption expenditure is ab and investment in the economy is bc so that aggregate demand in the economy at Rs. 2000 income is Rs. 2000a + ab + bc = Rs. 2000c.

Again $C_h + C_g + I$ line is drawn parallel to $C_h + C_g$ line because here investment is taken as given and is independent of the level of income. The introduction of government consumption expenditure in income determination means that saving which is a leakage in the process of income growth has to be equated to investment + government

consumption expenditure (where both are injections). This is not shown diagrammatically. You have to take it as an exercise and draw a suitable diagram to explain equilibrium level of income.

Fig. 15.13



You should try to incorporate government consumption expenditure in Fig. 15.11 to get an equilibrium level of income.

15.4.8 Introduction of Rest of the World Transactions in Goods Market Equilibrium

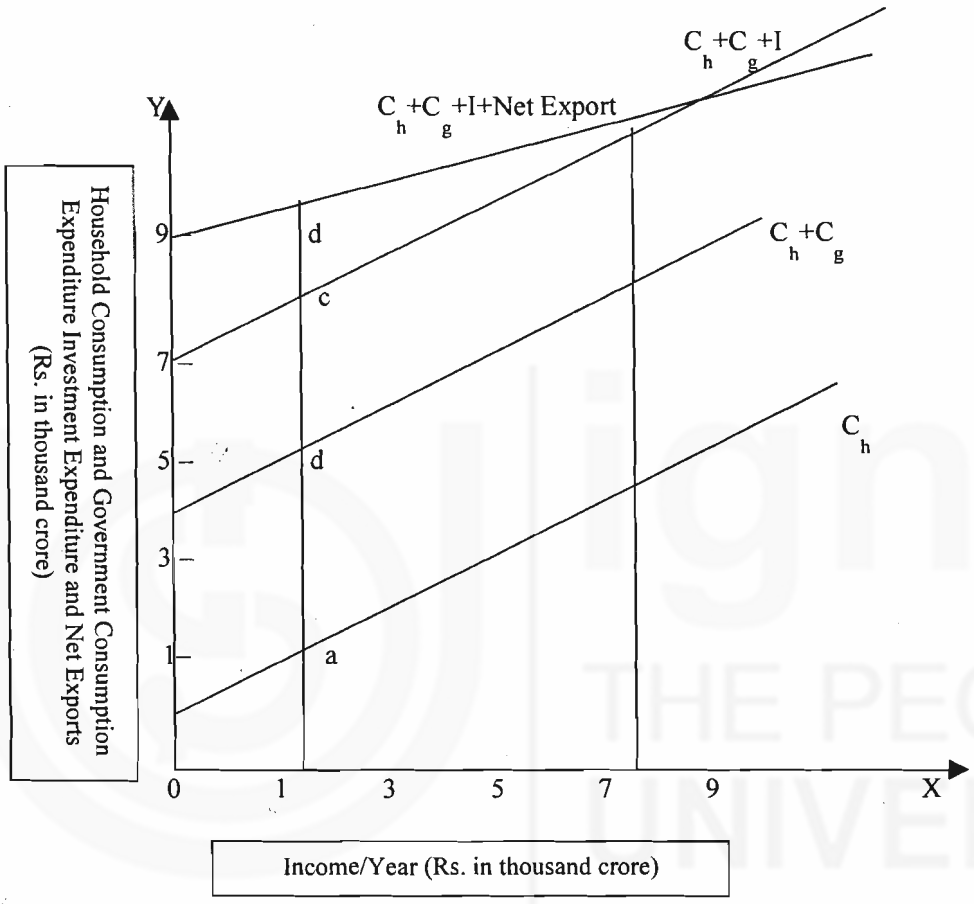
Net Foreign Demand: The modern economies are open economies. They export and import a number of commodities and services. Exports of a country constitute foreigners' demand for its products, whereas imports imply the country's demand for foreign goods. The difference between exports and imports, or net exports constitutes the net foreign demand for a country's output. The net foreign demand is the fourth component of aggregate demand in an economy. It is a function of factors like cost differences in the production of commodities in different countries, income of the people, their tastes and preferences, trading policies of the government, exchange rate, international movements of capital and so on. These factors are diverse and influence exports and imports in a complex manner. Besides, a country's foreign trade may also be guided by non-economic considerations such as political alignment and membership of trading block. Consequently, it may not be possible to clearly predict the behaviour of net foreign demand.

Exports of an economy do not depend upon the level of income of that economy because how much is exported is more dependent on the quality of goods, how far they can be sold at competitive prices and so many other factors. Therefore, exports are taken as an exogenous variable and thus taken as given at different levels of income. Imports on the other hand are dependent on the level of income, or we can state $M = mY$, where M is imports and Y is national income and m is the proportion of imports to national income. Thus, as incomes rise the level of imports increase.

If we talk of net exports i.e., exports minus imports, we would find that as national income rises since imports keep rising and exports are given, net exports will keep falling.

Net exports can be incorporated in aggregate demand of an economy by adding net exports to $C_h + C_g + I$ line, which is shown in Fig. 15.14.

Fig. 15.14



If we study the figure carefully, it is noticed that everything else is the same as in Fig. 15.13 and only new item added is net exports. At Rs. 2000 income, Rs. 2000a is household consumption expenditure, ab is government consumption expenditure, and bc investment expenditure. Aggregate demand of the economy will be equal to Rs. 2000 a + ab + bc + cd = Rs. 2000d at Rs. 2000 level of income. This we have to do at each level of income to get aggregate demand curve which is represented by $C_h + C_g + I$+ net exports line.

It is worth noting that the gap between $C_h + C_g + I$+ net exports line and $C_h + C_g + I$ line keeps falling showing those new exports expenditure keeps falling with increase in the level of income.

Exports act as injection while imports as leakage in income flow and since injections must equalise leakages for equilibrium level of income to be achieved we will find that household consumption expenditure + autonomous government consumption expenditure + autonomous investment expenditure + exports expenditure (which are all injections) should be equal to savings + imports expenditure + tax revenue not introduced in our system (which are all leakages).

The reader should try to incorporate net export expenditure in Fig. 15.11 to get an equilibrium level of income.

Check Your Progress 1

1) i) Why does equilibrium level of income get determined at a point where aggregate demand curve consisting of only household consumption and government consumption intersects the 45° line?

.....

.....

.....

ii) Draw a suitable figure and show the process which leads to a fall in level of income from a point where savings are more than investment. Assume that there is no government and no foreign trade.

.....

.....

.....

iii) What will be the shape of consumption line if consumption at 0 level of income is also 0 and marginal propensity to consume is constant.

.....

.....

.....

.....

2) Assume a straight-line consumption function passing through origin, how will you show average propensity to consume and marginal propensity to consume? Will they be equal?

.....

.....

.....

.....

3) Show diagrammatically that equilibrium level of income will rise if consumption at 0 level of income rises given a constant marginal propensity to consume. Assume that there are no investment, government expenditure and export expenditure.

.....

.....

.....

.....

15.5 CHANGES IN EQUILIBRIUM LEVEL OF INCOME

15.5.1 What Causes Change in National Income?

Changes in AD will obviously alter the equilibrium income. An increase in AD, for example, will raise the level of income whereas a fall will have the opposite effect. The aggregate demand can change not only with a change in the level of autonomous

investment but also with a change in Government consumption expenditure and export expenditure.

15.5.2 Investment Multiplier

To illustrate the effect of an increase in AD, let us assume that autonomous investment increases from Rs. 200 crore to Rs. 400 crore at each level of income, altering AD as shown in Table 15.6. The aggregate demand goes up and the equilibrium income increases from Rs. 5000 crore to Rs. 6000 crore. Thus, an increase in investment by Rs. 200 crore raises the income by Rs. 1000 crore. In other words, there is a multiple effect on income due to a given change in investment. This multiple effect is called investment multiplier. In our example, an increase in investment by Rs. 200 crore increases income by Rs. 1000 crore, giving a multiplier of 1000/200 = 5. The investment multiplier may be denoted by K and we can derive it as follows:

TABLE 15.6 : Aggregate Demand and Aggregate Supply Schedule

(Rs. crore)

Income (Y)/Yr	Consumption C/Yr	Saving (S)/Yr	Investment (I)/Yr	AD/ Yr	AS/ Yr	Stock Variation of income	Behaviour
1	2	3	4	5	6	7	8
0	800	-800	400	1200	0	-1200	Increases
1000	1600	-600	400	2000	1000	-1000	Increases
2000	2400	-400	400	2800	2000	-800	Increases
3000	3200	-200	400	3600	3000	-600	Increases
4000	4000	0	400	4400	4000	-400	Increases
5000	4800	200	400	5200	5000	-200	Increases
6000	5600	400	400	6000	6000	0	In equili-brium
7000	6400	600	400	6800	7000	200	Decreases

Let $Y = C + I$

So that $\Delta Y = \Delta C + \Delta I$

or, $K = \frac{\Delta Y}{\Delta Y - \Delta C}$, i.e., $\Delta I = \Delta Y - \Delta C$

or, $K = \frac{\frac{\Delta Y}{\Delta Y}}{\frac{\Delta Y}{\Delta Y} - \frac{\Delta C}{\Delta Y}}$

(dividing both the numerator and denominator by ΔY)

or, $K = \frac{1}{1 - MPC}$

or, $K = \frac{1}{MPC}$

(Since MPC + MPS = 1, 1- MPC = MPS)

Given the changes in investment, or aggregate demand and the values of MPC, or MPS, we can work out the size of the multiplier. The higher the value of MPC or lower the value of MPS, the higher is the size of K, as shown in Table 15.7.

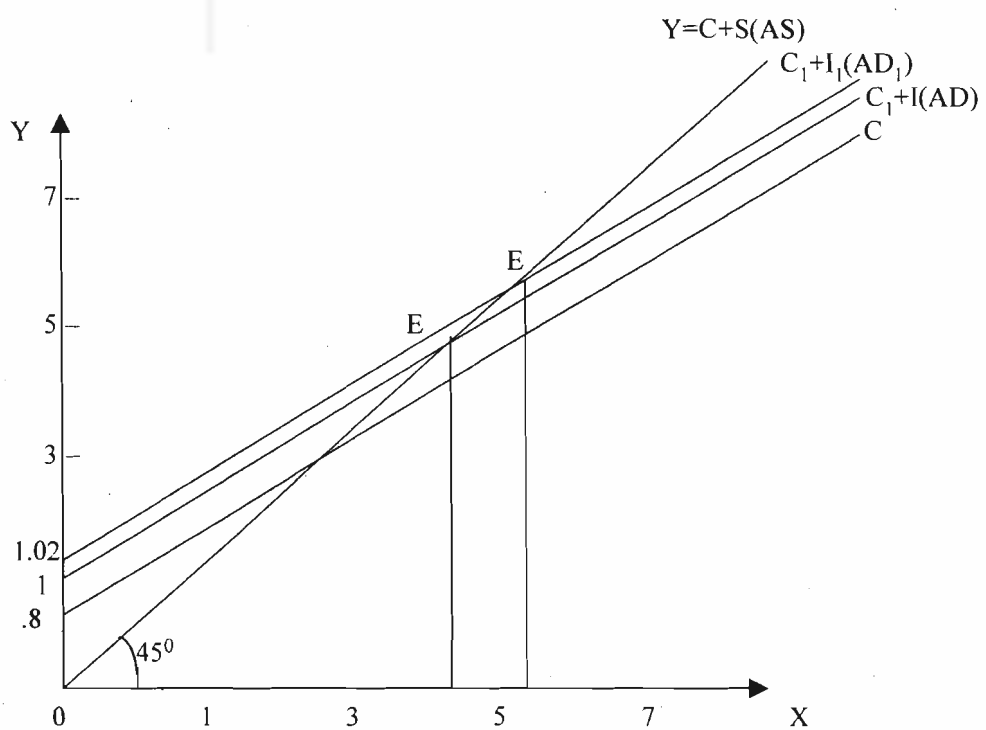
Just as an increase in aggregate demand leads to a multiple increase in income, a fall in aggregate demand reduces the income by a multiple factors in aggregate demand. In other words, multiplier works for increase as well as decrease in aggregate demand.

Table 15.7 : Size of Investment Multiplier

MPC	MPS	K
1	2	3
1.0	0	(Infinite)
0.9	0.1	10
0.8	0.2	5
0.7	0.3	10/3
0.6	0.4	10/4
0.5	0.5	2
0.4	0.6	10/6
0.3	0.7	10/7
0.2	0.8	10/8
0.1	0.9	10/9
0	1.0	1

Increase in income due to a given increase in aggregate demand is shown diagrammatically in Fig. 15.15.

Fig15.15



An increase in investment shifts AD curve from $C_1 + I$ to $C_1 + I_1$. The equilibrium point also shifts from E to E_1 and the equilibrium income increases from Rs. 5000 crore to Rs. 6000 crore. The size of K = 1000/200 = 5.

The size of investment multiplier, K, depends on MPC or MPS. This is shown below :

$$K = \frac{\Delta Y}{\Delta I}$$

Where K = size of the multiplier,
ΔY = change in income,
ΔI = change in investment.

K has already been shown equal to 1 / 1-MPC. Higher the value MPC higher will be the value of investment multiplier. For example, if MPC = .5, multiplier K is 2 and if MPC = .75, multiplier K is 4.

K is also equal to 1/MPS. Higher the value of MPS lower will be the value of investment multiplier. For example, if MPS = .5, multiplier K is 2 and if MPS = .75 multiplier K is 4/3 = 1.33.

Investment multiplier can also be derived differently using national income identities. You know from discussions in units of Block 6 that Y = C + I, where C is consumption expenditure, I is investment expenditure and Y is national income, assuming there is no government and rest of the world transactions. Level of consumption depends on income. In fact, there is a direct relationship between C and Y. This relationship may be put in the form of say, C = cY where c is MPC so that Y = cY + I where I is given to us irrespective of level of Y or we can say Y- cY = I, or Y (1-c) = I or, Y= I / 1-c .....(i)

If investment expenditure rises by ΔI then income will rise by ΔY. We can introduce ΔI and ΔY in equation (i). We get,

$$Y + \Delta Y = \frac{1}{1-c}(I + \Delta I)$$

$$= \frac{1}{1-c} I + \frac{1}{1-c} \Delta I$$

By definition $Y = \left[\frac{1}{1-c}\right] I$, Therefore, Y gets cancelled with

$\frac{1}{1-c} I$ and we are left with $\Delta Y = \frac{1}{1-c} \Delta I$

or, $\frac{\Delta Y}{\Delta I} = \frac{1}{1-c} = K$

15.5.3 Government Expenditure Multiplier

Just like investment multiplier, we can, derive government expenditure multiplier, which is defined as equal to $\Delta Y / \Delta G$. The derivation of the value of government expenditure multiplier is shown below:

$$Y = C + I + G \quad \text{.....(i)}$$

$$C = cY$$

substituting $C = cY$ in (i) we get

$$Y = cY + I + G$$

$$\text{or, } Y - cY = I + G$$

$$\text{or, } Y(1-c) = I + G$$

$$\text{or, } Y = 1/1\text{-}c\ [I+G] \quad \text{.....(ii)}$$

Increasing G by ΔG we get

$$Y + \Delta Y = 1/1\text{-}c\ [I+G+\Delta G] \quad \text{.....(iii)}$$

$$\text{or, } Y + \Delta Y = 1/1\text{-}c\ [I+G] + 1\Delta G\ /1\text{-}c$$

Since $Y = 1/1\text{-}c\ [I + G]$ as shown in (ii), Y on the left hand side of the equation gets cancelled with $1/1\text{-}c\ [I+G]$ on the right hand side and thus, we are left with

$$\Delta Y = \frac{1}{1-c}\ [\Delta G]$$

$$\text{or, } \frac{\Delta Y}{\Delta G} = \frac{1}{1-c}$$

It can be seen that the value of investment multiplier is equal to the government expenditure multiplier.

You should try to show government expenditure multiplier diagrammatically.

15.5.4 Tax Multiplier

So far we have not introduced tax in national income determination. The tax imposed on income can be lump sum, which is imposed on everybody irrespective of the level of income enjoyed. The other way income tax can be imposed is by deciding a tax rate, i.e., percentage of income, which will be taxed. For example, if $t = .2$ it implies 20% of the income will be taxed. The tax revenue will be given by tY.

Let us suppose that the tax imposed is lump sum tax T. The imposition of tax implies that income at the disposal of households will be Y-T, which can be used by them either to finance consumption expenditure or to save. If c is marginal propensity to consume, then consumption expenditure C will be equal to c (Y-T).

Tax multiplier is defined as $\Delta Y/\Delta T$ where ΔY is change in income consequent to change in tax ΔT. See that imposition of tax will leave less income at the disposal of the people to finance consumption expenditure given the MPC and vice versa. The tax multiplier can be derived as follows:

$$Y = C + I + G \quad \text{.....(i)}$$

$$C = c\ (Y\text{-}T)$$

and substituting the value of C in equation (i) we get,

$Y = c(Y-T) + I + G$

or, $Y = cY - cT + I + G$

or, $Y-cY = -cT + I + G$

or, $Y(1-c) = -cT + I + G$

or $Y = 1/1-c\ [-cT + I + G]$ .....(ii)

Let us increase T by ΔT, which will reduce income Y by, say, ΔY. Introducing ΔT and ΔY in equation (ii) we get

$$Y + \Delta Y = \frac{1}{1-c}\ [-c(T+\Delta T) + I + G]$$

$$\text{or, } Y + \Delta Y = \frac{1}{1-c}\ [-cT - c\Delta T + I + G]$$

$$\text{or } Y + \Delta Y = \frac{1}{1-c}\ [-cT + I + G] - \frac{c\Delta T}{1-c}$$

since $Y = 1/1-c\ [-cT+I+G]$ as shown in equation (ii), we can cancel Y on the left hand side with

$\frac{1}{1-c}\ [-cT + I + G]$ to be left with

$$\Delta Y = \frac{-c\Delta T}{1-c}$$

$$\text{or, } \frac{\Delta Y}{\Delta T} = \frac{1}{1-c} \quad \text{.....(iii)}$$

It is interesting to see that the value of Government expenditure multiplier is more than the value of tax multiplier. For example, if c = .5, tax multiplier is -1 while with c = .5, Government expenditure multiplier would have been 2. The implication of this is very clear. With a given MPC of .5, if Government raises its consumption expenditure by, say, Rs. 10 crore then, national income will rise by Rs. 10x2 = Rs. 20 crores. On the other hand, if Government raises tax revenue by Rs. 10 crore then national income will fall by Rs. 10 x 1 = Rs. 10crore.

The reason for tax multiplier being less than Government expenditure multiplier lies in the fact that when tax is imposed the whole of tax is not paid by reducing consumption; partly it is paid by reducing consumption and partly by reducing savings. As you already know savings act as a leakage in income change process.

15.5.5 Balanced Budget Multiplier

If you have followed well section 15.5.4, it will not be difficult for you to arrive at balanced budget multiplier. If Government expenditure is increased by ΔG and is financed by raising lump sum tax, ΔT so that $\Delta G = \Delta T$, then balanced budget multiplier will be equal to 1. This can be shown as follows :

$\Delta G = \Delta T$ by definition, and

$Y = C + I + G$ ...(i) and

$C = c\ (Y-T)$

Substituting the value of $C = c\ (Y-T)$ in (i) we get

$Y = c\ (Y-T) + I + G$

or, $Y = cY-cT + I + G$

or, $Y - cY = -cT + I + G$

or, $Y\ (1-c) = -cT + I + G$

$$\text{or, } Y = \frac{1}{1-c}\ [-cT + I + G] \quad ...(ii)$$

Now increase G by ΔG and T by ΔT such that $\Delta T = \Delta G$. We get

$$Y + \Delta Y = \frac{1}{1-c}\ [-c\ (T + \Delta T) + I + G + \Delta G]$$

$$\text{or, } Y + \Delta Y = \frac{1}{1-c}\ [-cT - c\Delta T + I + G + \Delta G\ ]$$

$$\text{or, } Y + \Delta Y = \frac{1}{1-c}\ [-cT + I + G\ ] - \frac{c\Delta T}{1-c} + \frac{\Delta G}{1-c}$$

since $\Delta T = \Delta G$.

Since equation (ii) tells us that

$$Y = \frac{1}{1-c}\ [\ -cT + I + G\ ]$$

we can cancel Y on the left-hand side with

$$\frac{1}{1-c}\ [-cT + I + G] \text{ in the right}$$

and are left with

$$\Delta Y = -\ \frac{c\Delta T}{1-c} + \frac{\Delta G}{1-c}$$

$$\text{or, } \Delta Y = \frac{\Delta G(-c-1)}{1-c}$$

$$= \frac{\Delta G.(-c-1)}{1-c}$$

$$\text{or, } \frac{\Delta Y}{\Delta G} = \frac{1-c}{1-c} = 1 \quad \text{..... (iii)}$$

This is known as balanced budget multiplier. The implication of balanced budget multiplier is, if Government raises its consumption expenditure by Rs. 10 crore and lump sum tax is also raised by Rs. 10 crore, then national income will rise by Rs. 10 crore. Had the Government increased its consumption expenditure by Rs. 10 crore without increasing lump sum tax, the national income would have increased by more than the increase in Government consumption expenditure.

15.5.6 Export Multiplier

We are in a position now to introduce export multiplier. As already explained in Section 15.4.8 exports are exogeneously determined and imports are an increasing function of national income, Y. The import function can be simply put as $M = mY$. Let us try to derive export multiplier, which is defined as $\Delta Y/\Delta E$ where ΔY is change in income and ΔE is change in exports. This can be derived as follows :

$$Y = C+I+G+E-M \quad \text{..... (i) and}$$

$$C = c\,(Y-T).$$

Substituting the value of C and M in (i), we get

$$Y = c\,(Y-T) + I + G + E - mY$$

$$\text{or, } Y = cY - cT + I + G + E - mY$$

$$\text{or, } Y - cY + mY = -cT + I + G + E$$

$$\text{or, } Y\,[1-c+m] = -cT + I + G + E$$

$$\text{or, } Y = \frac{1}{1-c+m}\,[-cT + I + G + E] \quad \text{.....(ii)}$$

Let us increase the value of E by ΔE, which will lead to increase

in national income by ΔY such that

$$Y + \Delta Y = \frac{1}{1-c+m}\,[-cT + I + G + E + \Delta E]$$

$$\text{or, } Y + \Delta Y = \frac{1}{1-c+m}\,[-cT + I + G + E] + \frac{\Delta E}{1-c+m} \quad \text{.....(iii)}$$

$$\text{since } \frac{1}{1-c+m}\,[-cT + I + G + E]$$

as shown in equation (ii), we can cancel Y on the left-hand side with

$$\frac{1}{1-c+m}\,[-cT + 1 + G + E]$$

on the right had side of equation (iii) and are left with

$$\Delta Y = \frac{\Delta E}{1-c+m}$$

$$\text{or, } \frac{\Delta Y}{\Delta G} = \frac{1}{1-c+m} \quad \text{.....(iv)}$$

Let us take an example. If Rs. 10 crore, with the values of raises E c and m are .5 and .3 then export multiplier is equal to

$$\frac{1}{1-.5+.3} = \frac{1}{1-.2}$$

$$= \frac{1}{.8} = \frac{10}{8} = 1.25$$

This implies that an increase of exports by Rs. 10 crore will lead to increase in national income by Rs. 12.50 crore. The value of export multiplier, as you can see, is less than investment multiplier and Government expenditure multiplier.

Check Your Progress 2

1) i) Suppose investment expenditure goes up by Rs. 100 crore and MPC = .33. What is the extent to which national income will increase.

.....

.....

.....

.....

ii) What would happen to national income if instead of investment expenditure rising by Rs. 100 crore, the Government expenditure rises by Rs. 100 crore, with given MPC = .33

.....

.....

.....

.....

iii) What would happen to national income if whole of increase in Government consumption expenditure is met by raising lump sum tax by Rs. 100 crore, with given MPC = .33.

.....

.....

.....

.....

2) Why is the value of export multiplier less than investment multiplier?

.....

.....

.....

.....

3) How will you derive transfer income multiplier if it is given in lump sum? (Transfer income is defined as government giving cash money to households,

say, in the form of old-age pensions).

.....

.....

.....

.....

15.6 INTEREST RATE, INCOME LEVEL AND EQUILIBRIUM IN GOODS MARKET

You have seen in Section 15.4.3 that consumption expenditure is an increasing function of income level. Interest rate can also influence consumption expenditure which we take as given. Similarly in 15.4.4 you have seen that investment expenditure is a decreasing function of interest rate. Income level can also influence investment expenditure which we have taken as given. Consumption expenditure together with investment expenditure constitutes aggregate demand and given the aggregate supply determines equilibrium level of national income. While talking about aggregate demand let us take government consumption expenditure and net exports expenditure as given though these variables can also be easily incorporated in our analysis.

Now, let us consider the equilibrium in goods market with interest rate and national income as independent variables. We will be able to get innumerable points where savings (leakages) will be equal to investment (injections) giving us equilibrium positions of national income. But we will not be able to get equilibrium of the economy until money market equilibrium positions are incorporated in our analysis. Money market equilibrium will be introduced in Unit 16 and equilibrium of the economy considering both goods market and money market will be discussed in Unit 17. Before we discuss equilibrium in goods market with interest rate and national income as independent variables, it is essential to introduce the tool of four-part diagram.

15.6.1 Four-part Diagram and Derivation of IS Curve

We make use of four diagrams, which are related to each other and are inter-connected while deriving these curves. Value of a variable found out from one diagram is used by the second. This gives us the value of another variable employed by the third diagram and finally in the fourth diagram we are able to get equilibrium position with two independent variables.

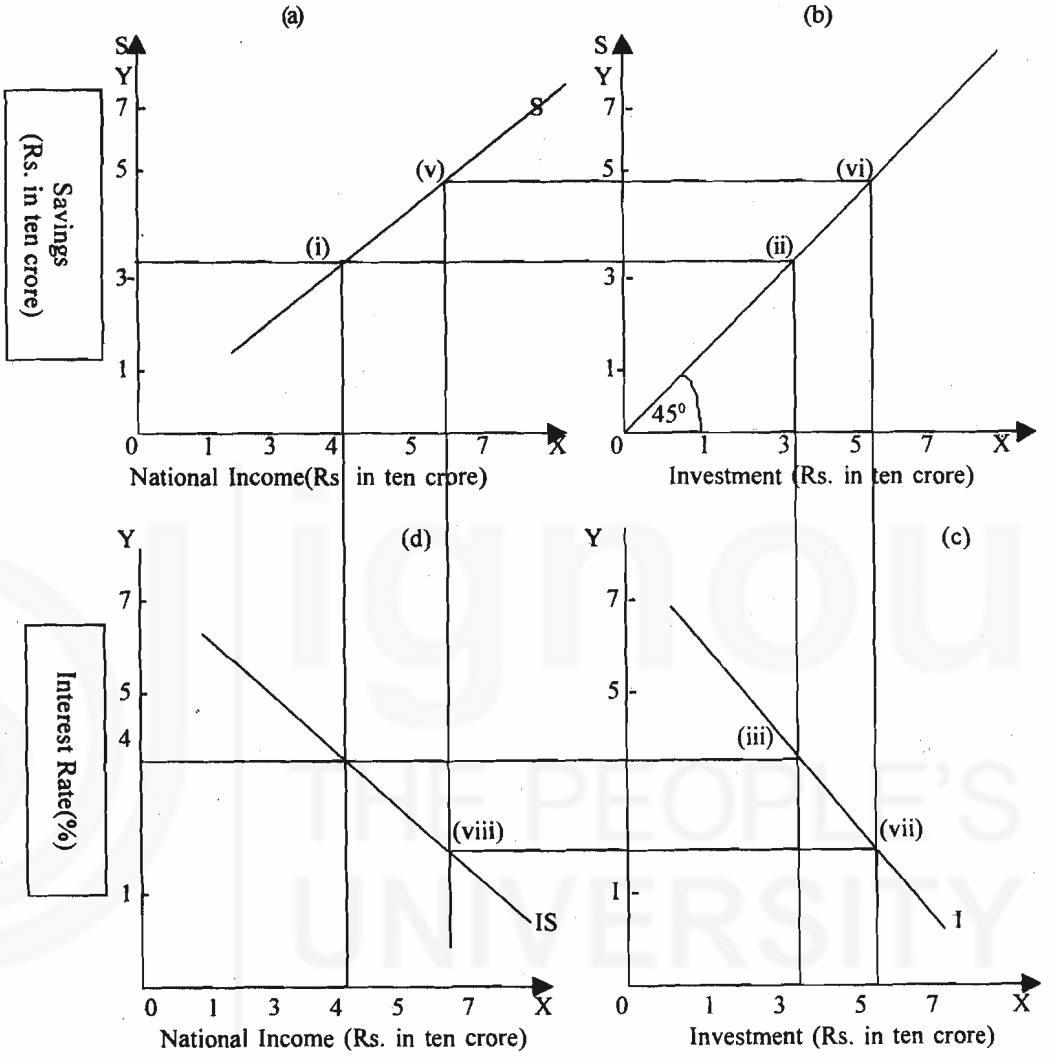
Explanation of Fig. 15.16

In this diagram part (a) shows the direct relationship between savings and national income where these variables are represented on Y-axis and X-axis respectively. It can be seen that at point (i) savings are Rs. 30 crore at national income equal to Rs. 4000 crore. In part (b) while savings are represented on Y-axis, investment is measured on X-axis. For equilibrium, which is possible only when we draw 45° line from the origin such that savings of Rs. 30 crore are equated to investment of Rs. 30 crore at point (ii). In part (c) investment is represented on X-axis and interest rate on Y-axis. Recall that interest rate determines investment as explained in Section 15.4.4. For investment of Rs. 30 crore to take place there has to exist interest rate of 4% as shown by point (iii).

Finally, in part (d) interest rate is shown on Y-axis and national income on X-axis. Interest rate of 4% derived from part (c) and national income of part (a) equal to Rs.4000 crores are indicated at point (iv) and at the same time savings and investment are equated. Thus, point (iv) in part (d) is a point on IS curve which indicates equality

of savings with investment. We have got one point on the IS curve but still we do not know the shape of IS curve i.e., whether it is upward sloping or downward sloping, or parallel to either X-axis or Y-axis.

Fig. 15.16



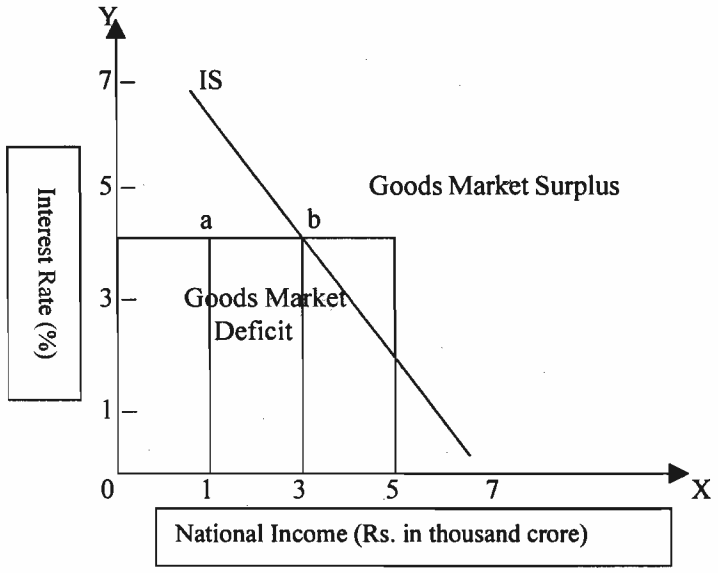
For knowing the exact shape of IS curve we have to derive another point on it. This curve which can be done by identifying points (v), (yi), (vii) and (viii), at another level of national income say Rs.5000 crore in part (a), (b), (c) and (d) respectively.

We have been able to get points (iv) and (viii) in part (d) on IS curve. Similarly we can get a series of other points on IS curve and by joining all such points get the shape of the IS curve. We notice that it is downward sloping and formed by a locus of points at which various combinations of interest rate and national income meet. Also at the same time it coincides with the points of equality of savings and investment in an economy. Throughout the IS curve goods market is in equilibrium.

15.6.2 Deficit and Surplus in Goods Market using IS curve

Let us draw an IS curve to identify deficit and surplus in goods market. This is done in Fig. 15.17.

Fig. 15.17



The slope of an IS curve depends on the slopes of saving and investment curves. The steeper the saving curves, the steeper would be the IS curve given the slope of the investment curve. Similarly steeper the investment curve steeper would be the IS curve given saving curve.

Let us study the position of IS curve at point a in Fig. 15.17. We already know from Fig. 15.16 that along the IS curve throughout I is equal to S at various combinations of interest rate and national income. At point a interest rate is 4% and national income is Rs. 5000 crore. But for goods market to be in equilibrium at interest rate 4% national income required is Rs. 3000 crore. Obviously, at point a since national income required for goods market to be in equilibrium is Rs. 3000 crore, higher national income means higher savings, more than what is required for it to become equal to investment. Since interest rate is still 4% and investment level is unchanged, savings have become more than investment. Saving more than investment means aggregate supply is more than aggregate demand and hence there is surplus in goods market. In fact any point above IS means surplus in goods market.

By the same logic when savings are less than investment, there would be deficit in goods market since aggregate demand is more than aggregate supply. In fact any point below the IS curve indicates a position of shortage in goods market. It is only on the IS curve that there is equilibrium.

15.6.3 Shift of IS Curve

If we study Fig. 15.16 and shift S curves in part (a) to the left it implies more savings at each level of national income. Similarly if S curve is shifted to the right, IS curve will shift to the right, assuming of course, that I curve remains unchanged.

Similarly in part (c) if we shift I curve to the right implying more investment at all levels of interest rate we would notice that IS curve shifts to the right. By the same logic if I curve shifts to the left, IS curve will also shift to the left assuming S curve remains unchanged.

Thus, in terms of IS curve mechanism, a rise in investment is indicated by a shift of the IS curve to the right so that at a given interest rate national income would rise and the extent of rise will be given by

$$\Delta I \times \frac{1}{1 - MPC}$$

where ΔI is an increase in investment and MPC is marginal propensity to consume. We already know 1/1-MPC is the value of the investment multiplier. Similarly a fall in investment is indicated by a shift of IS curve to the left implying a fall of national income at a given interest rate which is given by

$$\Delta I \times \frac{1}{1 - MPC}$$

Check Your Progress 3

1) What is the difference between S and IS curves?

.....

.....

.....

.....

2) When will IS curve be parallel to Y-axis?

.....

.....

.....

.....

3) What exactly is meant by deficit and surplus in goods market?

.....

.....

.....

.....

15.7 LET US SUM UP

In this unit we have given you an account of determination of equilibrium level of national income. This is done by first assuming that the economy consists of only two kinds of spending units, households and firms.

Concepts of injections and leakages in income flow have been introduced to get equilibrium level of income in goods market by assuming non-existence of money market.

Aggregate demand and aggregate supply concepts are introduced and linked with injections and leakages.

A separate discussion has been given on consumption and investment functions and these are related to 45° line to get equilibrium level of income.

At a later stage government and rest of the world transactions are introduced to get equilibrium level of income. Changes in equilibrium level of income are discussed subsequently. In this context, concepts of investment multiplier, government expenditure multiplier, tax multiplier, balanced budget multiplier and export multiplier are also introduced.

Lastly, the concept of fourth part diagram is introduced to derive IS curve. Concepts of deficit and surplus in goods market are highlighted with the help of IS technique. Determinants of the shape of IS curve and factors responsible for the shifts in IS curve are also discussed in this unit.

15.8 KEY WORDS

Assumptions : Whenever we formulate a theory we take, initially, certain variables and situations as given, to simplify our understanding of theory, and these are known as assumptions. Such assumptions can be relaxed subsequently to make a theory more realistic.

Autonomous Expenditure : It is an expenditure of household, or government consumption or investment or export, which is not linked to level of income.

Aggregate Demand : It is the demand for goods and services for the economy as a whole. This consists of demand for households, government, firms and rest of the world. The factors determining each of these components of aggregate demand are different.

Aggregate Supply : It is the utilisation of potential productive capacity of an economy given the level of factors of production.

Average Propensity to Consume : It is defined as C/Y where C is the level of consumption expenditure and Y is the level of national income.

Autonomous Investment : It is the level of investment, which takes place irrespective of the level of income.

Subsistence Level of Consumption : It is the level of consumption expenditure which takes place when level of income is zero.

Balanced Budget Multiplier: It is defined as $\Delta Y/\Delta G$ ($=\Delta T$) where ΔG is change in government expenditure which is defined as equal to change in tax revenue ΔT and ΔY is change in level of national income. If tax is lump sum the value $\Delta Y/\Delta G$ ($=\Delta T$) = 1

Consumer Goods : These are the goods demanded by households to satisfy their wants directly.

Consumption Function : It gives the relationship between level of consumption expenditure and level of income, or $C = f(Y)$, where C is consumption, Y is income and f represents the functional relationship. A simple form of consumption function is $C = A+cY$, where C is consumption, A is level of consumption when income is zero and c is marginal propensity to consume and Y is national income.

Dependent Variable : It is a variable the value of which depends on how an independent variable behaves. For example, when $C=f(Y)$, Y is an independent variable and C is consumption dependent on Y.

Deficit in Goods Market : It represents a situation when aggregate demand of an economy is more than aggregate supply, given the price level.

Equilibrium Level of Income : It is that level of income at which injections are equal to leakages or where aggregate demand is equal to aggregate supply.

Expected Returns : These are the revenues expected to be received over the expected life of a machine by selling the products produced with the help of investment in a machine.

Exports : It is that part of domestic production, which is bought by rest of the world.

Equilibrium in Goods Market: It is a situation where aggregate demand is equal to aggregate supply or where injections are equal to leakages in the income flow.

Exchange Rate : It is that rate at which domestic currency of a country is exchanged with a foreign currency.

Excess Capacity : It is a situation where the installed capacity of machines is not actually utilised. It happens primarily because there is lack of effective demand or where there is lack of complementary resources required to utilise an installed capacity of the machine.

Four Part Diagram : It is a combination of four diagrams where the independent variable used in one diagram is utilised in another to get equilibrium in an economy.

Full Employment level of Income : It is that level of income generated by an economy by producing goods and services at which resources are fully utilised.

Factor Market : It is a market in which factors of production, land, labour, capital and enterprise are demanded and supplied at a given price.

Firms : These are the units that produce goods and services.

Government Consumption Expenditure : It is the expenditure incurred by general government to produce and collectively consume goods and services. It consists of compensation of employees and net purchase of intermediate goods bought from firms.

Government Expenditure Multiplier : It is defined as $\Delta Y/\Delta G$ where ΔY is change in national income and ΔG is change in government expenditure. The value of this equals 1/1-MPC where MPC is marginal propensity to consume.

Goods Market : It is the market where aggregate demand and aggregate supply of goods and services are recorded.

Households : These are the units which supply factors services and get factor income to utilise it to buy goods and services produced by firms or to save.

Income Expenditure Line : It is 45° line at which income measured on X-axis equals expenditures incurred by various transactions recorded on Y-axis.

Induced Investment : It is that level of investment, which is dependent upon level of income or change in level of income.

Intended savings : These are the savings, which are functionally related to income.

Investment Multiplier : It is defined as $\Delta Y/\Delta I$ where ΔY is change in national income and ΔI is change in investment. It is equal to 1/1-MPC, where MPC is marginal propensity to consume.

Induced Expenditure : It is the consumption and investment expenditure, which is related to level or change in level of income.

Interest rate : It is the difference between forward price and spot price of a given sum of money expressed as a percentage of spot price.

Investment goods : These are the goods used to generate a further flow of goods and services in an economy.

Independent Variables : These are the variables, taken for granted or, are independently determined.

Injections : These are the increases in magnitude of autonomous variables, which generates income flow.

IS Curve : It is a locus of points at which savings are equal to investment at various combinations of interest rate and national income.

Investment Function : This gives a functional relationship between investment and the variables, mostly interest rate given the marginal efficiency of capital on which it is dependent.

Import Function : It is represented as M_a+mY, where M is imports M_a is level of imports irrespective of level of income and m is marginal propensity to import ($\Delta M/\Delta Y$) and Y is national income.

Linear Relationship : It is the relationship between independent and dependent variables such that we are able to get straight line mapping of points.

45° Line : It is a line drawn from origin with an angle of 45°. At any point on this line the perpendicular drawn on X-axis is equal to the one drawn on Y-axis.

Leakage : These are the factors, which lead to slowing down of the income change process. Saving, import, tax are some such examples.

Money Market : It is a market in which there is demand for and supply of money.

Marginal Efficiency of Capital : It is that critical rate of discount at which the discounted value of expected returns becomes equal to supply price of capital.

Marginal Propensity to Consume : It is defined as $\Delta C/\Delta Y$, where ΔC is change in consumption and ΔY is change in income.

Net Foreign Demand : It is the difference between exports and imports.

Multiplier : It is the value by which a change in autonomous variable should be multiplied to get change in level of national income or some other macro aggregates.

Marginal Propensity to Import : It is defined as $\Delta M/\Delta Y$ where ΔM is change in imports and ΔY is change in national income.

Present Value : It is the value of variable at present, which is derived by discounting its expected value over future periods.

Product Market : It is a market where goods and services are traded according to demand and supply.

Productive Capacity : It is the capacity of the economy to generate the flow of goods and services given the factors of production and technology.

Real Market : It is another name of goods market.

Short-run Analysis : An analysis where certain variables operating in the long run are ignored.

Savings : It is defined as Y-C where Y is national income and C is level of consumption.

Saving-Investment Equality : It is a situation of equilibrium in the goods or real market where intended savings are equal to intended investment.

Surplus in Goods Market : It is a situation where aggregate supply is more than aggregate demand for an economy.

Tax : It is a compulsory payment made by various transactors of an economy to the government.

Tax Rate : It is a percentage of income, which has to be paid to the government in the form of tax.

15.8 SOME USEFUL BOOKS

Shapiro, Edward, 1984, *Macro Economics Analysis* (5th Edition), Galgotia Publications: New Delhi (Chapters 4-8)

Ackley, Gardner, 1987, *Macroeconomics: Theory and Policy*, Macmillan Publishing Co., INc. New York (chapters 6-12).

Branson, William, H, 1979, *Macro Economic Theory and Policy*, (2nd Edition), Harper and Row Publishers, New York (Chapters 3,10,11,15)

15.9 ANSWERS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) i) It is only at a point where aggregate demand curve intersects the 45° line. We get equilibrium level of income because it is only at this point where intended investment is equal to intended savings. Any point away from this will lead to unintended increase or decrease in inventories, which will force national income move back to equilibrium level.

ii) At point 'a' investment is bc and savings are ac. Thus, savings are more than investment, which implies consumer goods produced are more than their demand which leads to unintended increase in stock and therefore, in the next round production will fall. So factor incomes and national income will fall until it reaches od' level at which savings are equal to investment. This of course assumes that price level remain unchanged.

iii) The shape of consumption curve will be as shown Figure 15.2.

Consumption curve will pass through origin O and marginal propensity is given by the angle COX, that is, constant throughout.

2) Using figure in 1 (iii) we have seen marginal propensity to consume is given by angle COX and average propensity to consume is also given by angle COX. Both marginal and average propensities to consume are equal.

3) Read Sub-section 15.4.5 and write your answer.

Check Your Progress 2

1) i) $\Delta Y = \Delta DI \times \frac{1}{1 - MPC}$

$= 100 \times \frac{1}{1-.33}$

$= 100 \times \frac{1}{1-.66}$

$= \frac{100 \times 100}{66} = \frac{10000}{66}$ = Rs. 151.5 crore

ii) National income will still rise by Rs. 151.5 crore

iii) $\frac{\Delta Y}{\Delta G(=\Delta T)} = 1$

or, $\Delta Y = \Delta G (=\Delta T) \times 1$

ΔY = Rs. 100 crore

National income will increase by Rs.100 crore

2) $\Delta Y/\Delta E$ is export multiplier that is given by

$$\frac{1}{1-MPC+m}$$

and on the other hand investment multiplier $\Delta Y/\Delta I$ is given

by $\frac{1}{1-MPC}$

and therefore the difference .

3) $Y = C+I+G$

since $C = c\,[Y+Tr]$,

$Y = c\,[(Y+Tr)] +I+G$

or, $Y = cY+cTr+I+G$

or, $Y-cY = cTr+I+G$

or, $Y(1-c) = cTr+I+G$

or, $Y = \frac{1}{1-c}\,[cTr+I+G]$

Increase T i.e. transfers by ΔTr leads to increase in national income by ΔY.

$$Y+\Delta Y = \frac{1}{1-c}\,[c\,(Tr+\Delta Tr) +I+G]$$

$$\text{or, } Y+\Delta Y = \frac{1}{1-c}\,[cTr+I+G] + \frac{c\Delta Tr}{1-c}$$

since $Y = \frac{1}{1-m}\,[cTr+I+G]$

we can cancel them from each side

$$\text{or } \Delta Y = \frac{c}{1-c}\,\Delta Tr \text{ or, } \frac{\Delta Y}{\Delta Tr} = \frac{c}{1-c}$$

Check Your Progress 3

1) S curve gives the relationship between savings and national income. IS curve on the other hand, gives us points where savings are equal to investment at various combinations of interest rate and national income.

2) IS curve will be parallel to Y-axis if investment is insensitive to interest rate.

3) Surplus in goods market means a situation where savings are more than investment or where aggregate supply is more than aggregate demand. Similarly deficit in goods market means a situation where savings are less than investment or where aggregate supply is less than aggregate demand.

UNIT 16 EQUILIBRIUM IN MONEY MARKET

Structure

16.0 OBJECTIVES

On going through this unit you would be able to explain:

- exact meaning of money supply;
- the reasons for which money is demanded;
- meaning of money market;
- meaning of equilibrium in money market;
- what is LM curve;
- determinants of the slope of LM curve; and
- factors responsible for shift in LM curve.

16.1 INTRODUCTION

This unit relates to money market. The need to study money market arises from the fact that this market along with goods market determines the equilibrium level of income and interest rate. We known that introduction of money is necessitated by the difficulties encountered in the operation of barter system of exchange. You would be introduced to the nature of supply of money and demand for money to help in understanding the equilibrium in money market. Since money market can be in equilibrium at various combinations of interest rates and national income, LM curve will be introduced along with the factors determining its slope as well as shift.

16.2 ROLE OF MONEY

Money is anything, which is generally acceptable as a means of payment in the settlement of transactions. It is commonly used as a medium of exchange or a means of transferring purchasing power. In absence of money, people exchange goods for goods (See Unit 18 - Section 18.1). A more detailed treatment of problems of barter system, evolution of money and demand for money is attempted in Unit 18, Block 8.

16.3 SUPPLY OF MONEY

Supply of money is a stock, which can be measured at a point of time. It is taken to be autonomously determined.

Measures of Money Supply

Once we have settled on a theoretical definition of money, we can identify empirically the things that serve as money in an economy. Then, the total stock of money of various kinds at a particular point of time can be computed. By repeated measurements at different points of time, a whole time series of money supply can be constructed. This will show the time behaviour of money supply. Coupled with other data and helped by theory, this information can be used to throw light on the effect of changes in the supply of money on several key variables such as income, prices, wages, employment, rate of interest and balance of payments, and how to control changes in the supply of money to attain certain policy goals.

At the outset, we must note two things about any measure of money supply. First, supply of money refers to its stock at any point of time. This is because money is a stock variable in contrast with a flow variable, such as real income, which refers to its rate per unit of time (say, per year). It is the change in stock of money, say, per year, which is a flow.

Second, stock of money always refers to the stock held by the public. This is always smaller than the total stock of money in existence. The term public is defined to include all economic units, households, firms and institutions, except the producers of money, the government and the banking system. For the most common definition of money, the government means the Central Government plus all state governments; the banking system means the RBI plus all banks, which accept demand deposits. This means that the word public is inclusive of all local authorities, non-bank financial institutions, and non-departmental public-sector undertakings (such as Hindusthan Steel, Indian Airlines, etc.) and even the foreign central banks and governments and the International Monetary Fund who hold a part of Indian money in India in the form of deposits with the RBI. In other words, in the standard measures of money supply, money held by the government and the banking system is not included.

The primary reason for measuring stock of money in this way is that this separates producers or suppliers of money from holders or demanders for it. For both monetary analysis and policy formulation, such a separation is essential.

The measurement of money supply is an empirical matter. We study various measures of money supply published by the RBI. Till 1967-68 the RBI used to publish only a single measure of money supply (M) defined as a sum of currency and demand deposits, both held by the public. Following that convention, now we call it the narrow measure of money supply. From 1967-68 the RBI started publishing additionally a broader measure of money supply, called 'aggregate monetary resources' (AMR). It was defined empirically as money narrowly defined plus the time deposits of banks held by the public. From April 1977 yet another change was introduced. Since then the RBI has been publishing data on four alternative measures of money supply in place of earlier two. The new measures are denoted by M_1, M_2, M_3 and M_4. The two earlier measures were represented by M and AMR. The respective empirical definitions of these measures are given below :

M or M_1 = C+DD+OD

M_2 = M_1 + savings deposits with post office saving banks,

AMR or M_3 = M_1 + net time deposits of banks,

M_4 = M_3 + total deposits with the Post Office Savings Organization (excluding National Savings Certificates)

In the above definitions,

C = currency held by the public,

DD = net demand deposits of banks,

OD = 'other deposits' of the RBI.

Contents of each of the components of M_1 to M_4 are explained briefly below.

Currency consists of paper currency as well as coins. Paper currency is predominant in the form of Reserve Bank of India currency notes of the denomination of Rs.2 and above (rupees five, ten, twenty, fifty, one hundred and five hundred notes). In addition, we also have small amounts of Government of India rupee one notes. Though made of paper, they are counted as rupee one coins. Together with rupee one coins and other small coins, they constitute the small-coins component of money supply. They are direct monetary liability of the Government of India but put into circulation by the RBI as the agent of the Central Government. The RBI does this by holding stocks of government currency on the one hand and by maintaining full convertibility of the small coin compare currency into the rest of the country's currency and vice versa.

Items included in any measure of money supply are the net demand deposits of banks. Net instead of total demand deposits of banks is because we have defined money and any one of its components as something held by the public only and total deposits include both deposits from the public and inter-bank deposits. The latter are deposits which one bank holds with others. Since they are not held by the public, they are netted out of the total demand deposits to arrive at net demand deposits. We may remind you that demand deposits comprise the current-account deposits and the demand deposit portion of savings deposits, which can be withdrawn by cheque .

'Other deposits' of the RBI are its deposits other than those held by the government (the Central and state governments), banks, and a few others. They include demand deposits of quasi-government institutions (like the IDBI), foreign central banks and governments, the IMF and the World Bank etc. Empirically, whatever the measure of money supply, these 'other deposits' of the RBI constitute a very small proportion, less than one per cent, of the total money supply. Therefore, no harm will be done, if in our future discussion we ignore these other deposits.

The following additional points about the new measures of money supply vis-a-vis the old need to be noted.

1) M_1 is only a revised measure of M, RBI's old measure of money supply. The revision is not conceptual, but only in terms of coverage. The new series gives a better coverage of the co-operative banking sector. Formerly, only the demand liabilities of the State Co-operative Banks were included in money supply. Other tiers of the co-operative banking sector were neglected on account of the non-availability of data. In the new series, net (i.e., excluding inter-bank) demand deposits of State Co-operative Banks, Central

Co-operative Banks and a segment of Primary Co-operative Banks consisting of (i) urban co-operative banks and (ii) salary earners' credit societies are included. Similarly, M_3 is the revised version of the series on AMR with extended coverage for the co-operative banking sector.

2) The new series M_2 and M_4 have been devised to accommodate the Post Office deposits.

3) The RBI views the four new measures of money stock to represent different degrees of liquidity. It has specified them in the descending order of liquidity, M_1 being the most liquid and M_4 the least liquid of the four.

Which of the alternative measures of money supply to choose and why? We cannot attempt an answer here, as it will involve going into questions of monetary theory, policy, and empirical testing. It should suffice to say at this stage that the most common measure of money supply is that provided by M or M_1.

Till 1978 the RBI also used to concentrate most of its accounting analysis on this narrow measure of money supply. But things have changed since. Due to the introduction of a change in 1978 in the division of saving deposits of banks as between demand deposits and time deposits the data on M for post-1978 years are no longer comparable with those for previous years. This has compelled the RBI to conduct its accounting analysis of changes in money supply in terms of M_3.

We do not have any basis to either explain the sources of increase in M (or M_1) or AMR (or M_3) or to evaluate such increases as socially beneficial or injurious. But they are very important questions of monetary theory and policy.

16.4 DEMAND FOR MONEY

To be able to analyse the effects of changes in the stock of money, it is essential to study the equilibrium of money market. Money is an asset of the holding public. As such, it must have a demand for and a supply of money. The demand for money comes from the general public. The supply of money comes from its producers, the government and the banking system. The money market comprises these demanders and suppliers.

The supply of money is autonomously given by the monetary authority. The demand for money, which is the keystone of monetary theory, needs to be carefully understood.

Money is a stock variable. The stock refers to the quantity money at a point of time. As an asset, the demand for money is the public's demand to hold money, whatever be the motive for holding. The length of the time period for which money is held is not an essential consideration. Holding cash in one's pocket for spending is as much a part of the demand for money as burying currency notes in a pot under ground.

Monetary analysis is aggregative. Therefore the demand for money relevant to this analysis is also the aggregate demand. It is the sum of all the money demanded by individual members of the public, whether households or firms.

Theory of demand for money is mainly concerned with answering the question like what are the determinants of public's demand for money? A related question is, why does the public demand money? Several explanations have been offered in reply. Each explanation has its own implications for explaining further the economic effects of changing money supply. The theory of the demand for

money has been widely tested empirically. As a consequence, much empirical evidence has accumulated around it.

Nominal versus Real Cash Balances

At the outset it is important to introduce the distinction between nominal and real cash balances. Cash balances are another term for money. Nominal cash balances are money of the current purchasing power of a unit of money (say, a rupee). Real cash balances are money of some base-year purchasing power. A nominal rupee is nominally always a rupee. But its purchasing power in terms of real goods and services can vary from time to time with changes in the general price level. Then, it is said that the real value (purchasing power) of a (nominal) rupee has been changing over time. For making real-value comparisons, first a reference point or a base year is chosen. Suppose we choose 1991-92 as our base year with the wholesale price index number = 100. A rupee in this year had certain amount of purchasing power at the prices prevailing during the year. Suppose the average value of this index number for the year 2000-2001 was 400. Then, the purchasing power or value of a nominal rupee at the 2000-2001 prices was only one fourth of what it was at the 1991-92 prices, that is, one rupee in could buy only as much goods and services as only 25 paise could buy in 1991-92. The intervening four-fold rise in prices has reduced the real value of a rupee to that of a quarter. If some other year (say, 1961-62) with a lower price level than that of 1991-92 were chosen, the real value of a nominal rupee in 2000-2001 would turn out to be lower still in comparison with its real value in the new base year. Thus, technically, real cash balances are given by nominal cash balances deflated by the price level. That is, if M and P are used to denote nominal money and price level respectively, then the real cash balances will be given by M / P. The latter will be measured in terms of the purchasing power of money in the year, which serves as the base year for the price index number measuring P.

Whenever changes take place in P, the distinction between nominal and real cash balances becomes relevant. To analyse the demand for money, we have to decide whether what the public wants to hold is some nominal quantity of money or some real quantity of it. This is one way of bringing in P as one of the determinants of the nominal demand for money. J.M. Keynes had identified three different motives for holding money or cash. He had termed them:

i) Transactionary motive,

ii) Precautionary motive, and

iii) Speculative motive.

Each of these motives gives rise to one type of demand for money. Thus, transactionary demand for money arises because people want to keep some cash with them to meet their day to day needs to purchase commodities etc. in the market. This demand for money will be dependent on the level of income. The precautionary demand for money arises as people want to keep some additional amount in cash to meet unforeseen emergencies. This amount will also depend upon level of income. These two components constitute,

$$M_t^d = M_t^d (Y)$$

The speculative motive gives rise to another component of demand for money. Some people keep some more cash to make use of opportunities that may arise in the market. This is a type of business objective and have r, the rate of interest becomes main determinant of the demand. Thus,

$$M^d_{sp} = M^d_{sp}(r) .$$

The overall demand for money Keynes had defined will therefore be :

$$M^d = M_t^d + M^d_{sp} = M^d$$

In other words, total demand for money is sum of transactionary, precautionary and speculative demands for money. While the first two components depend on the level of income and the last one on rate of interest, the overall demand for money depends on both income and rate of interest. We have discussed different aspects of demand for money in Unit 18 - in Sections 18.3, 18.4 and 18.5.

Check Your Progress 1

1) How does the introduction of money facilitate the exchange process in an economy?

.....

.....

.....

2) Explain M_1 and M_3 measures of money supply.

.....

.....

.....

16.5 MONEY MARKET

The introduction of money market along with goods or real market helps arrive at equilibrium of an economy. Therefore, it is essential for us to discuss the equilibrium in money market separately. Before we talk of equilibrium in money market, it is better to be clear about its meaning.

16.5.1 Meaning of Money Market

Money market like any other, say, commodity or factor market has the forces of demand and supply operating to determine the equilibrium. In this market as well quantity demanded becomes equal to amount supplied.

The role of money has been undergoing a change as visualised by different economists. The classical economists like Irving Fisher and AC Pigou thought of money performing the role of meeting transactions to buy goods and services. Therefore the operation of money market was supposed to have no impact on goods or real market. It was J.M. Keynes who emphasized that money can be demanded not only to meet day to day transactions, but also as the most liquid asset. Therefore, there is demand for money and given supply of money we can find the equilibrium in money market. The price for parting with money is termed as interest rate. Thus, when demand for money becomes equal to supply of money equilibrium interest rate is determined.

The forces working in the money market to determine interest rate can be shown with the help of the Fig. 16.1.

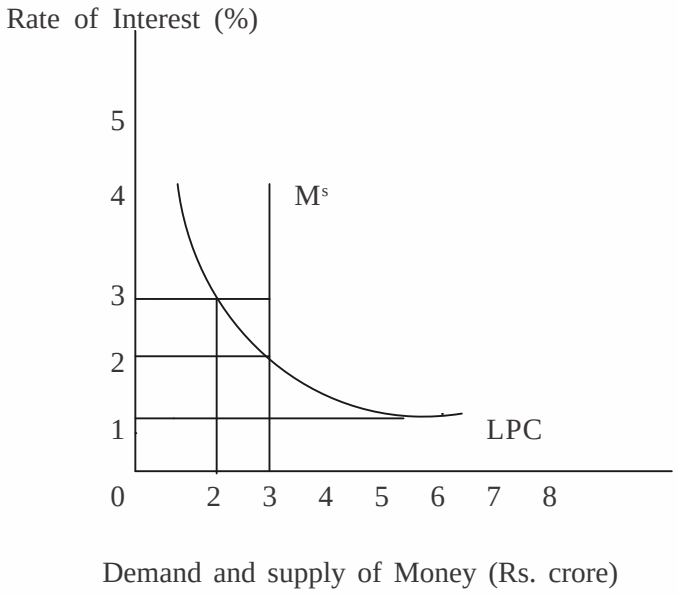


Fig. 16.1: Shows the forces working in the money market. The vertical straight line M^s shows the money supply determined by the monetary authorities. The speculative demand for money is shown by liquidity preference curve LPC. Money market equilibrium occurs at the point where demand equals supply.

16.5.2 Equilibrium in Money Market

In Fig. 16.1 demand and supply of money for speculative purposes (in crores of Rs.) is shown on X-axis and interest rate (in %) is shown on Y-axis. At interest rate 3% demand of money is Rs.2 crore and supply of money is Rs.3 crore which will push interest rate down and this process will continue until interest rate is 2% at which demand of money is equal to supply of money. On the other hand at interest rate of 1% demand of money is Rs.5 crore and supply of money is Rs.3 crore which will pull the interest rate up until it becomes equal to 2% at which again demand for money is equal to supply of money. The equilibrium in money market is a situation in which interest rate prevailing is 2% and demand and supply of money is Rs. 3 crore.

In Fig. 16.1 we have taken demand and supply of money for speculation purposes alone. If we want to consider total demand and supply of money we will have to incorporate demand of money for transactionary and precautionary purposes along with that of speculative purposes. This incorporation of the two is shown with the help of Fig. 16.2.

You already know that transactionary and precautionary demand of money (M_t) is an increasing function of income (Y) and speculative demand of money (M_{sp}) is a decreasing function of interest rate (r). Combining M_t and M_{sp} we get total demand of money (M^d). Supply of money (M^s) is independent of interest rate (r). Demand and supply of money is shown in Fig. 16.2.

Fig.16.2

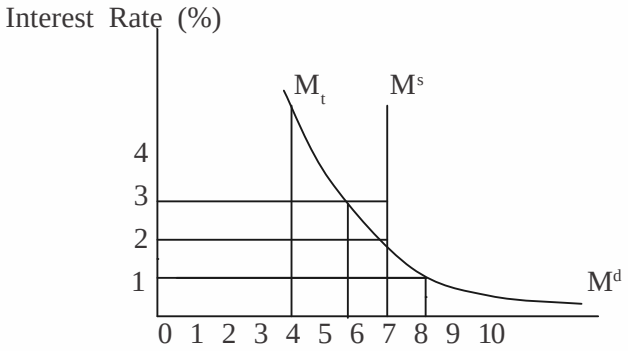


Fig. 16.2: Shows the two components of demand for money separately. A vertical line shows M_t the transactionary demand, as it is independent of the rate of interest. The other component, the speculative demand, is shown as a downward sloping curve beginning at the M_t line. Intersection with the policy-determined money supply, M^s, gives the equilibrium.

In Fig. 16.2, M_t i.e., Rs.4 crore is shown as vertical line M_1 which is parallel to Y-axis. It is independent of r and from Rs.4 crore onwards it slopes downwards to indicate speculative demand of money (M_{sp}). If M_t and M_{sp} are combined we get total demand for money which is indicated by M^d curve. Supply of money is taken to be Rs.7 crore such that at 2% interest rate M^d becomes equal to M^s. At rate of interest 3% M^s is Rs7 crore and M^d is Rs6 crore forcing r to come down to 2%. Similarly, at interest rate 1% M^s is Rs.7 crore and M^d is Rs 8 crore which forces r to go up to 2%. It is finally at r equals to 2% that $M^d = M^s$ = Rs.7 crore which represents the equilibrium in the money market.

16.5.3 LM Curve and its Derivation

We are now in a position to introduce LM curve i.e. a curve at which $M^d = M^s$ or speculative demand for money (M_{sp}) is equal to supply of money available for speculation (M^s_{sp}) at the various combinations of r and Y. We would employ the technique of four-part diagram introduced in Unit 15 to derive the LM curve. The shape of LM curve is derived in Fig. 16.3.

Fig.6.3

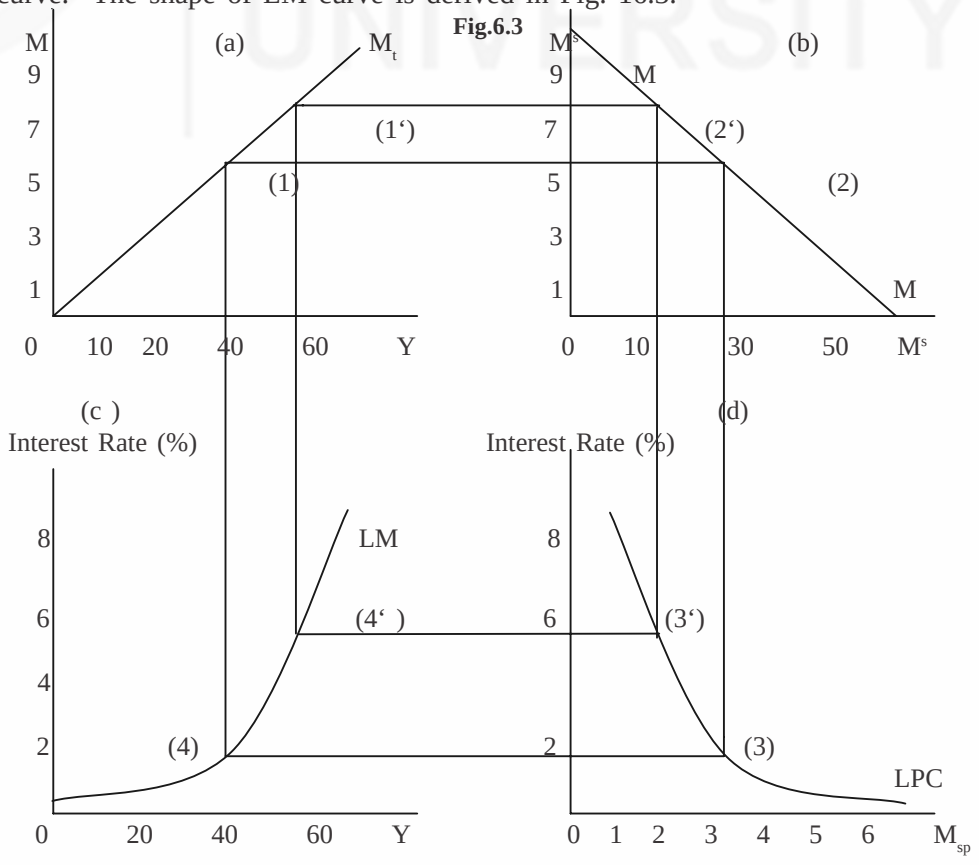


Fig.16.3: Shows derivation of the LM curve. Part (a) shows transactions demand for money M_t and both the axes in part (b) show the supply of money M^s. In part (c), M_{sp} is the speculative demand for money. M_{sp} should be equal to money available for speculative M^s_{sp} purposes, determined in part (b) to have money market in equilibrium. In part (d), we have brought togather the levels of income Y and rates of interest r associated with such points of equilibrium to form the LM curve .

Fig. 16.3 has four parts (a), (b), (c) and (d). In part (a) transaction plus precautionary demand for money (M_t) is shown on Y-axis and income on X-axis. When income, Y, is 40, M_t is 5 as shown by point (1). In part (b) money supply (M^s) is shown on X-axis as well as on Y-axis. Money supply (M^s) is taken to be 8. Point (2) in part (b) indicates that given M_t equals 5, out of $M^s = 8$, 3 is left by way of supply of money available for speculation (M^s_{sp}). In part (c) speculative demand for money (M_{sp}) is represented on X-axis and interest rate (r) on Y-axis. M^s_{sp} of 3 in part (b) must be equalized to M_{sp} for equilibrium in money market. Since M_{sp} is 3 in part (b) it is possible as shown in part (c) at r equals to 2%. This is down by point (3) in part (c). Finally, in part (d) interest rate (r) is represented on Y-axis and income (Y) on X-axis. We started with Y equals 40 in part (a), which gave us M_t equal to 5. Out of M^s equals to 8, 5 is M_t which leads to 3 being left by way of supply of money available for speculation. An equality of M^s_{sp} and M_{sp} at 3 to attain equilibrium in money market, requires to be r equal to 2%. Part (d) of Fig. 16.3 shows that such an equilibrium is possible only when Y equals 40 and r equals 2%. This combination of Y = 40 and r = 2% is shown by point (4) in part (d). Point (4) is on the LM curve as at this point demand for money equals supply of money.

To get another point on the LM curve start with Y = 50 which gives M_t = 7. This in turn leaves M^s_{sp} = 1, which must be equalized to M_{sp} = 1. Such a condition is satisfied at r = 6%. At r = 6% and Y = 50 we have an equilibrium in money market which is indicated by point (4'). Thus, point (1') of part (a) helps us to arrive at point (2') of part (b) which, in turn, helps us to arrive at point (3') of part (c) and finally point (1') of part (a) and point (3') of part (c) together, help us to arrive at point (4') of part (d).

Just like point (4) and (4‘), we can arrive at a number of other points on LM curve at different combinations of r and Y. Joining of such points we get LM curve.

It can be easily seen that LM curve in its initial stage has been shown as parallel to X-axis and subsequently as rising from left to right. The initial parallel shape of LM curve is a direct consequence of liquidity preference curve (LPC) of part (c) being shown parallel to X-axis at a very low interest rate.

16.5.4 Slope and Shift of LM Curve

After we have learnt the derivation of LM curve in Section 16.5.3 let us study the factors on which slope of LM curve depends. Two such factors are slope of M_t curve in part (a) and slope of LPC in part (c) of Fig. 16.3 This we will show with the help of Fig. 16.4.

In Fig. 16.4 four-part approach has been adopted. M_t, MM, LPC and LM are the curves, which have been introduced in Figure 16.3. Suppose a new M_t curve M'_t is introduced in part (a) of Fig. 16.4. M'_t curve is steeper than M_t curve showing that a greater proportion of Y is demanded for transaction and precaution purposes. Let us see the effect of this on LM curve. At Y = 40 in Fig.16.3, M_t was 5, which was shown by point (1). Now it is 7 and is shown by point (1^1). At M^1_t equal to 7, 1 is supply of money available for speculation purposes as shown by point (2^1) in part (b).

Fig.16.4

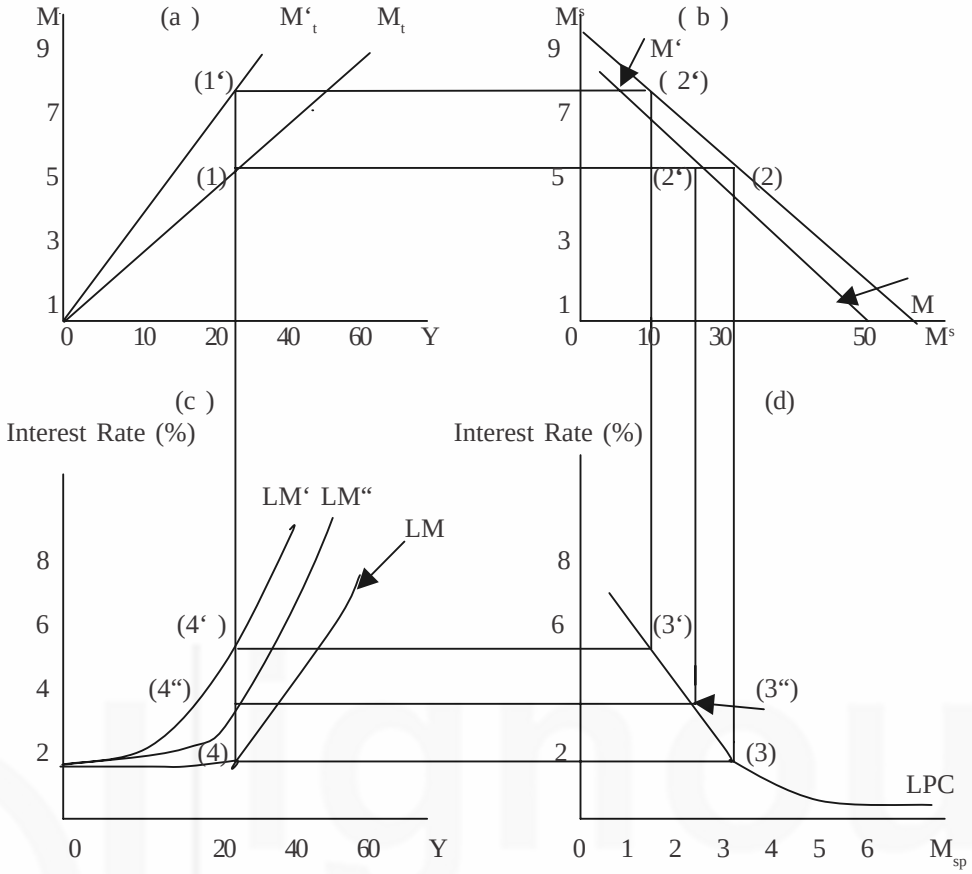


Fig.16.4: Highlights the factors that determine slope of the LM curve. As the proportion of money supply needed for transactions and precautionary purposes rise, lesser amount remains to be available for speculative activities. Thus, at the same level of income, money market equilibrium will occur at a higher rate of interest only. This makes the new LM curve steeper . If the liquidity preference curve LPC was made steeper, that is, M_{sp} was less responsive to the rate of interest, again the LM will become steeper.

To equalize supply and demand of money for speculation purposes we get point (3^1) in part (c) which indicates 6% interest rate (r). Now in part (d) we get point (4^1) given at r of 6% and Y = 40 Y. By joining all such points we get LM' curve which is steeper than LM. Thus steeper MT curve leads to steeper LM curve.

Similarly, it can be shown that if LPC is made steeper, i.e., M_{sp} becoming less responsive to r, again LM curve will become steeper.

Shift of LM curve

The shift in LM curve is a consequence of the change in money supply. In Fig. 16.4 in part (b) we have drawn a new supply curve of money indicated by M'M curve which shows the supply of money has fallen from 8 to 7. Thus, at M_t equal to 5 in part (a) we are able to get point (2^{11}) in part (b) which gives point (3^{11}) in part (c) indicating r equal to 3. Finally in part (d) we get a point (4^{11}) . If all such points on M'M curve are traced we get a new LM curve indicated by LM" . So, a fall in money supply leads to a shift in LM curve to the left. Similarly if MM curve in part (b) is shifted to the right LM curve will shift to the right. The change in money supply over LPC when it becomes parallel to X-axis will not shift the LM curve. It is because of this reason LM curve even if it shifts to the right or left continues to originate from the parallel portion of the LM curve.

16.5.5 Surplus and Deficit in Money Market

Let us draw a LM curve to show surplus and deficit in money market. This is done in Fig. 16.5.

Fig.16.5

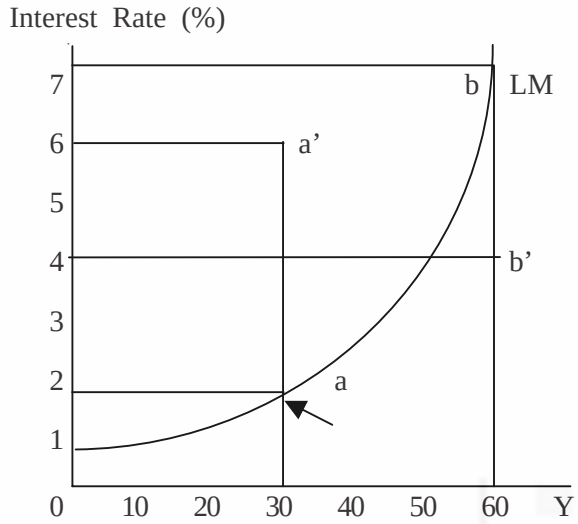


Fig.16.5: focuses on the possibilities of surplus and deficit in money market. A point above the LM curve will represent a higher rate of interest. It leads to squeeze in speculative demand for money and thus emerges a surplus of funds in money market. On the other hand, a point below the LM curve, the reverse will happen. Speculators find the rate of interest too low and their demand for money rises, leading to a dearth of funds in the market.

By definition, there is equilibrium in money market along the LM curve. Consider points a and a'. At a, i.e., with r = 2% and Y = 30 there is an equilibrium in money market. On the other hand, at a', r = 6% , which is more than 2%. At high rate of interest, supply of money is more than demand for money, which shows a situation of surplus in the money market. Therefore, any point above the LM curve represents a situation of surplus in money market. Similarly we can consider points b and b'. At b an equilibrium is established at Y = 60 and r is 7.5%. By definition, at b' demand for money is more than supply. So, point b' represents a situation of deficit in the market. Thus, any point below the LM curve represents a situation of deficit in money market.

16.5.6 Role of LM Curve in Determination of Equilibrium level of National Income

In Unit 15 you have learnt how the level of national income is determined. But, while discussing this only goods market or real market was considered. The working of multipliers - investment, government expenditure, tax, balanced budget and export, was discussed assuming absence of money market. The introduction of money market (i.e., LM curve) along with goods market (or IS curve) makes quite a difference to arrive at new level of national income with a given change in investment, government expenditure, tax or exports.

The role of LM curve in equilibrium level of national income can be illustrated with the help of an example. Suppose there is an increase in government expenditure, which will bring about an increase in national income. That in turn will increase

the demand for money for transaction and precautionary purposes leading to a fall in its supply available for speculation. The fall in supply of money for speculation leads to rise in interest rate given the liquidity preference curve. The rise in interest rate leads to fall in level of investment given the marginal efficiency of capital. The fall in investment results in fall in income. It is worth noting that in the absence of money market an increase in government expenditure would have just led to an increase in national income depending on investment multiplier. The introduction of money market leads to a situation where private investment starts getting reduced, forcing the national income to come down. So, finally national income level does not increase as much as it would have had there been no money market. This is referred to in macro-economics as "crowding out" effect, i.e., expansion of government expenditure causes national income to rise, which , in turn , raises the transactions demand for money leaving out a smaller amount available for speculative purposes. This leads to rise in rate of interest. As a result of change in the rate of interest private investment falls. Thus , a rise in government expenditure crowds out the private investment. So, we can say investment multiplier will be less in a situation where LM curve is introduced along with IS curve. Or, full effect of multiplier will not be felt on national income.

Check Your Progress 2

1) Explain the effect of an increase in money supply when LPC becomes parallel to X-axis.

.....

.....

.....

2) What will be the shape of LM curve if speculative demand of money is insensitive to interest rate?

.....

.....

.....

3) Explain the meaning of "crowding out" effect.

.....

.....

.....

.....

16.6 LET US SUM UP

In this unit we have given you the limitations in the working of barter system of exchange and the need for introduction of money. The various functions performed by money have been introduced. There are four measures of money supply and each one of them has been discussed. Demand of money has been divided into transactionary, precautionary and speculative purposes.

Factors determining demand for money have been discussed exhaustively using Keynesian approach.

The supply of money and demand for money have been introduced primarily to

discuss money market and LM curve. With the help of four-part diagram approach LM curve has first been derived and later on the factors determining its slope and shift are introduced.

Finally, the concepts of surplus and deficit in money market are introduced along with the role played by LM curve in determining equilibrium level of national income.

16.7 KEY WORDS

Barter : It is a system of exchange in which one commodity/ service is exchanged for another.

'Crowding-out' Effect : It is a situation in which private investment is crowded or choked by public expenditure because of interest rate being influenced by changes in level of income.

Deferred Payments : This is another name for future payments or where payments are postponed to the future.

Double-Coincidence: It is one of the problems faced in the operation of barter system of exchange. It refers to finding a person who has the good which you require and is also needing the good which you possess in surplus.

Financial Assets : These are the assets consisting of shares, bonds, debentures, bills of exchange etc.

Inter-bank deposits : These are the deposits of one bank with other banks.

Indeterminate exchange rate : It is an exchange rate which is not determined on scientific basis. This occurs in barter system of exchange.

LM curve : It is a locus of points at which at different combinations of income and interest rate demand for money or liquidity is equal to supply of money available.

Liquid Assets : These are the assets, which can be converted into money without loss of much time and without undergoing monetary loss.

Money Market : It is a market where demand and supply of money together determine interest rate.

Real Cash Balances : These are the money holdings adjusted by price index which are kept by a transactor.

Stock Variables : These are the variables which can be measured at a point of time.

16.8 SOME USEFUL BOOK

Shapiro Edward, 1984, *Macro Economic Analysis* (5th Edition); Galgotia Publications: New Delhi (Chapters 4-8).

Hansen, Alven H. 1953, *A Guide to Keynes, International* Students Edition, McGraw-Hill Book Company, Inc: Tokyo (Chapters 1-6).

16.9 ANSWERS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) The introduction of money facilitates the process of exchange because the problems of indivisibility of goods and problem of double-coincidence faced in barter can be got over. Moreover, the indeterminateness of terms of exchange can be solved with the introduction of money.

2) M_1 = Currency held by the public + net demand deposits of banks + other deposits of the Reserve Bank of India

$M_3 = M_1$ + net time deposits of banks.

Check Your Progress 2

1) An increase in money supply will not reduce interest rate in the range of LPC when it becomes parallel to X-axis. This is referred to as a situation of liquidity trap.

2) Using the four-part diagram we will see that if speculative demand of money is insensitive to interest rate or in part (c) LPC curve becomes parallel to Y-axis, LM curve will also become parallel to Y-axis.

3) 'Crowding out' effect is a situation when an increase in public investment or, government expenditure is not allowed to have an expected multiplier effect on income because private investment begins to fall due to an increase in interest rate. Thus, the full multiplier effect is not felt.

UNIT 17 INTEGRATION OF GOODS AND MONEY MARKETS

Structure

17.0 OBJECTIVES

On going through this unit you would be able to explain:

- Goods Market equilibrium;
- Money market equilibrium;
- Integration of Goods and Money markets;
- Situations of disequilibrium in goods and money markets;
- Steps to restore equilibrium;
- Classical approach and its policy implications; and
- Keynesian approach and its policy implications.

17.1 INTRODUCTION

This unit relates to integration of goods and money markets. You have already been introduced to these markets in Units 15 and 16. It has been seen that equilibrium in goods market can be achieved at various combinations of interest rate and national income. Same is the case in respect of money market. In this unit we want to integrate goods and money markets in such a way that we get one combination of interest rate and national income at which both the markets are in equilibrium.

The technique of IS-LM will be employed to integrate goods and money markets. It will also be shown that if there is disequilibrium in both the markets, or in one of these, there exists a mechanism which helps restore equilibrium.

The integration of goods and money markets will also enable you to distinguish between Classical and Keynesian approaches to appreciate the policy implications of each.

17.2 GOODS MARKET EQUILIBRIUM

The goods market equilibrium has already been introduced to you in Unit 15. The equilibrium in goods market or what is termed as real market is achieved where injections are equalized to leakages. Injections consist of autonomous investment, government expenditure and exports , which are taken as autonomous variables. Leakages on the other hand, consist of savings, taxes and imports assumed to be dependent on interest rate or national income. Thus, at various combinations of interest rate and national income we can get different points of equilibrium in goods market. This also helps us to identify the situations of surplus and deficit in goods market.

17.2.1 Constituents of Goods Market

Various constituents of goods market are investment, government expenditure and exports. We are assuming away the flow of foreign capital and also taking price level as given. All these constituents termed as injections are taken as autonomous variables i.e., they are not influenced by national income. Other constituents of goods market are savings, taxes and imports, which are termed as leakages. Saving is taken as an increasing function of national income though interest rate can also influence the level of saving. Taxes are taken to be an increasing function of national income. Sometimes savings are taken to be an increasing function of disposable income, which is defined as national income minus direct taxes. Similarly imports are taken to be an increasing function of national income though we can have elements of imports which may be autonomously given.

17.2.2 Investment-Savings Equality

As you already know from Unit 15, in an economy where there is no government and further if it is a closed economy, i.e., there are no exports and imports an equilibrium in goods market is ensured where (ex-ante or, planned) investment becomes equal to (ex-ante or, planned) savings. This equality, in turn, ensures that aggregate demand in the economy is sufficiently of the level of aggregate supply so that there is no unplanned increase or decrease in inventories. A level of national income where aggregate demand is equal to aggregate supply is termed as equilibrium level of national income, which also indicates equilibrium in goods market.

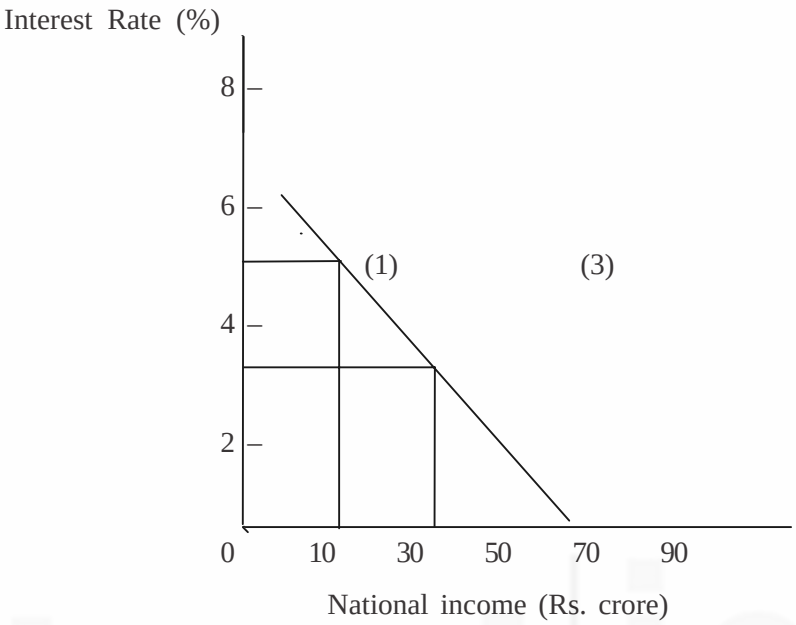
Any other level of national income where investment is not equal to savings (or where aggregate demand is not equal to aggregate supply) represents disequilibrium in goods market. Disequilibrium can be of two types: (1) Surplus in goods market, which represents a situation of aggregate supply being more than aggregate demand, or savings being more than investment and (2) deficit in goods market, which, represents a situation of aggregate supply being less than aggregate demand, or savings being less than investment.

17.2.3 Investment-Savings Equality at Different Combinations of Interest Rate and National Income

Equilibrium in goods market at various combinations of interest rate and national income, as shown in Unit 15 is represented by IS curve. The IS curve represents equality of investment with savings, or equilibrium in goods market,

at various combinations of interest rate and national income is reproduced in Fig. 17.1 for ready reference, to be referred to in Section 17.4.1.

Fig.17.1



In Fig. 17.1: we are recalling the IS curve made in part (d) of the four- part diagram in Unit 15. The vertical axis shows rate of interest in per cent and the horizontal axis shows the level of national income. Each point on IS curve shows equilibrium in goods market.

In Fig. 17.1, national income is measured on X-axis in rupees crore and interest rate in % on Y-axis. On the IS line, at point (1) at 5% interest rate and Rs. 20 crore of national income investment is equal to saving. Similarly at point (2) where 3% is the rate of interest and Rs.40 crore national income, again investment is equal to saving. We can get many other points on IS line at which investment-saving equality will take place. Point (3) indicates surplus in goods market. Similarly point (4) indicates deficit in goods market. In fact all points above IS line indicate surplus and those below IS line show deficit in goods market.

17.3 MONEY MARKET EQUILIBRIUM

Money market equilibrium has already been introduced to you in Unit 16. The equilibrium in money market is achieved where demand for money is equal to supply of money. Supply of money is a stock, which can be measured at a point of time. Demand for money consists of demand for transactionary and precautionary purposes, which are lumped together. The transaction demand for money is represented by M_t. Demand for money also consists of demand for speculative purposes, which is represented by M_{sp}. So, the total demand for money concerning M_t is an increasing function of national income and that of M_{sp} is a decreasing function of interest rate. Thus, at various combinations of interest rate and national income we can get various situations of equilibrium in money market. This also helps us to identify situations of surplus and deficit in money market.

17.3.1 Constituents of Money Market

Constituents of money market are supply of money, demand for money for transactionary and precautionary purposes, and demand for money for speculative purposes. From among these, supply of money is autonomously given. The demand for money for transactionary and precautionary purposes is taken to be an

increasing function of national income. However, demand for money for speculative purposes is a decreasing function of interest rate. At a very low interest rate demand of money for speculative purposes becomes infinite. These constituents of money market are to be considered together to arrive at equilibrium in money market.

17.3.2 Demand-Supply Equality of Money

As known to you from Unit 16 equilibrium in money market is ensured where demand for money is equal to supply of money. This equality, in turn, ensures demand- supply equality of money available for speculation purposes.

Fig.17.2

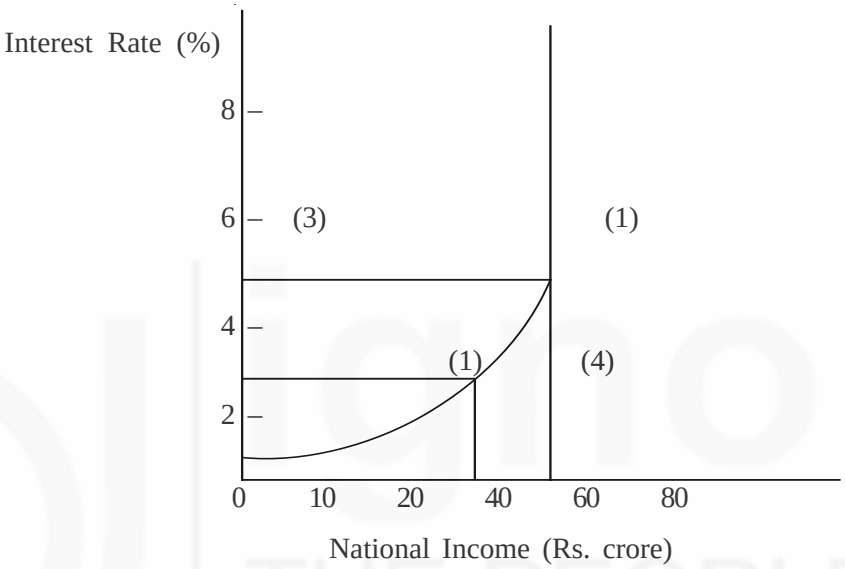


Fig.17.2 is nothing but L M curve designed in Unit 16. Here too we have shown the per cent rate of interest on the vertical and national income on the horizontal axis. Each point on the L M curve shows equilibrium in the money market.

In Fig. 17.2 national income is measured on X-axis and interest rate on Y-axis. On the LM line, at point (1) where the interest of rate is 5% and national income is Rs.45 crore, demand for money (L) is equal to supply of money (M). Similarly at point (2) at 3% interest rate and Rs.40 crore of national income again demand for money is equal to supply of money. We can get many other points on the LM line at which demand-supply equality of money will take place. Point (3) indicates surplus in money market. Similarly point (4) indicates deficit in money market. Disequilibrium in money market is a situation where demand for money is not equal to supply of money. Disequilibrium can be of two types: (1) surplus in money market which is a situation of supply of money being more than demand for money and (2) deficit in money market which represents a situation of supply of money being less than demand for money. This LM curve will be used in Section 17.4 along with the curve of Fig. 17.1 to determine simultaneous equilibrium in both goods and money markets.

17.4 INTEGRATION OF GOODS AND MONEY MARKETS

Integration of goods and money markets involves taking both the markets together against given interest rate and national income. As you already know from sections 17.2 and 17.3 above that goods and money markets are represented by IS and LM curves respectively. An integration of two markets will involve taking

interest rate on Y-axis and national income on X-axis in a two-dimensional diagram and plotting IS and LM curves to get the point of equilibrium. We will also have to work out the mechanism by which equilibrium in one or both the markets is restored in case of an disturbance.

17.4.1 IS and LM at Different Combinations of Interest Rate and National Income

We will make use of the information contained in Figs. of 17.1 and 17.2 and put these together in Fig. 17.3.

Fig.17.3

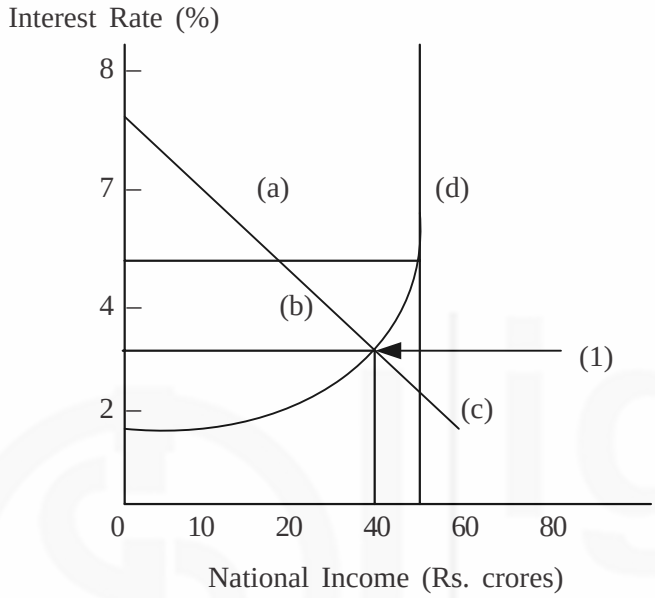


Fig. 17.3: Brings together the Figs. 17.1 and 17.2. The IS curve shows equilibrium in goods market and the L M curve shows that in money market. The point of intersection between these two curves gives us that combination of rate of interest and national income at which both goods and money markets attain equilibrium simultaneously.

The points on IS and LM curves represent equilibrium in goods market and money market respectively at various combinations of interest rate and national income. Region (a) in Fig. 17.3 indicates surplus in goods market as well as in money market. Region (b) indicates a situation of deficit in goods market and surplus in money market. Region (c) is indicative of deficit in goods market as well as deficit in money market. Region (d) gives us surplus in goods market and deficit in money market. It is only at point (1) where IS and LM curves intersect that we have equilibrium in goods market as well as in money market and this simultaneous equilibrium in both the markets is achieved at a combination of interest rate of 3% and national income of Rs. 40 crore. At no other point simultaneous equilibrium in both markets will be achieved.

17.4.2 Steps to Restore Equilibrium in Goods and Money Markets

Let us redraw Fig. 17.3 to spell out the steps to be taken to restore simultaneous equilibrium in goods and money markets. In case of a situation of disequilibrium this will be explained in Fig. 17.4.

Fig.17.4

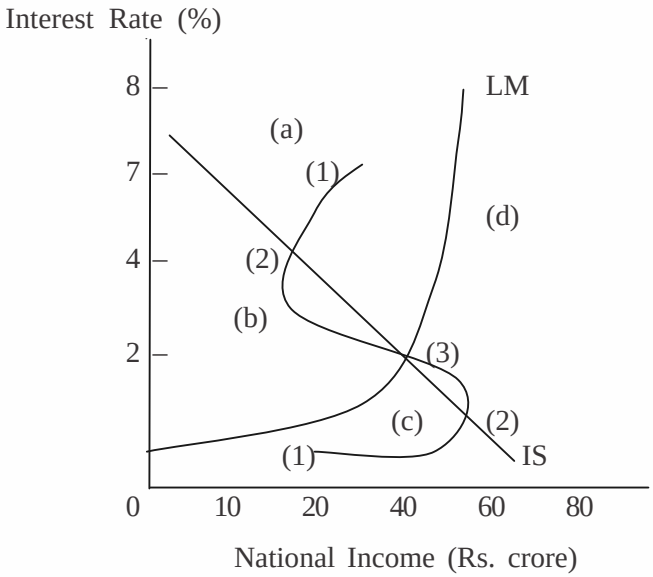


Fig.17.4 Explains that if the economy happens to be away from the point of intersection of IS and L M curves, that is, at any point in the regions (a), (b), (c) or (d) then what steps would restore the equilibrium again.

In Fig. 17.4 the process of restoration of equilibrium from disequilibrium is shown. Let us take point (1) in region (a), which represents a situation of surplus in goods as well as money markets. In order to correct surplus in goods market, we require contractionary fiscal policy, i.e., reducing government expenditure, or autonomous investment, or increasing taxes, or reducing exports. This will lead to a fall in national income. The direction to the left at point (1) indicates reduction of national income. Similarly, to correct surplus in money market we need to reduce interest rate or adopt expansionary monetary policy. The direction downward at point (1) indicates reduction of interest rate. As we go along adopting contractionary fiscal policy and expansionary monetary policy we cross over to region (b). In point (2) is in region (b), we face the situation of deficit in goods market and surplus in money market. In order to correct this situation we start adopting expansionary fiscal policy and expansionary monetary policy, which finally leads us to point (3), which represents simultaneous equilibrium in goods and money markets.

We could have started from point (1') of region (c) representing deficit in goods as well as money markets for which we adopt contractionary monetary policy and expansionary fiscal policy. As we continue pursuing these policies, we reach point (2') of region (d) showing surplus in goods market and deficit in money market. So, from point (1') we move to point (2') adopting contractionary fiscal as well as monetary policy. This leads us finally to point (3) where simultaneous equilibrium in both the markets is attained.

How quickly equilibrium in both the markets is restored depends on slopes of IS and LM curves. If IS curve is steep given the LM curve, i.e., point (3) will be achieved quickly.

Check Your Progress 1

1) Refer to Figure 17.4 take a point in region (d) Work out the steps to restore equilibrium point (3).

.....

.....

.....

.....

.....

2) What is the reason for the LM curve to be parallel to X-axis at a very low level of interest rate?

.....

.....

.....

.....

.....

3) Distinguish between contractionary monetary and fiscal policies.

.....

.....

.....

.....

.....

17.5 CLASSICAL APPROACH WITH THE HELP OF IS-LM TECHNIQUE

Using IS-LM technique developed in Section 17.4.1 we can distinguish between classical and Keynesian approaches. The classical economists were of the view that money is demanded only for transactionary purposes. They also maintained that an economy operates always at full employment of income and there cannot exist unemployment because of shortage of aggregate demand. J.B. Say propounded, what is termed as Say's law of market, according to which every supply creates its own demand. In case we assume non-existence of savings, Say's law of market implies that in order to produce more goods and services we require the services of factors of production and for rendering factor services they receive factor incomes. Factor incomes received by factors of production are used to purchase goods and services produced by enterprises. A factor of production is employed upto a point where the remuneration given to it becomes equal to marginal revenue product - as stated by marginal productivity theory of distribution. This way, the sum of factor remunerations adds up to value of total product. Thus, there cannot be shortage of demand. If savings are introduced factor income received by factors of production will be used partly to finance consumption expenditure and partly for savings purposes. Consumption expenditure is used to purchase consumer goods and services and savings are used to finance investment activity. If savings are equal to investment again there cannot be shortage of demand in the economy. Saving-investment equality in classical approach is ensured by interest rate variation and it keeps changing until savings are equalized to investment. Demand for labour is equalized to supply of labour by variation in real wage rate. The working of classical approach can be illustrated with the help of IS-LM technique as shown in Fig. 17.5.

Fig. 17.5

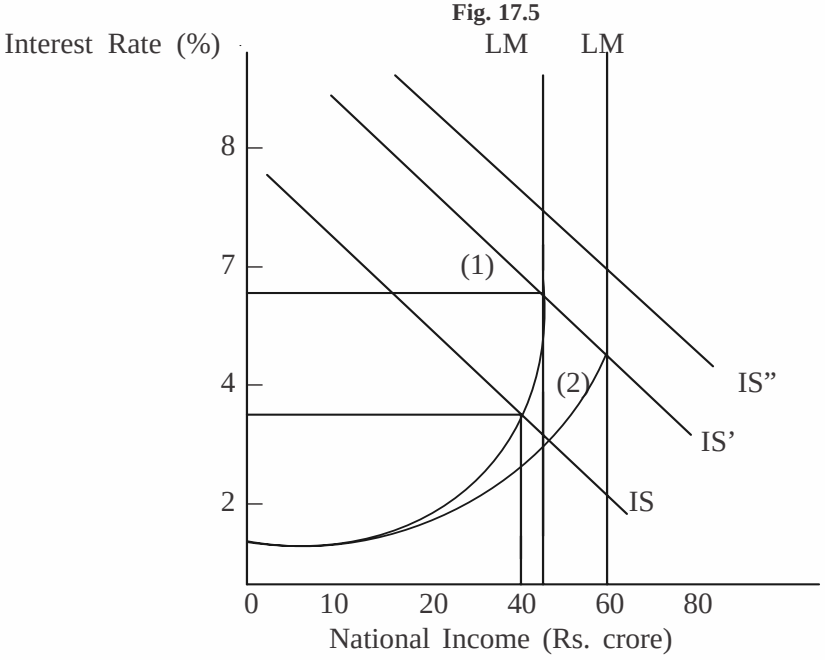


Fig.17.5 tells us that the portion of LM curve, which is parallel to Y-axis conforms to classical approach. In such situation an increase in government expenditure reflected in the right wards shift of IS curve does not lead to any increase in the national income - only rate of interest rises. Here the only way to raise the national income is to expand the supply of money. In other words, the fiscal policy is totally ineffective here. Only the monetary policy can work.

The shape of LM curve beyond point (1) when it becomes parallel to Y-axis is referred to as classical range. The meaning of such a feature is, if investment, or government expenditure rises as shown by shift of IS curve from IS' to IS" it will result in increase in interest rate from 5% to 7.5% keeping the level of national income at Rs.45 crore. If national income is to be raised from Rs.45 to Rs.55 crores the only way out is to increase the money supply which is shown by shift of LM curve from LM to LM' given the IS' curve.

It can be easily seen that the slope of IS curve matters a lot in determining the effect of monetary policy. If IS curve is steeper than IS' more money supply has to be resorted to achieve a new national income level of Rs.55 crore, or we can put the same thing by saying that monetary policy will be less effective if IS curve is steep. Similarly flatter the IS curve more effective will be the monetary policy.

17.5.1 Policy Implications of Classical Approach

It is time to see the policy implications of the classical approach. Refer to the Fig. 17.5. In this figure we are at equilibrium. The level of national income is Rs.45 crores at which both goods and money markets are in equilibrium. Now, if the full employment level of national income is Rs.55 crore at which resources are fully employed, this can be achieved only by increasing the money supply which is indicated by shift of the LM curve from LM to LM' given the IS curve. The extent of shift of the LM curve to the right will depend on slope of the IS curve. In Fig. 17.5 the IS is given by IS'. The full employment equilibrium level of income is given by the intersection of IS' with LM' at point (2). On the other hand, if IS curve had been steeper than IS' the extent of shift of LM curve would have been more, implying thereby, greater increase in money supply would be required to achieve full employment level of national income.

Thus, we realize immediately that in the classical range when LM curve becomes parallel to Y-axis the only policy, which can work, is to shift the equilibrium level of income. The monetary policy and the fiscal policy, if adopted, would result in change of interest rate. For instance if expansionary fiscal policy

had been adopted, as shown by the shift of IS curve from IS' to IS'', the interest rate would have increased from 5% to 7.5% keeping the level of national income unaffected at Rs.45 crore.

Therefore, monetary policy is the only policy which will be effective in the classical range. The degree of effectiveness of such a policy will depend on the slope of the IS curve. Steeper the IS curve less effective will be the monetary policy or flatter the IS curve more effective will be the monetary policy. The reason why interest rate rises with adoption of expansionary fiscal policy lies in the fact that increase in investment, or government expenditure will result in an increase in national income depending on the value of multiplier. This rise in national income will result in demand for money for transaction purposes. Given the money supply, if demand for money for transaction purposes increases, the supply of money available for speculation purposes falls. This leads to an increase in interest rate. We can also put the same thing by saying that as public investment or government expenditure increases, private investment falls resulting in the operation of "crowding out" effect.

17.6 KEYNESIAN APPROACH WITH THE HELP OF IS-LM TECHNIQUE

Unlike the classical range where LM curve is parallel to Y-axis, we have another extreme where LM curve becomes parallel to X-axis and the demand of money is infinite. Keynesian approach is at its effective best in this range. It has been depicted with the help of Fig. 17.6.

Fig.17.6

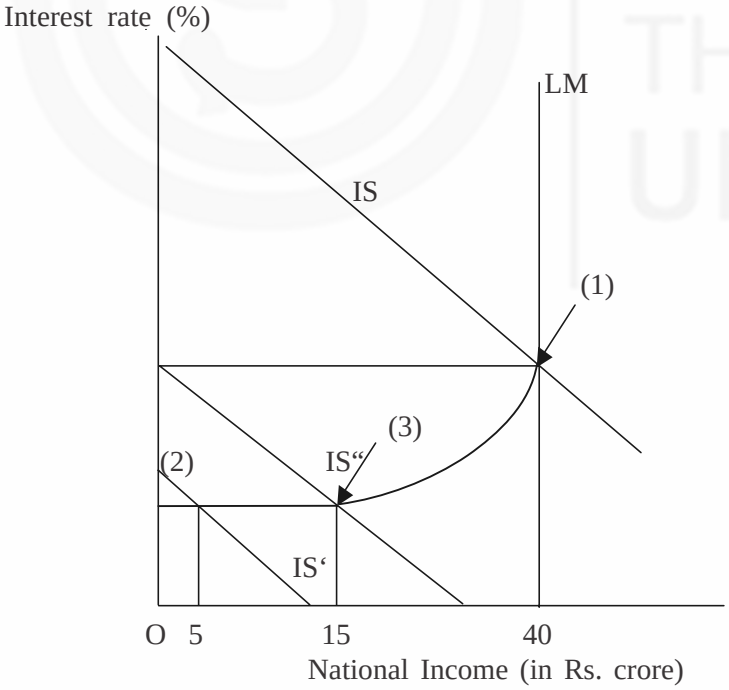


Fig. 17.6: Focuses on that portion of L M curve, which is parallel to X-axis. This is the region in which the Keynesian approach is most effective. The economy is deep in 'liquidity trap' and it does not respond to monetary policy changes. But the fiscal policy works very effectively here.

In Fig. 17.6 IS-LM curves are drawn with their intersection at point (1). At this point interest rate is 3% and national income is Rs.40 crore. At 0.5% interest rate, the LM curve is drawn parallel to X-axis which is also referred to as a

situation of liquidity trap. In this region, demand of money for speculation purposes is infinite. It is called Keynesian range.

In Keynesian range or liquidity trap range, equilibrium level of national income is Rs. 5 crore and 0.5% interest rate. This is shown at point (2). If national income is to be increased from Rs.5 crore to Rs.15 crore it will not be achievable since interest rate of .5% will remain unchanged. On the other hand, the an increase in autonomous investment or government expenditure through multiplier will be in a position to achieve national income of Rs.15 crores. As you already know increase in autonomous investment or government expenditure is reflected by a shift of IS curve to the right. So, in this case, IS curve has to move to the right from IS' to IS'' to increase national income from Rs.5 crore to Rs.15 crore. This shift of IS curve is what is referred to as the adoption of expansionary fiscal policy. Similarly to reduce national income we have to adopt contractionary fiscal policy.

It can be seen that the suitable policy to be adopted in Keynesian range is fiscal policy i.e., the policy of changing government expenditure and for taxation.

17.6.1 Policy Implications of Keynesian Approach

In Fig. 17.6 the initial equilibrium level is given at point (2). If the full employment level of national income is Rs. 15 crore, it can be achieved by increasing autonomous investment or government expenditure or reducing tax revenue which is indicated by the shift of IS curve from IS' to IS'' given that LM curve is parallel to X-axis. It is important to see that the slope of the IS curve is of no consequence in achieving full employment level of income. The full employment equilibrium level of income is given by the intersection of IS'' with LM as represented by point (3).

Thus, in the Keynesian range where LM curve becomes parallel to X-axis, the only policy, which can work to shift equilibrium level of income, is fiscal policy. The monetary policy, if adopted, would be unsuccessful because it operates through changes in rate of interest, which is ineffective here

Another important point to be noted is the "crowding out" effect which was referred to in Section 17.5.1. It will be totally non-existent in Keynesian range. The implication of the non-existence of "crowding out" effect is that full multiplier effect as a consequence of change in autonomous investment or government expenditure will be felt on changing the level of national income. It is same thing as saying that Keynes assumes away the money market in his analysis in liquidity trap range.

Finally, we can refer to the intermediate range, which lies between Keynesian range and classical ranges. It is not shown but you can analyse the situation in Fig. 17.6 by taking the starting point (1) at which interest rate is 3% and national income is Rs.40 crore. Assume full employment national income is Rs.42 crore. It can be shown that this new full employment equilibrium of national income can be achieved either by shifting IS curve to the right keeping LM curve unchanged, or shifting LM curve to the right keeping IS curve unchanged or shifting both IS and LM to the right. Obviously shift in IS or LM curves to the right will be less if there is simultaneous shift of the both as compared to a situation where only one of them is to be shifted.

The relative effectiveness of monetary and fiscal policies to achieve full employment equilibrium level of income depends on the slopes of IS and LM curves. The general rule is that flatter the IS curve given the LM curve the more effective is the fiscal policy. Similarly steeper the IS curve, given the LM curve, monetary

policy is more effective. Moreover, flatter the LM curve with a given IS curve, monetary policy is more effective. You should take this as an exercise and show diagrammatically the four cases mentioned above concerning the relative effectiveness of monetary and fiscal policies in the intermediate range.

Check Your Progress 2

1) Explain the steps involved in the working of 'crowding out' effect.

.....

.....

.....

2) Which policy, monetary or fiscal, will you recommend if LM curve becomes flatter given IS curve?

.....

.....

.....

.....

3) If tax is raised, argue how does it lead to less fall in National Income if Money Market is introduced along with Goods Market?

.....

.....

.....

17.7 LET US SUM UP

In this unit we have given you an idea about goods market equilibrium by introducing its constituents. Investment-saving equality has been discussed to get the equilibrium in goods market. Various combinations of interest rate and national income are introduced to arrive at investment-saving equality. Similarly, money market equilibrium, its constituent's equality between supply and demand for money at various combinations of interest rate and national income are discussed.

Integration of goods and money markets is discussed with the help of IS and LM curves to get a unique combination of interest rate and national income at which both goods market and money market are in equilibrium. How to achieve equilibrium in both the markets if they are in disequilibrium has also been discussed.

Finally, using IS-LM technique a detailed discussion has been attempted concerning classical, Keynesian and intermediate range. The policy implications of classical and Keynesian approaches have been spelt out at the end.

17.8 KEY WORDS

'Crowding-out' Effect : It is a situation in which private investment is crowded or choked by public expenditure because of interest rate being influenced by changes in level of income.

Double-Coincidence : It is one of the problems faced in the operation of barter system of exchange. It refers to finding a person who has the good which you require and is also needing the good which you possess in surplus.

Financial Assets : These are the assets consisting of shares, bonds, debentures, bills of exchange etc.

Inter-bank Deposits : These are the deposits of one bank with other banks.

Indeterminate Exchange Rate : It is an exchange rate which is not determined on scientific basis. This occurs in barter system of exchange.

LM Curve : It is a locus of points at which at different combinations of income and interest rate demand for money or liquidity is equal to supply of money available.

Liquidity Trap : It is that portion of LM curve where change in money supply has no influence on interest rate or it is a situation where the demand for money is infinity at a low interest rate.

Liquid Assets : These are the assets, which can be converted into money without loss of much time and without undergoing monetary loss.

Money Market : It is a market where demand and supply of money together determine interest rate.

17.9 SOME USEFUL BOOKS

Lipsey. Richard G. 1983, *An Introduction to Positive Economics,* (6th Edition), E.L.B.S. and Weidenfeld and Nicolson, London (Chapters 35-36).

Shapiro, Edward, 1984, *Macro Economic Analysis* (5th Edition), Galgotia Publications, New Delhi (Chapters 1 to 12).

Samuelson, Paul A. 1973, *Economics* (9th edition), International Student Edition, McGraw-Hill Kogakusho Ltd Tokyo (Chapters 11-12).

17.10 HINTS/ANSWERS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) In region (d) we have deficit in money market and surplus in goods markets. Therefore, we will adopt contractionary monetary policy to raise the interest rate and contractionary fiscal policy lowering national income, or direction will be one facing upwards and other facing leftwards. As we keep adopting those policies we will cross to region (a) and from there we will adopt the route to reach point (3) as shown in Fig. 17.4

2) At a very low interest rate everybody expects future interest rate to rise. A consequence of this is the expectations of price of bonds to fall. Therefore, people would like to keep their wealth in the form of liquidity or money. Hence demand for money is infinite at a very low interest rate.

3) Contractionary monetary policy is a policy when money supply is reduced or

interest rate is raised. Since rise in interest rate lowers investment and hence national income, we call it a contractionary monetary policy.

Contractionary fiscal policy is an intervention when government expenditure is reduced keeping tax revenue unchanged or when tax revenue is raised keeping government expenditure unchanged . This leads to lowering of national income through expenditure and tax multipliers.

Check Your Progress 2

1) As public investment or government expenditure is increased national income rises due to multiplier effect. A rise in national income leads to rise in demand of money for transaction purposes and results in less supply of money available for speculative purposes. This raises interest rate which, in turn, leads to fall in private investment and therefore, the national income.

Hence, an initial increase in autonomous expenditure does not lead to as much increase in national income as it would have happened had there been no money market existing along with goods market.

2) Fiscal policy would be more suitable since we are approaching Keynesian range with LM curve becoming flatter.

3) Tax multiplier is given by [-c]/[1–c], where c is marginal propensity to consume.

A rise in tax leads to fall in national income depending upon the tax multiplier [-c] / [1 – c]. This leads to a fall in level of demand for money for transactions purposes. Consequently, supply of money available for speculation purposes rises which, in turn, leads to fall in interest rate and through that to a rise in national income.

Thus national income finally does not fall as much as it would have if money market did not exist along with goods market.

UNIT 18 QUANTITY THEORY OF MONEY

Structure

18.0 OBJECTIVES

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

- understand the importance of money and functions it performs;
- know why or for what reason individuals demand money;
- ascertain theoretical implications of demand for money; and
- compare the theories of demand for money and their policy implications.

18.1 INTRODUCTION

In a barter economy, goods and services were exchanged directly for goods and services. For example, if person 'A' has surplus shoes and she wanted rice and another person 'B' has surplus rice and she wanted shoes and they are able find each other. Only then the exchange will be possible such a system of exchange is called Barter. The barter system is also referred to as the direct exchange of one good or, service for another without mediation of money.

Problems of a Barter System

The barter system has many disadvantages. The main four of these are as follows:

i) **Lack of a common measure of value**: There is no common measure of value in a barter system. All commodities do not possess equal value. Suppose 'A' has rice and 'B' has wheat. How to decide that so much of wheat is equal to so much of rice? As there is no common measure of value, the ratio will be fixed according to the intensity of reciprocal of demand by 'A' and 'B';

ii) **Absolute insistence on double coincidence of demand**: The functioning of barter system necessitates a double coincidence of demand on the part of those involved in exchange of goods and services. It is absolutely necessary for one party to want exactly what the other party is offering in exchange and *vice versa*. If this double coincidence of demand is not matched exactly, no exchange

will take place. Suppose, for example a farmer wants to sell her wheat and wants to buy shoes; then he has to find a person who has surplus shoes to sell and wants to buy wheat in exchange;

iii) **Indivisibility of commodities**: There are many goods, which cannot be divided. Suppose 'A' has a horse and 'B' has 1 kg of rice. Both of them agree to an exchange transaction. But according to the ratio of exchange arbitrarily fixed between them a portion of horse is equal to the value of 1 kg of rice. The transaction cannot take place because it is unthinkable to cut out a portion of living horse for exchange; and

iv) **Difficulties in storing wealth**: Another problem under the barter system is related to storing of wealth for future use. Most of the goods like wheat, horses, leather etc., lack sufficient durability and deteriorate with passage of time. Therefore, they cannot be conveniently stored for future use.

The barter system was compatible only with simple and primitive economies where the material needs of the people were few and everyone wore the same kind of clothes, the same kind of food and engaged in similar activities. But as the civilization advanced, wants multiplied and a certain degree of division of labour was achieved. In the changed economic scenario the problems of barter system became highly pronounced. Such a feature encouraged some suitable alternative medium of exchange, which would eliminate above-mentioned difficulties of barter system and facilitate transactions. This led to the innovation of 'Money'.

Money is not needed for its own sake as one needs food, clothes and a house for living but to mediate transactions. It has purchasing power, which enables us to exchange goods and services. This feature makes money a unique commodity.

18.2 MONEY: FEATURES AND FUNCTIONS

We are all familiar with money in our day-to-day transactions. Perhaps the oldest and simplest role of money has been the 'medium of exchange' for all economic transactions. This is because money is acceptable to everybody. In short, money is anything, which is acceptable as a means of payment in the settlement of all transactions, including debt. It is commonly used as medium of exchange or means of transferring purchasing power. General acceptability as a means of payment or as a medium of exchange is the unique feature of money. It does not need to be converted into something else before it can be spent or used for settlement of debt. What makes money 'money' is the belief held by everyone that it would be accepted as such by all others in the economy.

Functions of Money

The function of money have been well summed up in a couplet:

Money is a matter of functions four
A medium, a measure, a standard, a store

Thus, the four broad functions money are:

i) **Medium of Exchange:** We saw in the last section that in barter economy there must be a double coincidence of demand for a transaction to occur. Money takes care of this problem: persons A, B, and C can sell their output to others for money with which they can buy the things they require. This property of money is associated with development of commerce and trade on a larger scale. Emergence of money has also made the transactions quicker and economical to settle.

ii) **Unit of Value:** Money customarily serves as a common unit of account or measure of value in terms of which the values of all goods and services are expressed. This makes possible to have a meaningful accounting system by adding up the values of a variety of goods whose quantities are measured in different physical units. Important example of value totals is the national income estimates of a country, total money cost of a project, total sale proceeds of a firm producing many products etc. This makes comparisons of various kinds across time and across regions possible. It has truly been said that it has been possible for economics to grow as a science, because it analyses social behaviour concerned with production, exchange, distribution and consumption of goods and services whose value could be measured in a common unit, money.

iii) **Standard of Deferred Payment:** Money also serves as a standard or unit in terms of which deferred or future payments are stated and settled. This applies to payments of interest, rents, salaries, pensions, insurance premium, etc. In a money using economy, the bulk of deferred payments are stipulated in money terms. But large fluctuations in the value of money (because of inflation or deflation of prices) makes money not only a poor measure of value, but also a poor standard of deferred payment.

iv) **Store of Value:** Money also serves as a store of value, i.e., people can hold their wealth in the form of money. This function is derived from the use of money as a medium of exchange in a two-fold manner: First, the use of money as a medium of exchange decomposes a single barter into two separate transactions of purchase and sale. Under barter, purchase and sale are necessarily simultaneous operations. The use of money separates the two transactions in time. This will require that the medium of exchange also serves as a store of value. This is encouraged by unique feature of money that is, generalized purchasing power, and as such only perfectly liquid asset. No doubt, money is not the only store of value. There are other assets like gold, shares, and bonds. But money is unique as a store of value in that it alone is perfectly liquid. That is, it alone serves as a generally acceptable means of payment. The fluctuation in the value of money that affects its functions as a measure of value and as a standard of deferred payment also influences its role as a store of value.

Check Your Progress 1

1) What are the problems in having a horse as a medium of exchange?

.....

.....

.....

.....

2) Now you know the functions of money, i.e., it should be stable, divisible, durable and portable. Rank the following commodities according to their closeness to money, stating reasons, in the order of your preference:

i) sugar; ii) horse; iii) salt; iv) prepared *idli*; v) gold.

.....

.....

.....

.....

3) What are the differences between money and barter economy?

.....

.....

.....

.....

18.3 ALTERNATIVE THEORIES OF DEMAND FOR MONEY

To be able to analyze the effects of changes in the stock of money, it is essential to study the equilibrium of the money market. Money is an asset for the holding public, therefore, the public must have a demand for it and a supply of it, and so also a market for it. The demand for money comes from the general public while its supply comes from the government and the banking system, whose liability the money is. The money market, then is, simply the market comprising these demanders and suppliers of money. In this unit, we shall assume that supply of money is autonomously given by the monetary authority.

Money is a stock variable. The stock of it refers to its quantity at a point of time. As an asset, the demand for it is the public's demand to hold money, whatever the reason (motive) for holding it and whatever length of the time period for which it is held. *Holding cash in one's pocket for spending it is as much a part of the demand for money as burying currency notes in a pot.* The several motives for holding money will be studied later in this unit as an explanation for various theories of demand for money.

We will analyse the aggregate demand for money, i.e., the demand for money of the public as a whole. Thus, we will consider, the sum of all the money demanded by individual members of the public, whether households or firms.

Theories of demand for money are mainly concerned with the question: what are the determinants of public's demand for money and why? A related question is: why does the public demand money? Several explanations have been offered in reply. This unit will explain them in detail. Each explanation has its own implications for explaining the economic effects of changes in supply of money.

The classical theory of demand for money, popularly known as the **Quantity theory of Money** (**QTM**), is basically is a theory of the price-level. However, under the influence of Keynes, the theory of demand for money became a theory of rate of interest, output and employment. Friedman later tried to rescue the *quantity theory of money* through his restatement. In his version of the theory of demand for money he completely neglected the Keynes' classification of motives for holding money and corresponding components of the demand for money. Instead of motives, he identified the key determinants of the demand for money. The essentials of his theory have been set out in format of the classical theory, and relatively much less in terms of Keynesian theory.

Theories of demand for money could be divided into three broad categories, viz.;

a) Classical theory of demand for money or the QTM;

b) Keynesian theory of demand for money; and

c) Friedman's restatement of Classical QTM.

18.4 QUANTITY THEORY OF MONEY— CLASSICAL APPROACH OR THEORY OF PRICE LEVEL

Several versions of classical *Quantity Theory of Money* are popular. One version, also known as **transactions version** is due to Fisher. It is also called **Fisher equation of exchange:**

M.V = P.T

Where

T is number of transaction of average size

M is defined as quantity money,

V is velocity of circulation of money, and

P is the average price level.

where **T** is a proxy for level of income.

The classical macroeconomic theory relies on the QTM as the theory of demand for money. This theory says that it is the quantity of money in the hands of the public that determines how high or low the price level will be. Such a conclusion has been reached since level of output in the classical model is always at the full capacity (or full employment) level.

It is assumed that output in classical system is 'given' or constant for the duration of the analysis. There T is fixed and it is a proxy for national income. Velocity of circulation of money (V) is dependent on the payment behaviour of people and is, therefore, a long term constant. It is defined as a number of times a rupee changes hands during a given accounting period.

Given as above definitions, product PT will represent product of number of average sized transaction and average price, which is equal to the total amount of money needed to help facilitate sale/purchase of total output. On the other hand, components of the product MV shows how many rupees are in circulation and how many times each is used for payments. Thus, MV equals the amount of money available for transaction. When money available equals money needed, then will be equilibrium in the system.

Re-arranging the terms of the equation of exchange, MV = PT we get:

$$P = \left(\frac{V}{T}\right) M$$

Since V and T are both constants, this form of equation gives us a direct relationship between money supply and price level. If M doubles, P will also double. If M is reduced by half, Price level will also be halved. In this sense, classical quantity theory of money can be called a theory of price level.

According to another approach the classical QTM the demand for money can be described as the following relationships with 'nominal output'

M . v = P . y

where

M = Demand for money

v = Velocity of money circulation

P = Price level

y = Real output level

The above identity is converted into the QTM under the assumption that **v** and **y** are constant or stable in the short run. With **v** and **y** being constant, the assumption that price level is passive means that **P** depends on changes in **M** rather than changes in **M** depend on changes in **P**. These assumptions give us the nice and straightforward result that any short run increase (or decrease) in **M** must lead to proportional rise (or fall) in **P**. With any one or more of these assumptions not valid would imply that the proportionality is unlikely to hold between **M** and **P**.

Say's Law: S.B. Say had propounded another theory about aggregative economic behaviour of a society. His idea is : 'Supply creates its own demand'. In simple terms, it means that in the process of production a society generates sufficient amount of income, which is enough to help its members purchase that output. This theory is critically dependent on implicit assumption of there being perfect price flexibility in the society. If prices are sufficiently flexible, any quality, which is produced, can be sold out. If output exceeds existing demand, prices must fall. If supply in any time period is less than the market demand, a rise in prices will be only natural consequences. This applies to all the commodities and factors of production. That is why, sometimes it (Says' law) is regarded as a natural consequence of perfect competition.

18.5 KEYNESIAN THEORY OF DEMAND FOR MONEY

This theory was formulated by Keynes in his famous book "*The General Theory of Employment, Interest and Money*", Published in 1936.

To understand Keynes' theory two questions need to be separated: first, why is money demanded? and second, what are the determinants of demand for money? Both these questions are inter-linked. Keynesian demand for money has 3 components, they are transaction demand, precautionary demand and the speculative demand.

Keynes made the demand for money a function of 2 variables; namely,

i) money income, or **Y**; and

ii) rate of interest, or **r**.

In functional form:

$$M^d = M^d (Y, r)$$

Keynes retained the transactions approach (explained above) to the demand for money under which demand for money is hypothesized to be a function of nominal income. But, according to him, this only explained the transaction demand for money and not the entire demand for money. The revolutionary insight of Keynes has been the speculative demand for money component. Through it Keynes made this part of the demand for money a declining function of rate of interest, which is purely a monetary phenomenon and solely influenced by the monetary influences in the economy. The speculative demand for money arises from the speculative motives for holding money due to changes in the rate of interest in the market and uncertainty about them.

18.5.1 Transaction Demand for Money

Money is needed to carry out day-to-day transactions. There are discrepancies between receipts of income (say, once in a week or a month) and the expenditures of a person. A person may be assumed to incur expenditure almost daily throughout the week till her income is exhausted. Thus, while receipts of income are discrete, expenditure is almost continuous. Because of this discrepancy, it is necessary that individuals have cash at their disposal for meeting their current (or daily) expenditures. This demand for money is called the **transaction demand for money**. The level of income determines the control over goods and services in the market. Given the payments habit of the community, an individual has to have cash at her disposal to meet these expenditure requirements. An individual with higher level of income has a greater demand for goods and services (in general) than an individual with a lower level of income. It means that transactions are directly related to the level of income. In other words, more cash is needed by an individual with a higher level of income compared to one with lower level of income. Thus,

$$\mathbf{M^d_t = M^d_t(Y)}$$

The classical economists, the Quantity Theorists, also considered the transaction demand for money, which emphasizes the role of money as medium of exchange. However, the precautionary and the 'speculative' demand for money are Keynes' additional sources of 'liquidity preference' (or, demand for money).

For simplicity, we can say that transaction demand for money, is a constant proportion, **k**, of the level of national income, **Y**

$$\textbf{i.e.,} \quad \mathbf{M^d_t = k.Y = k \cdot P \cdot y} \qquad \mathbf{0<k<1}$$

This equation says that if the level of income (nominal) is say, Rs. 800 crore and **k** = 2/5, then, the transaction demand for money in the economy would amount to Rs. 320 crore [800x(2/5) = 320]. This means that the economy would requires Rs. 320 crore of money in order to finance smoothly production and sale worth of goods and services of Rs. 800 crore. If the level of national income (nominal) goes up to Rs. 1000 crore and **k** remains at the same level of 2/5 then the transaction demand for money will be Rs. 400 crore.

We know that $\mathbf{M^d_t = k .Y}$

it follows then,

Fig. 18.1

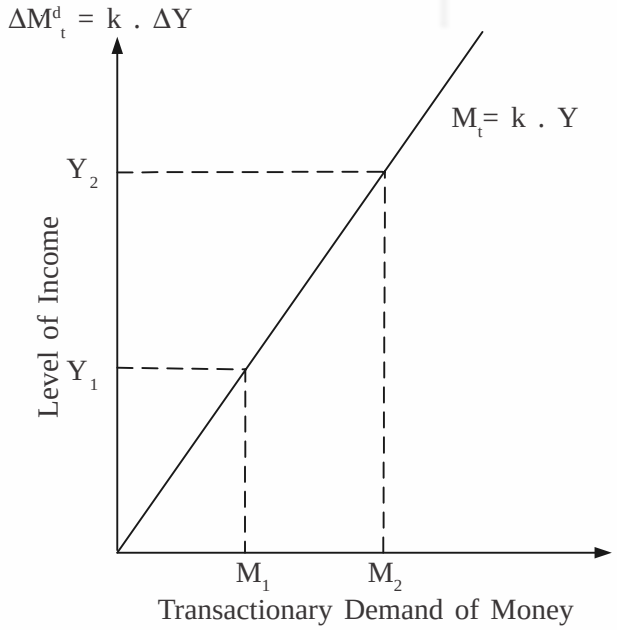


Fig. 18.1: Depicts transactionary Demand for money as a proportion k of money income. As income rises by Y_2-Y_1. The demand for money goes up by M_2-M_1. Note that M_2-M_1=k(Y_2-Y_1)

where

ΔM^d_t = change in the transaction demand for money

ΔY = change in the level of national income(nominal)

In case of a decline of Rs. 200 crore in the level of national income(nominal) there will be a decline in the transaction demand for money by Rs. 80 crore.

The functional relationship between transaction demand for money(M^d_t) and the level of nominal national income (**Y**) is depicted in Figure 18.1. In this figure, the transaction demand for money is shown on X-axis and level of nominal national income (**Y**) is shown on the Y-axis, At OY_1 level of national income, OM_1 money will be demanded to meet transactions demand and at the OY_2 level, OM_2 amount will be demanded. This means that as the level of national income (nominal) increases from OY_1 to OY_2, the additional transaction demand for money would be $M_1M_2 = OM_2\text{-}OM_1$.

Fig. 18.2

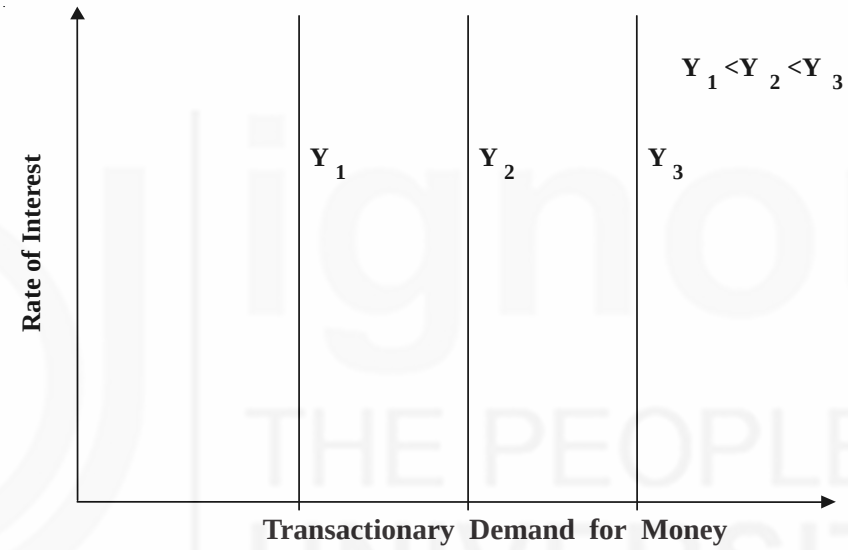


Fig. 18.2 : This figures emphasizes the idea that rate of interest does not affect *transactionary demand for money.* Y_1, Y_2 and Y_3 represent different levels of income and money needed by people at these levels is constant, for each level, separately irrespective of the rate of interest.

Note: Y indicates level of Income

According to transaction demand theory, the determinant of demand for money is the level of national income (nominal). The transaction demand for money is not affected by the rate of interest. Figure 18.2 shows three different level of National income(nominal) where $Y_1<Y_2<Y_3$. This figure illustrates the relation, or rather lack of relation, between rate of interest & transaction demand for money. Here vertical axis depicts the rate of interest and horizontal axis depicts transaction demand for money. The fact that each curve is a vertical straight line means that the rate of interest does not affect the transaction demand for money.

18.5.2 Precautionary Demand for Money (M^d_p)

The Precautionary demand for money arises out of the need for any contingent payments/expenditures. Individuals and firms desire to hold cash balances for covering events of a more uncertain nature like accidents, prolonged illness, sudden change in technology forcing firms to replace machinery to stay competitive. These are referred to as **precautionary demand for money**. Like the transaction demand for money, precautionary demand for money is also closely related to the level of income. At the higher level of income, individuals and firms may keep more cash balances for meeting

unforeseen situations. Thus, the precautionary demand for money is also a function of level of **Y**:

$$M^d_p = g(Y)$$

Keynes aggregated transaction and precautionary demands for money and pointed out that these two demands are a stable function of the level of national income(nominal). The rate of interest as an important determinant of demand for money enters through the third motive, the speculative demand for money.

Check Your Progress 2

1) What are the important assumptions of classical Quantity Theory of Money?

.....

.....

.....

.....

2) The transaction demand for money depends on three factors. What are they?

.....

.....

.....

.....

18.5.3 Speculative Demand for Money, M^d_{sp}

In addition to working as a medium of exchange, money also serves a role of store of value. The speculative demand for money is the demand for money as an asset or as a store of value. This is considered by Keynes, as 'Liquidity Preference Proper'. This was truly novel and revolutionary element of Keynes' theory of demand for money. Through it Keynes made (a part of) the demand for money a declining function of rate of interest. The speculative demand for money constitutes the main pillar of Keynes' revolution in monetary theory and Keynes' attack on Quantity Theory of Money.

Fig. 18.3

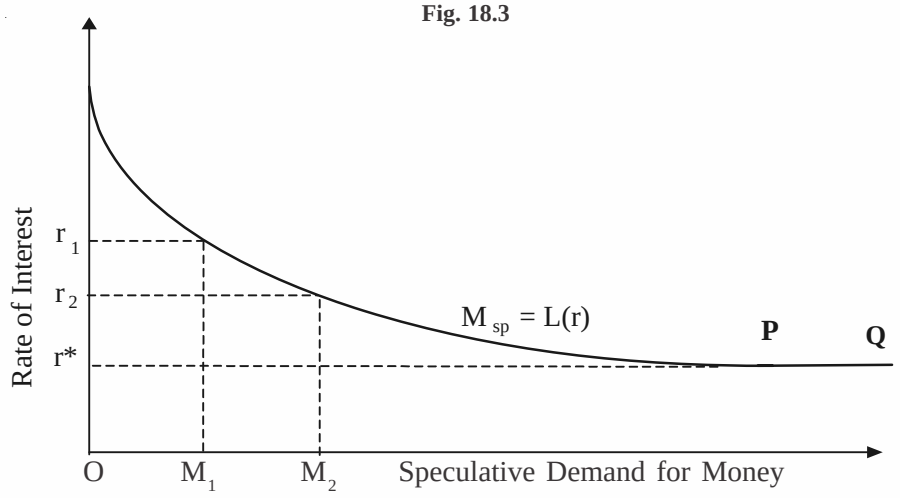


Fig. 18.3: Sepculative Demand for money is a function of rate of interest. At very low rate of interest, r*, the demand for money function becomes parallel to horizontal axis-the stretch PQ. This is known as *liguidity trap*.

The speculative demand for money arises from the speculative motive for holding money. This arises from the variability of interest rates in the market and uncertainty about them. For simplicity, Keynes' assumed all securities (bonds, shares etc.) to be of only one type, i.e., perpetual bonds. These perpetual bonds are the only non-money financial assets, which compete with money in the asset portfolio of the public. Money doesn't earn its holders any interest income but its capital value. Bonds on the other hand, yield interest income to their holders. But this income can be more than wiped out if bond prices fall in future. We can show algebraically that the price of a (perpetual) bond is given by the reciprocal of the market rate of interest times the coupon rate of interest (payable on the bond).

Economic agents hold a part of their wealth in the form of financial assets. In the two asset model of money and bonds (perpetual), bond prices keep increasing sometime with the change in the rate of interest. Therefore, they are subject to capital gains or losses. Thus, for a bondholder the return from bond holding per period (say a year) per rupee is the rate of interest. At the time of making decision about an investment in bonds, the market rate of interest will be a given datum to an economic agent, but the future rate of interest, gain or loss, will have to be anticipated. Hence the element of speculation in the bond as well as the money market comes in.

Keynes' asserted that as rate of interest falls, speculators find it profitable to keep more cash with them so that they can benefit from the opportunities which may arise. If, on the other hand, rate of interest goes up, then holding cost will go up. As a result they will demand less liquid cash. Thus, at high interest rate, r_1 speculators demand for cash may be zero. As rate of interest falls to say, r_1 they demand a larger quantity of money, M_1. A still lower rate r_2 indues speculation to demand M_2 amount of money. This inverse relation between aggregate speculative demand for money and rate of interest is shower in Fig. 18.3.

Keynes discussed the possibility of the existence of 'liquidity trap' which refers to a situation when at a certain very low level of rate of interest, **r**, the speculative demand for money becomes perfectly elastic (PQ stretch in Fig. 3). This will happen when at r, very low rate of interest, all the asset holders become bears so that none of them is willing to hold bonds and everyone wants to move into cash. The **r** seems as the banking system and credit institutions have to survive on the rate of interest they earn. Hence, that is an institutional rock bottom below, which the rate of interest cannot fall.

In equation form, speculative demand for money is a function of rate of interest;

i.e., $\mathbf{M^d_{sp} = L(r)}$

We have noted earlier that the transaction and precautionary demands for money depend on the level of **Y**. Moreover, this relationship is proportional one given by the proportionality factor, **k**. We can now aggregate the demand for money, which is given by

$$\mathbf{M^d = k.(P.y) + L(r)}$$

i.e., the demand for money has two components - one depending on the level of nominal income and the other on the rate of interest. This demand for money function is shown in Fig. 18.4

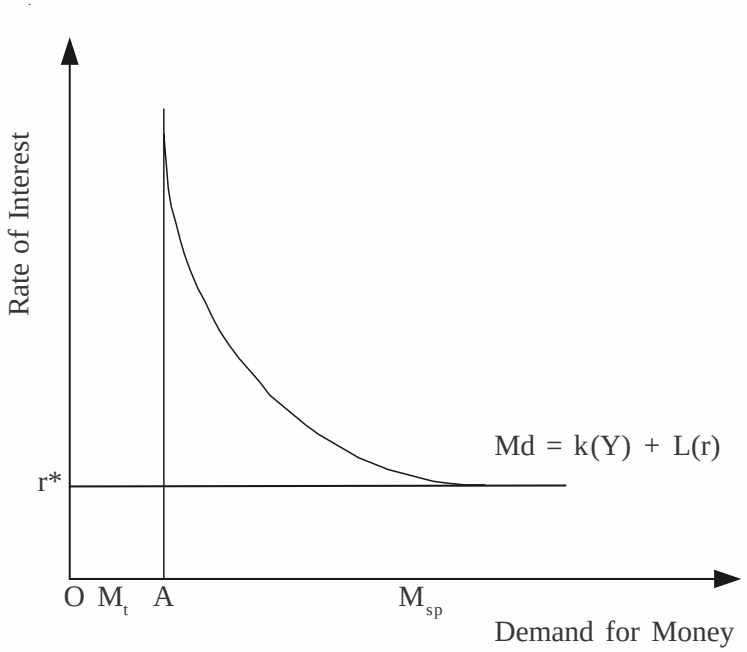


Fig. 18.4 : Total demand for money depends on both, income and the rate of interest. It is shown on interest-money plane. For each level of money income Y, we will get one such schedule. Note that OA=Transactions demand for money while rest of the segment represents speculative demand.

In this unit, supply of money($\mathbf{M^s}$) is assumed to be exogenous, i.e., it is determined outside the system. In other words, monetary authority (the Reserve Bank of India and Government of India in case of India) determines the supply of money and firms and households do not play any role in influencing supply of money.

Thus, $\mathbf{M^s} = \overline{M}$ (given)

where $\mathbf{M^s}$ = supply of money.

Now for the money market to be in equilibrium, the demand for money should equal the supply of money, i.e., actual money holding or, cash balances of the public should match the total needed or, desired balances

or,

$$\mathbf{M^d = k.P.y + L(r) = \overline{M} = M^s}$$

From this equation, it follows that combination of **Y** and **i** must be such that people's demand for money equals supply of money by the monetary authority (see Fig. 18.5)

In Fig. 18.5, $\mathbf{M^d_t = OA}$, and $\mathbf{M^s = OB}$, demand for money must equal OB so that $\mathbf{M_{sp} = L(i)}$ must equal **OB - OA = AB.** Hence the money market equilibrium occurs when $\mathbf{r = r^e}$ so that $\mathbf{M_{sp} = AB.}$ Thus, it is through the rate of interest that the money market equilibrium occurs. Note that here **P, k, y &** $\mathbf{M^s}$ are all given.

So, according to Keynes', the rate of interest is purely a monetary phenomenon, determined by the demand for and supply of money. This is in sharp contrast to the QTM of classical economists. For them rate of interest is a real variable determined by the commodity market by the equation between the supply of real saving and the demand for investment.

Impact of Changes in Supply of Money

You may call that changes in supply of money would mean the $\mathbf{M^s}$ schedule in the Fig. 18.5 would shift to $\mathbf{M^s_1}$, what would be its consequences? Increase in supply of money would push down the rate of interest in the money market, which would make the bears active. As a result of this the *speculative demand for money* would start rising till the point where the excess supply of money gets absorbed by the rise in demand for money and equilibrium is restored in the money market at $\mathbf{r_1}$ (lower rate of interest) and $\mathbf{M_1}$ (higher quantity of money) levels.

Fig. 18.5

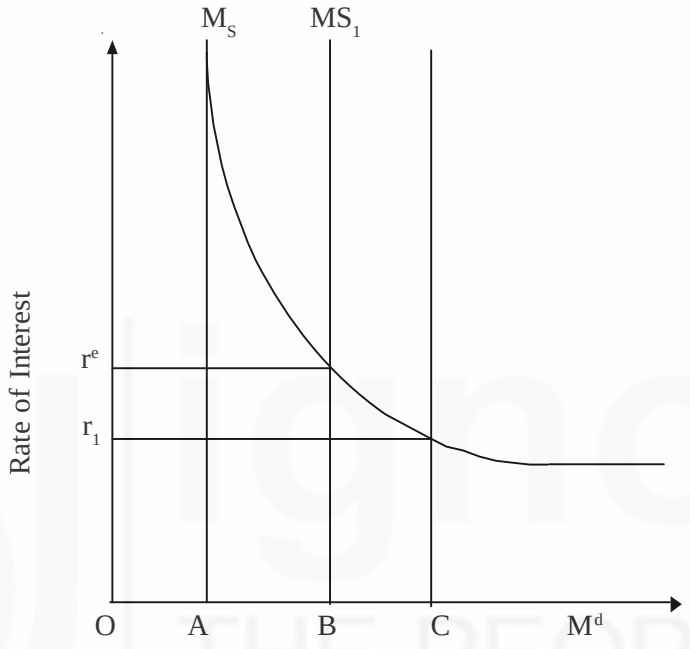


Fig. 18.5 : M^d shows aggregate demand for money and M^s is policy determined aggregate supply of money at a given level of income. Equilibrium rate of interest is $M^d=M^s$, that is r_e. Note that total money demanded =OB, out of which OA is transactionary component and AB is the speculative component. If money supply is raised ot M_s, new equilibrium rate will Be r_1. The speculative demand component will rise and become equal to AC.

Check Your Progress 3

1) Point out the main differences between classical and Keynesian theory of demand for money?

.....

.....

.....

.....

2) Explain the difference between speculative and transaction demand for money using diagram.

.....

.....

.....

.....

3) State three differences between money and bonds. Explain in four lines.

.....

.....

.....

.....

4) Briefly explain the concept of 'liquidity preference' in six-to eight lines with the help of diagram.

.....

.....

.....

.....

5) What happens when supply of money is increased?

.....

.....

.....

.....

6) Explain the concept of liquidity trap.

.....

.....

.....

.....

18.6 QUANTITY THEORY OF MONEY- MODERN VERSION

Milton Friedman, a Chicago economists, has made a restatement of the Quantity Theory of Money. His version is referred to as Monetarism or the New Quantity Theory. It is the modern version of the classical QTM. Friedman's theory of the demand for money is partly Keynesian and partly classical (or, non-Keynesian). The classical QTM is basically a theory of general price-level. For Friedman, however, it is primarily a theory of the demand for money. It is non-Keynesian in the sense that Friedman neglects completely Keynes' classification of the motives for holding money and the corresponding component demands for money.

In the classical macroeconomics, role of money is emphasized as a medium of exchange. Friedman, in his version of the theory, regards money as an asset or a capital good also (similar to Keynes' theory). According to him, wealth can be held not only in the form of (i) money and (ii) bonds as analyzed by Keynes, but also in the form of (iii) equities, (iv) physical goods (durable and semi-durable consumer goods, structures and real property), and (v) human wealth. Human wealth as embodied in human beings is the form of their potential to earn income. Total wealth is one of the key determinants of Friedman's demand for money. In practice, especially the wealth in human form it is very difficult to estimate. He, to overcome this problem, took permanent income as an approximate index of wealth.

We shall now discuss the expected rates of return on various forms of wealth, **w,** which is another key determinant of the demand for money.

i) *Money:* Friedman used broader definition of money to include not only currency, but also demand deposits(DDs) and time deposits(TDs) with the commercial banks. Money provides security and convenience to the holder. Currency is barren and provide no yield. Deposits with the banks bring in nominal interest. A change in the p-level affects the value of money: if the p-level rises (falls) money depreciates (appreciates) in value.

ii) *Bonds:* These are debentures or government securities that yields interest.

iii) *Equities:* These are shares in corporate companies. Owners of equities are the shareholders of a company. Shares bring dividend.

iv) *Physical goods:* Capital gains can be realised if nominal price of such physical goods increases. For example in India most desired physical good may be gold, land and houses as their value appreciates very fast.

v) *Human Wealth:* It is the income earning capacity of a person. This capacity is determined by the person's capabilities in terms of health, education, working conditions etc.

vi) *Change in rate of interest:* Change in the rate of interest depends on the production and demand conditions in an economy.

vii) *Expected inflation:* Expected rate of inflation depends on the expectations of growth of output level and past rates of inflation.

18.7 THREE VERSION COMPARED

Now we shall discuss one issue, which makes for the crucial distinction between Keynesian approach and Friedman's approach. This concerns the stability of demand function for money. By stability, in this case, we mean that functional relationship between money demanded and the variable influencing it is not subject to frequent changes.

According to the monetarist approach, led by Friedman, the classical demand for money function **Mv = Py** is quite stable. In other words, the velocity of money **v** is relatively stable. Given the stability of the value of **v**, the supply of money determines **Y** in the short run. Indeed, the essence of monetarism is that money supply is the major determinant of nominal income(**Y**) growth. Monetarism, like the Keynesian multiplier is a theory of determination of aggregate demand. However, according to the monetarists, crudely speaking, "*Only money matters*", in determining aggregate demand while fiscal policy does not. For, given **v**, it is **M** that determines nominal income. Suppose **v** were not stable but an upward rising function of the cost of holding money so that people want to minimize their cash balances. In other words, velocity of circulation of a unit of currency increases per period of time. The government's fiscal policy, by stimulating public expenditure may partially reduce the induced private investment by driving up the rate of interest. This results in an increase in **v** and hence, in nominal income. Thus, even without an increase in **M**, fiscal policy can raise nominal income if **v** is an increasing function of the rate of interest. The constancy of **v** as assumed by the monetarists, rules out any importance of the fiscal policy.

However, the volatility of the value of v and the existence of liquidity trap, led the Keynesians to delimit the importance of monetary policy. In the simple Keynesian system money is sidelined completely and exclusive importance is attached to the

fiscal policy in determining the level of aggregate demand of the economy.

The Keynesian policy makers were at their hey-day in fifties and sixties. The resurgence of inflation in the seventies has given the monetarists a certain respectability. The controversy between the Keynesian and the monetarist theories and policies still continues. However, Samuelson and Nordhaus note that there has been convergence to a certain degree "*from disagreement into the synthesis of modern mainstream macroeconomics*".

Check Your Progress 4

1) Explain briefly the difference between Keynesian and Friedman's version of demand for money.

.....

.....

.....

.....

18.8 LET US SUM UP

In this unit we dealt with problems associated with a barter economy, which, later on, led to the invention of money to rectify such problems. Money performs certain crucial economic functions by being a unit of account, a medium of exchange and a store of value. In **QTM**, the classical economists regarded money as only a veil beneath which all the real transactions in the economy occur. They argued that money is neutral and it does not affect real variables at all. The only advantage of money being to operate as a lubricant and thus, hastening economic transactions by reducing the costs involved in exchange. As discussed above, in a barter, this cost is considerable requiring double coincidence of demands. Keynes, in addition to the transactions demand for money described by the **QTM**, explained two other motives of holding money - the precautionary and the speculative motives. In his analysis, the classical dichotomy ceases to exist and real and monetary (nominal) variables are interrelated and interdependent. Friedman's monetarist analysis placed money at the centre of the stage in determining aggregate nominal income. Keynes was skeptical about the role of monetary policy in determining aggregate output, especially in the context of the **Great Depression**, which was the context of his 'General Theory'. He gave prime importance to the fiscal policy for the economic recovery after the Depression. Friedman, having satisfied himself about that the velocity of money is constant argued that the fiscal or, monetary policy does not help. The essence of monetarism is, thus, "*money does not matter*", i.e., it does not affect the real variables. In other words classical QTM is basically true.

18.9 KEY WORDS

Bears : They are operators in the share/bond market who expect the price of share/bond to fall.

Bond Market : It is a market concerning demand and supply of bonds, which together determine its price.

Bulls : They are operators in the share/bond market who expect the price of share/bond to rise.

Barter Economy : A direct exchange based economy without the mediation of money.

Capital gains/loss : The positive difference between the selling price of an asset and its cost price. If the selling price is less than the cost price, the difference is called capital loss.

Classical QTM : Money acts as a veil beneath which all the economic activities take place. Change in money supply only affects the nominal variables leaving real variables unaffected. The crux of the theory is: supply of money determines the level of prices.

Classical Dichotomy : It basically is the argument that real and monetary economy operate independently of each other.

Income Velocity of money : The number of times in year, a unit of currency changes hands in transactions only for final goods and service.

Nominal Income : Value of income measured at the prevailing (or, current) price level.

Near-money : It is a term which is used for those assets which are not perfectly liquid but can be easily converted into liquidity or money.

Nominal cash balances : These are the money holdings of a transactor kept in the form of cash or money.

Normal rate of interest : Introduced by J.M. Keynes, it refers to that rate of interest, which tends to prevail in the market at 'normal' conditions.

NMFAS : It is a term used to represent non-money financial assets.

Permanent income : It is a weighted average of past and present values of income.

Precautionary Demand for Money : Precautionary motive induces public to hold money to provide for unforeseen contingencies requiring sudden expenses, unforeseen opportunities of advantageous purchases. This motive is a product of uncertainties of all kinds.**Speculative Demand for Money:** Demand for money as an asset to be utilized in bond (or, share) market because holders of such speculative balances may anticipate such fall in future bond prices which will make the loss of foregone interest look relatively smaller.

Portfolios : It refers to a mixture of shares of various types, bonds, debentures, bills of exchange maintained by a transactor.

Transaction Demand : The transactions motive gives rise to transaction demand for money, which refers to the demand for cash of the public for making current transactions of all kinds. This is basically the medium of exchange role of money.

Say's Law : It sates that, 'Supply creates its own demand'. In other words, output level in an economy is supply rather than demand determined.

18.10 SOME USEFUL BOOKS

Ackley, Gardner(1977), *Macroeconomics: Theory & Policy,* Macmillan, New York.

Bhaduri, Amit(1986), *Macroeconomics, The Dynamics of Commodity Production,* Macmaillan, London.

Gupta, Suraj B.(1982), *Monetary Economics,* S.Chand & Company, New Delhi.

IGNOU(1992), *Money, Output and The General Price Level,* in EEC-01, Fundamentals of Economics" School of Social Sciences, New Delhi.

Shapiro, Eduard(1985), *Macroeconomic Analysis,* Edward Brace Juanovich, New York.

18.11 ANSWERS/HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) The problem with having horse as a medium of exchange will be that when you want to exchange it for say 1 kg. of rice. You cannot do that as it is not possible to divide the horse.

2)

i) Value of sugar is likely to fluctuate with the season, though it is divisible and can be durable but is difficult to move it to wherever required.

ii) Value of a horse is likely to depreciate as the horse becomes older and as we saw in the Section 18.2 it is not divisible. Durability also is doubtful because it has life and its health is likely to be subject to various influences. It is also not easy to take the horse everywhere where the transactions are taking place. It is very unlikely candidate as a substitute for money.

iii) Salt is a somewhat better substitute for money as it is more or less stable in value, is divisible and also durable but it is difficult to be ported for facilitating the transactions due to its weights especially for the high-value goods.

iv) Idli is very poor substitute of money due to its perishable nature. Also it is neither stable in value nor portable. It satisfies only divisibility function of money

	STABLE	DIVISIBLE	DURABLE	PORTABLE
Sugar	X	√	√	√
Horse	X	X	X	X
Salt	√	√	√	X
Idli	X	√	X	X
Gold	√	√	√	√

UNIT 19 INFLATION AND UNEMPLOYMENT

Structure

19.0 OBJECTIVES

In this unit you will learn about prices and inflation, a phenomenon, which we come across in our everyday life. The aspects that we would look at are:

- prices and price levels, and their measurement;
- meaning of inflation;
- effect of inflation on various sections of society and the economy in general;
- causes of inflation;
- anti-inflationary polices for inflation; and
- relation between inflation and unemployment using the Phillips curve.

19.1 INTRODUCTION

We come across the term inflation in newspapers every day. The reason why it holds such importance is because of its adverse effects on an economy as well as people. A question that could arise at this point is in what way does inflation affect our everyday life? Let us illustrate with the help of a single household. Inflation, in simple words, is a steady rise in the prices of various goods and services. Given the level of the money income, a household consumes a group of commodities at a given price level. With inflation, the price level goes up. So with the same level of money income, this household can consume a smaller amount of the commodities than it was consuming earlier. Alternately, to maintain the earlier level of consumption this household now needs to have more money. For example, suppose the household has a monthly income of Rs.100, consumes the entire income on a single commodity A and does not save anything. If the price of commodity A is assumed to be Rs. 4 then the household consumes 25 units of A in a month. Now suppose, the price of commodity A goes up from Rs.4 to Rs.5, the household will be able to consume only 20 units of commodity A. To maintain the level of consumption at 25 units of A per month, the household needs to have a monthly income of Rs. 125. Thus, we

see that with inflation, one unit of money purchases a smaller amount of goods than what it was doing earlier. In other words, with inflation, purchasing power of money goes down.

In the example cited above, consumption of the household comprises one commodity only. But for a typical household, consumption involves a variety of goods and services. As a result, increase in the price of one commodity need not affect household consumption adversely if there is a decline in the price of some other good. Therefore, to ascertain the effect of inflation we need to take into account the change in the prices of all the goods consumed by the household. To do that we try to find the change in the general level of prices. Therefore, before defining inflation we need to understand the meaning of price and price level and the changes in these concepts.

19.2 PRICE, PRICE LEVEL AND THE MEASURES THEREOF

What are prices? What do we mean by the term price level? What is the difference between the two? And how do we measure price level? These are some of the question we try to answer in the present section.

In simplified terms price is defined as the rate at which goods and services are exchanged for money. It is the amount of money received for selling or, paid for buying, one unit of a commodity (or services) in an exchange economy.

The term price level is an aggregate concept. It relates to the price of a basket of goods and services. See that we do not refer to the price of a single commodity but to a group of goods and services taken as a whole. Therefore, when we talk of a change in the price level it is always in reference to a group of commodities. Since the prices of commodities differ, in order to measure a change in the price level of a group of commodities, it is necessary to use index numbers. More specifically, we have to use price index. Let us understand the idea of an **index number** in an elementary form.

19.2.1 Definition of Index Number

An index number is a device for comparing the general level of magnitude of a group of distinct, but related, variables in two or more time periods. A price index is used for comparing changes in the general level of prices of a group of commodities. Generally the index number refers to changes in the prices obtained over time. It is expressed by putting a particular period (called the base) equal to 100 and the price level for other periods are expressed relative to this base. For example, when we say, the whole sale price index has gone up this year with respect to last year, we are taking last year price level as the base or, the reference point = 100. With respect to it we measure the change in the price level this year.

The *price relative* of an individual item is the ratio of its current price to its price in a base period. The simplest price index for a given commodity can be expressed as

$$I_{t,o} = 100 \ (p_t / p_o) \ \ldots\ldots\ldots\ (1)$$

where p_t and p_o denote prices in the current period 't' and the base period '0' respectively.

For instance, if price of a kilo of potato goes up from Rs. 8 in 1995 to Rs. 10 in 1996, then the price index in this case would be:

$I_{1995, 1996}$ = 100 (10/8)
= 125

This index shows a 25 per cent increase in the price of a kilo of potato. In other words, you need 25% more money to maintain your consumption of potatoes at the same old level.

Check Your Progress 1

1) What is meant by prices? What does the price level imply?

.....

.....

.....

2) What is an index number?

.....

.....

.....

19.3 INFLATION DEFINED

With the background of prices and price level in view we go on to the definition of inflation. We mentioned earlier that Inflation is defined as a persistent rise or, a tendency towards persistent rise in the general level of prices. The adjective 'persistence' has to be taken note of. The reason is, if price level goes up today but again falls tomorrow then it may not imply inflation, but only short-term fluctuations in prices. The term general price level is also important since, over a period of time, prices of some commodities may have gone up while some may have actually fallen. As a result, on the whole, the average of these prices may remain constant or even go down. Similarly if the price of a group of commodities, which constitute a small fraction of the total value of output of the economy, would go up, then again it might not be inflationary as such. That is, the effect of rise in prices of such commodities might be too small so as to affect the average price level of all the commodities. Thus we see that inflation is a macroeconomic phenomenon and is not concerned with the rise in the price of a particular commodity, or, a small group of commodities.

Another aspect of inflation is that it need not be open. That is, one would not see any changes in the quoted prices of certain goods. This can happen in a controlled economy where rise in prices of essential commodities are artificially suppressed. In India, goods sold through the public distribution system (PDS) are sold at administered prices, which are maintained at a level much below the prices in the open market. This is known as ***suppressed inflation***, as commodities sold in the ration shops may be available at a cheaper rate, but carry a higher price in the market.

In Section 19.1, it was pointed out how inflation is likely to affect a household with fixed money income. In many cases, however, some of the income classes actually benefit from inflation or at the least may remain unaffected by it. The next section discusses how this happens.

19.4 EFFECTS OF INFLATION ON SOCIETY AND ECONOMY

Inflation affects various segments of the society in different ways. It is the poor and the fixed income earners who are almost always the worst affected. For instance, a large proportion of the Indian population are daily-wage earners who work on other people's farm or in small factories. They are employed on the basis of a daily wage or a piece-rate system. Given the huge army of unemployed in our country and the paucity of a employment opportunities (they arise during harvesting and sowing, in case of agriculture), those willing to work far exceed the number that can be employed. As a result, employers almost always find adequate number of workers willing to work, howsoever small the wage rate offered. Workers have very little bargaining power vis-à-vis employers. In times of inflation the unemployment factor plays a crucial role, as the workers are unable to bargain for higher wages to offset increase in prices. Though minimum wage rates are fixed by the government, they are revised after considerable gaps and in many cases are not implemented at all. Thus a worker getting a fixed rate of Rs 12 per garment stitched, would lose in times of inflation as purchasing power of Rs.12 would be declining continuously. The poor have an added disadvantage as they rarely have any savings to fall back upon in times of adversity.

There are others in the society who either gain during inflation or at least manage to maintain the same level of real income. Organized working class like government employees manages to keep pace with inflation to a large extent as their salary is indexed to inflation. Businessmen and entrepreneurs, who can raise prices of their goods and services, may sometimes gain from inflationary situations marked by scarcity of essential services. Thus we see that inflation affects the poor, the fixed income earners and the unorganized working class much more adversely than any other section of the society. On the whole, inflation redistributes income in favour of the rich ?making the rich richer and the poor poorer.

19.5 TYPES OF INFLATION

On the basis of the severity of inflation or, the rate of acceleration in prices we can divide inflation into three different types, namely, moderate, galloping and hyperinflation.

When the general price level increases slowly but steadily, it is known as ***moderate inflation***. Moderate inflation remains within a single digit level of less than 9 per cent. You can say that the increase in the price level stays within 'limits'. There are no major uncertainties regarding the price level in the future.

Steady and fairly high rate of increases in the general price level is known as ***galloping inflation***. The rate of inflation runs into two digits (20 per cent, 40 per cent etc.) and sometimes even as high as three digits (i.e., 200 per cent). Some Latin American countries like Brazil and Argentina had experienced inflation rates of over 100 per cent in the 1970s.

Hyperinflation is characterized by astronomical increases in the annual rate of inflation. There have been cases in history when the price index rose from 100 to 10,000,000,000 within a year and a half! In such situations, money ceases to be a store of value as well as a medium of exchange. The most recent example of hyperinflation is perhaps the case of Brazil in the latter half of the 1980s.

Check Your Progress 2

1) Does inflation affect various sections of a society in a similar way? If not, explain why?

.....

.....

.....

2) What are the various types of inflation? What are the criteria used for differentiating between them? Explain.

.....

.....

.....

3) What is the difference between open and suppressed inflation?

.....

.....

.....

19.6 CAUSES OF INFLATION

The causes underlying inflation can be generally divided according to the source through which inflation originates. As we learnt in price of a commodity is determined at the point at which its supply equals demand. This is known as the equilibrium price. If the demand goes up, price of the commodity would go up in order to restore the equilibrium. So is the case when there is a fall in supply of a commodity. In either case the price of the commodity goes up till the supply and demand are equalized. But the source of the change in one case originates from the supply side while in the other from the demand side. Sometimes, rise in the cost of production pushes up the supply schedule. This again leads to a rise in prices.

So, depending upon initial process, we can identify two types of inflation: **Demand-pull inflation** or, *demand-side inflation*; and **Cost-push inflation** or, *cost-side inflation*. Note that here, we refer to aggregate demand and aggregate supply in the economy.

19.6.1 Inflation: The Demand-Side

Factors, which influence an increase in aggregate demand with no change in the level of aggregate supply, can be said to cause demand-side inflation. These factors can be an increase in government spending, a decrease in savings rate and a lower rate of taxation, which leave greater amount of money or, increased disposable income with the public, and increase in money supply. Let us examine, how each of these factors generates inflationary tendencies.

I) **Inflation Caused by Increase in Government Spending**

Suppose the government decides to build roads. In the process many unemployed get employment and earn an income. This increases the number of people who have money to spend. With no change in aggregate supply in the economy, a situation of excess demand arises. There are two ways in which this excess demand can be met, viz., by increasing the production and supply of commodities or, by increasing the price level, which would then have negative impact on the demand. In the short run, more often than not, it is difficult to expand supply. Hence

the price level increases to equilibrate the aggregate supply and aggregate demand. An increase in demand arising out of the increased government spending could be depicted by a shift in the **aggregate demand** (AD) curve as shown in Fig. 19.1. In the figure the **aggregate supply** (AS) curve is shown as a vertical line implying that we have taken aggregate supply as fixed in the short run. Thus with unchanged supply, the national real income remains at Y and only the price level goes up to equilibrate demand and supply.

Fig.19.1

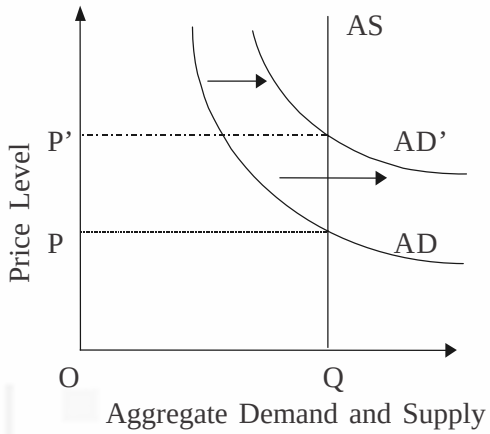


Fig.19.1 : Initially aggregate demand is AD and aggregate supply is AS all of OQ quality is purchased at OP price. Rise in government spending pulls OP the aggregate demand to level AD'. But in the long run, output cannot rise. Therefore, supply remains fixed at OQ level. People compete with one another to buy their requirements and in the process bid up prices to new and higher level OP'

II) **Keynesian Inflationary Gap**

A related, but a slightly different type, is the Keynesian inflationary gap. Recall that in Keynesian theory, investment plays the vital role in determining the national income. In the Keynesian system, the economy is divided into three broad sectors, namely, the household sector, the government sector and the private sector. The households earn money by offering their labour and other factor services and consume a part of it and save the rest. The private sector produces goods and services, earns profits and invests a part of that, say, for buying machines. The government sector gets revenue from these two sectors by way of taxes and spends it on building of roads, public services and so on. The total income of the economy or the national income consists of the goods and services produced by the above sectors. Expenditure, on the other hand, is the aggregate of consumption, C, investment, I, and government spending, G. Equilibrium in the Keynesian system is obtained where the income, Y, earned in the economy equals the expenditure on it. Assuming that there is no government spending, expenditure would then constitute consumption, C, by households, plus investment, I, by firms. Diagrammatically, equilibrium is obtained where the expenditure line, **C+I+G**, intersects the 45^0-degree line (depicting **income = expenditure**), as shown in Fig. 19.2. If government expenditure is raised, the point at which the expenditure line **C+I+G'** intersects the 45^0 degree line depicts an increase in the national income level. However, suppose for various reasons, the income level cannot be increased. Then we have a situation of excess demand equal to MN, which will be purely inflationary. MN is known as the ***inflationary gap***

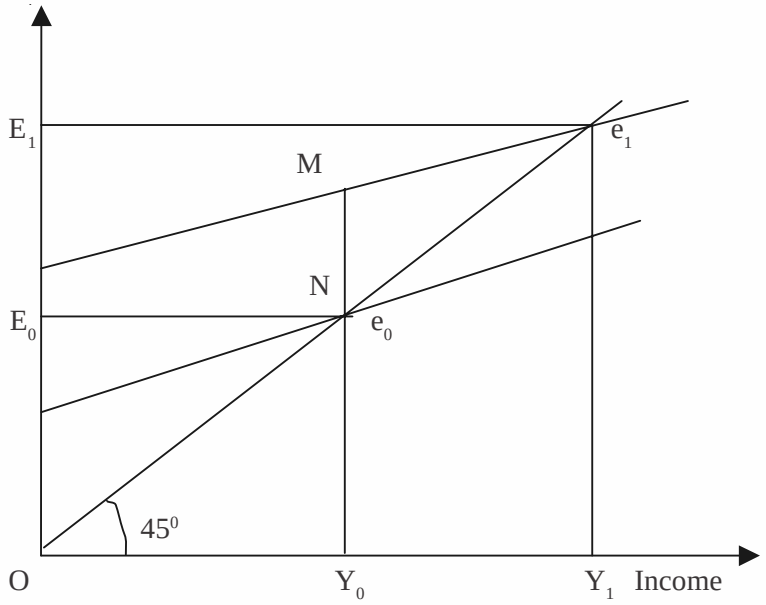


Fig.19.2 : A jump in government spending from G to G' pushes up the aggregate expenditure line from C+I+G to C+I+G'. Initially, e_0 was the equilibrium point and Y_0 was the equilibrium level of income and E_0 was the level of expenditure. But new equilibrium at e_1 means the aggregate expenditure and money income ought to be E_1 and Y_1 respectively. If, the real income cannot rise along with money income to new higher level (from Y_0 to Y_1) only the price will rise. Hence MN will be called the *Inflationary gap*.

III) **Inflation due to Increase in Money Supply**

The above argument holds when there is an increase in the money supply. With increased money supply, there is more money to spend at the disposal of general public. Again a situation of excess demand arises, i.e., a situation of disequilibrium and general price level goes up to restore equilibrium in the system. In this case there is no change in the aggregate supply curve as depicted in **Fig. 19.1**.

19.6.2 Inflation: The Supply-Side

Inflation arising out of movements in the aggregate supply curve with the aggregate demand curve remaining unchanged is known as supply-side inflation. Cost-push inflation, profit-push inflation and supply-shock inflation are three variants of this idea.

I) **Cost-push Inflation**

Cost-push inflation arises when either the labour unions (or the firms) exercise their market power to increase the wage rate (or the price of their products),. With an increase in the wage rate, producers find that the labour cost per unit of output have risen, and they respond by increasing the prices of goods to cover the higher cost of production. The workers, faced with higher prices, demand still higher wage rate, to which the producers respond by increasing the price of their commodities. A series of increase, in wage rates leads to a series of increase in price. This kind of inflation is known as ***wage-push inflation***. When the firms increase the price of their products to increase their profits, there is a demand for higher wage rate by the workers. Higher cost of production due to increases in the wage rate and prices of inputs makes the producers raise their prices further. Again a series of increase in the wage rate leads to a series of increase in the prices. This kind of inflation is known as ***profit-push inflation***.

In both the cases, each possible level of output will be supplied at a higher price level than before. As shown in Fig. 19.3, the *aggregate supply* curve (in this case we have taken a curved supply curve for analytical convenience) moves inwards showing that for the same level of output Y_1 the price now being charged is P_2, which is higher than P_1. Note, in this case, unlike in demand pull type of inflation, the output level goes down from Y_1 to Y_2.

A pre-requisite for inflation due to increases in the wage rate is that of unionisation of labour. In India labour is not organized in all sectors and there is very little empirical evidence showing inflation arising out of increases in the wage rate.

Similarly a pre-requisite for the firms to increase the prices is that the firms must be operating in an imperfect market. A firm, which has many competitors, would be unable to increase the price of its products because of the fear of losing its customers. On the other hand, fear of monopolistic or oligopolistic firms can increase their prices without the fear of losing out on demand.

Fig. 19.3

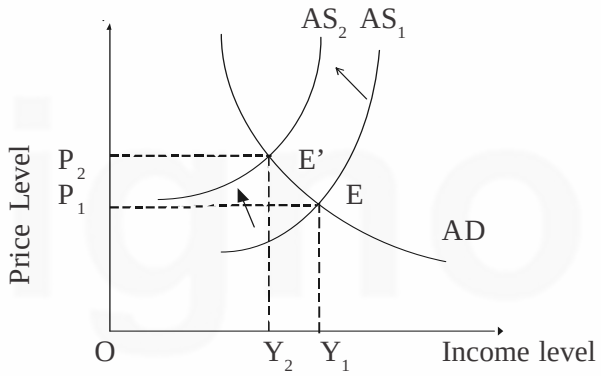


Fig. 19.3 : A rise in the cost of production pushes up the aggregate supply schedule to AS_2. In other words producers expect the higher price for the same level of output to supply. they were ready to sell Y_1 at P_1 earlier, but now demand P_2 level of price for it. The new equilibriun will take place at E' - where only Y_2 level of goods and services is produced and sold at P_2 price level. Note that equilibrium income Y_2 is less than Y_1 in pre-inflation income

II) **Supply-shock Inflation**

Inflation can arise out of either an unexpected or unforeseen sharp fall in the supply of commodities or a rise in the prices of commodities. Reasons for a situation lie out of the control of either the firms or the workers. It is known as supply-shock inflation. For instance, a crop failure due to an unfavourable weather condition would give rise to an all round shortage and lead to increase in the general price level. The above can be an example of the supply-shock inflation. Similarly, in 1973 and 1979, when oil prices were unexpectedly raised by the OPEC all the economies world wide experienced a massive rise in the general prices.This is another manifestation of supply-shock inflation.

19.7 STRUCTURAL INFLATION

Theories of inflation discussed so far have all been developed with particular reference to the developed countries. In most cases, they do not have the same applicability to inflationary experience of developing countries like India. Unlike developed countries of the West, the developing countries are characterized by a lack of adequate resources like capital, foreign exchange (for essential imports like machinery and technology), land and infrastructure (roads, railways, power

etc.). Further, over-population with the majority depending on agriculture for their livelihood means that there is a fragmentation of the land holdings. There are other institutional factors like land-ownership, technological backwardness and low rate of investment in agriculture. These features are typical of the developing economies. '**Structural theory of inflation**' explains inflation in the developing economies in terms of the structural features. Let us see, how these factors work.

I) **Food Shortages**

Majority of population in the developing economies live in the rural areas and depend on agriculture for their livelihood. With development, say, building of some new industry, some people get employment outside of agriculture and they settle down in urban areas. But, due to the various *structural features* such as highly unequal distribution of land-ownership and tenancy, technological backwardness and low rates of investment in agriculture, inadequate growth of the domestic supply of food in correspondence with an increase in demand arising from increasing urbanization and population prices increase. Further, the extreme dependence of agriculture on weather produces an acute shortage of food from time to time due to droughts, floods, etc. In years of food shortages, the prices of food grains increases very fast. Food being the key wage-good, an increase in its price tends to raise other prices as well. Therefore, some economists consider food prices to be the major factor, which leads to inflation in the developing economies.

II) **Scarcity of Foreign Exchange**

The industrial development of the developing economies requires a heavy import bill on account of import of capital goods, essential raw materials, and in several cases even food grains and other consumer goods. While the developing economies have a very high import requirement, their exports to the developed economies are very small for reasons like poor quality of goods. As a result, the foreign exchange that comes into the country through exports is a much smaller amount than the requirements of the economy. Due to this, the developing economies most of the times face foreign exchange shortages. Moreover, the shortages in the domestic supply of many basic inputs cannot easily be mitigated through imports. As a result, the prices of such goods increase, and the increase spreads to other prices. The result is all-round inflation.

Other structural factors, like socially unproductive private investment in land and precious metals like gold take away a sizable chunk of resources. These resources could have been otherwise invested in new industries, new machines, new roads, better irrigation facilities for agriculture and other productive investment, which could have led to faster development of these countries. According to the **structural approach to inflation,** the above factors and similar other structural features of the developing economies can explain the price rise situations better.

19.8 ANTI-INFLATIONARY POLICIES

Now we have a fairly good idea about the causes of inflation. Let us move on to the question of its remedies. What are the possible ways to control the inflation in an economy? But for recommending a cure, an analysis of the source of the problem, i.e., whether inflation is due to demand-pull factors or cost-push factors, is important. Why that is so, is what we show below.

Fig.19.4

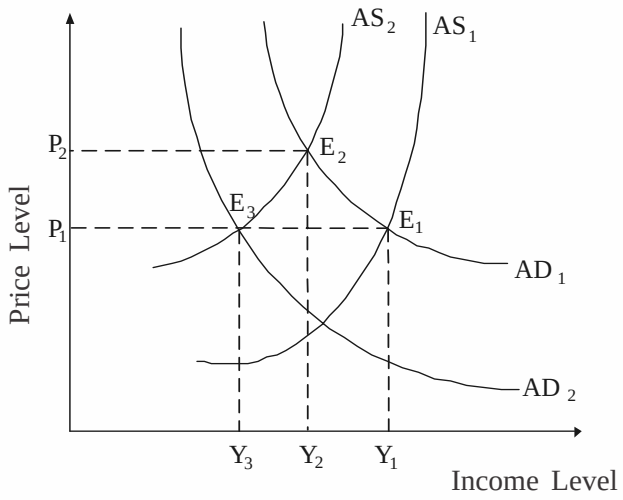


Fig.19.4 : Initial aggregate demand, AD_1 and aggregate supply, AS_1 curves intersect at E_1 where level of income is Y_1 and that of prices is P_1. However, a rise in cost of production pushes up aggregate supply to AS_2. This results in a rise in rise in prices to P_2. At such higher prices, only Y_2 level of income will be generated as the consumes cannot afford to buy larger amounts of goods and services. Adoption of demand management policies under such circumstances, will push down aggregate demand to AD_2. It may restore the prices to old level P_1 yet level of income at Y_3 is at the lowest level of income.

Suppose, inflation is of the demand-pull variant. This means that a higher level of disposable income with the public with no change in the supply function has given rise to inflation. So, in such situations, inflation can be controlled by simply reducing the extra disposable income in the hands of the people.

The Government can do this by either decreasing the money supply, or by increasing the incentives for savings by giving tax exemptions on savings. Reducing the money supply directly lowers the extra funds available and thus helps in bringing down the demand. Various incentives increase savings, reduce consumption of those who save and thus bring down the level of aggregate demand and the price level. Given the aggregate supply, a rise in the aggregate demand raises the price level. But the demand-regulating measures push back the aggregate demand to AD and restore the old price level.

If the source of inflation lies in a decrease in the aggregate supply, then prescribed policies above would hamper economic situation on the whole. Fig.19.4 shows this situation as well. If policies, which reduce aggregate demand level, are adopted, then the price level would go down. However, equilibrium level of output would also follow a similar trend. Such policies, therefore, would only decrease the demand for labour and create an all around increase in unemployment levels. For tackling supply-side inflation, what one needs to do is to focus on the supply side. Though the government cannot do anything to increase supply in the short run, it can adopt policies, which nullify the inflationary effect arising out of increase in cost of production.

One of the possibilities is to decrease taxes like sales tax and excise duties at various levels, which helps bring down the cost of production. The firms then can reduce the prices of their products and are able to sell larger quantities in the market. In case of situations like crop failures government can augment food supplies by releasing larger stocks through **public distribution system** and bring down prices. Outright sale (from buffer stocks) in the open market can also have similar effect. The extent, to which such deflationary policies are actually effective, depends on various other factors, which we won't go into here.

19.9 DEFLATION

Deflation is a situation where prices fall continuously or have a tendency to fall. This can arise when the aggregate demand is lower than the aggregate supply. Thus, deflation is characterized by a decrease fall in output, increase in unemployment and general slowing down of the economic activities. The Great Depression from 1929 to 1933 in the capitalist countries is an example of an acute deflation when the prices crashed, unemployment catapulted to astronomical heights and the income of these countries fell sharply.

19.10 STAGFLATION

In the Keynesian system an inflationary gap in the short-run would lead to an increase in the real national income and hence employment. Thus while the price level goes up, so does the output, which acts as a dampener on upward movement of the prices. Thus, inflation in the Keynesian system would be accompanied by an increase in the level of real output and employment. However, in 1970s, several countries experienced a peculiar situation. There were rising rate of inflation, which was accompanied, by not only rise in unemployment but also falling or stagnating output. This type of phenomenon is called stagflation.

Suppose the prevailing rate of inflation is 6 per cent. This prevailing rate builds expectations in the minds of people about its level for some time in future. Such an expectation determines the money wage rate to be negotiated by the labour unions and employers. The employers in response to increase in the money wage rate increase their prices, which then increase the rate of inflation in the present period itself. As a result of this unexpected increase in inflation (since inflation rate now is greater than 6 per cent) the labourers find that the earlier negotiated increase in money wage rate is not sufficient to protect them against the falling purchasing power of money and demand a still higher wage rate. Such increase in the money wage rate, to compensate for the new level of inflation, would result in the firms increasing their prices.

Thus the expectation about the future price level plays a crucial role in determining the actual price level today. And according to this view, the simultaneous impact of the remedial policy measures adopted and the lag in the adjustment between the expected inflation rate and the actual inflation rate results in stagflation.

As we saw in Section 19.8, the remedial policy measures initiated would not only bring down the prices but the national income as well. But the effect of such policies is not felt instantaneously. While such policy measures are in the process of exerting their impact, expected inflation is still catching up with actual inflation. That is, the upward pressure on the inflation rate exerted by a slower growth rate of nominal income would be a consequence of the restrictive policy measures followed to control inflation. Therefore, a situation is seen when stagnation in the output level goes hand in hand with the rising inflation.

19.11 INFLATION AND UNEMPLOYMENT: THE PHILLIPS CURVE

In this section we study the relation between the inflation rate and the unemployment rate and see how one affects the other.

For long economists have tried to find a relation between inflation and unemployment. A British economist, A. W. Phillips, studied the relation between the wage inflation (tendency of wages to rise continually) and unemployment. By studying data for the

British economy for the years 1861-1957, he found a correlation between the two, which seemed to suggest that the rate of unemployment and the rate of wage inflation are inversely related. Wage inflation without any change in other factors, we know, would lead to pure inflation of the cost-push variety. This implies a positive relation between the rate of inflation and the wage inflation. On the other hand, an increase in wage rate reduces the demand for labour and thus unemployment increases. This implies an inverse relation between inflation and unemployment. A direct relation between wage inflation and rate of inflation and an inverse one between wage inflation and unemployment imply an inverse relation between rate of inflation and unemployment. We can put these discussions as :

wage rate α **inflation rate**

wage rate α **(1/unemployment rate)**

Therefore: **inflation rate** α **(1/unemployment rate)**

where α **= sign of proportionality**.

There are two ways in which this can happen: When unemployment levels are low i.e., when fewer people are unemployed the unionized labour is in a better bargaining position to demand for an increase in the wage rate. On the other hand, the low unemployment rate and relative scarcity of labour are ordinarily times of buoyant demand and abundant profits. Thus, the firms would usually grant the demand for increase in wage rate than face the possibility of strikes and closure of such profitable production. In the reverse case of high unemployment and low profits neither would the labour unions press for higher wages nor would the firms give in to such demands.

The second explanation is the excess demand for labour. In periods of boom characterized by growing demand and profits. The demand for labour is likely to exceed its supply. As a result, the wage rate may goes up leading to increase in the rate of inflation. Therefore, as unemployment goes down inflation rate goes up. Note, these two possible explanations, however, are not mutually exclusive as both the factors can work together in periods of either boom or slowdown.

Thus the Phillips curve is a downward sloping curve with the horizontal axis representing unemployment rate and the vertical axis representing the inflation rate. It shows a trade-off between inflation and unemployment, implying that government policies can reduce the unemployment rate only by accepting a higher inflation rate and vice-versa. (Fig. 19.5)

However, the acceptability of the Phillips curve explanation suffered a setback in the 1970s when the capitalist countries experienced situations where rising inflation rate was accompanied by rising rates unemployment, i.e., a situation of stagflation.

There are many criticisms against the Phillips curve, but we enumerate just two of them.

The first argument says that Phillips curve holds true only in the short run and is not valid in the long run. This is because, an increase in the prices with the nominal wage rate remaining the same, implies that larger profits can be made for each unit of the commodity sold. And even bigger profits are possible if larger quantities of the commodities can be sold. Therefore, when prices go up, in order to reap the benefits of higher prices, producers expand the production of goods and services whose prices have gone up. Greater production means more people are employed. This happens in the short run. This is in contrast to the Phillips Curve relationship since increases in inflation leads to fall in unemployment.

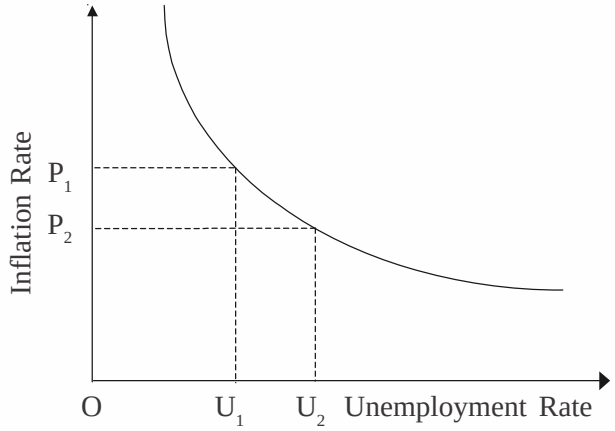


Fig.19.5 : The Phillip curve shows inverse relationship between rate of unemployment inflation. As unemployment to rise from U_1 to U_2, the rate of inflation comes down brought down from P_1 to P_2. Conversely, higher rate of inflation is the price a society has to pay to bring down the rate of unemployment

In the long run, things are very different. An increase in prices with the nominal wage rate remaining constant implies that the real wage, which is the ratio of nominal wage rate to general price level, declines. Sooner or later, workers realize that their real wages have gone down and they demand higher nominal wages in keeping with the inflation rate. An increase in nominal wage rate implies that the cost of production goes up and hence it is no longer as profitable for producers to raise production. Therefore, they cut down on production and consequently employment too would go down. Thus, while the inflation rate remains at the present level whereas, the unemployment rate goes up. Thus, the theory of trade-off between inflation and unemployment no longer remains valid. But we must not loose sight of the fact that the Phillips Curve analysis is based on the experience of 96 years, which saw two world wars and many momentous changes in the technology profile of the world.

The other argument runs as follows: It says that the Phillips curve ignores the role of capital in sparking off inflation and unemployment. Under capitalism, capitalists' aim to maximize profits leads them adopt new technologies, which are generally more capital-intensive. This tends to lower the need for labour and hence increases the rate of unemployment. At the same time these capitalists follow an aggressive mark-up pricing policy with a view to increase their profit margins leading to an inflationary spiral. Thus the theory of trade-off between inflation and unemployment no longer remains valid.

This argument seems to assume that there shall be no expansion in the overall demand for goods and services when the technological changes are raising productivity per worker by substantial amounts! Historical experience of the rising consumerism and mass production would not justify such an assumption.

Check Your Progress 3

1) What are the various causes of inflation? Explain with diagrams

.....

.....

.....

2) What is meant by Stagflation?

.....

.....

.....

3) What is deflation

.....

.....

.....

4) Give a reason why Phillips curve might not be valid.

.....

.....

.....

19.12 LET US SUM UP

In this unit we learnt that price is the rate at which goods and services are exchanged for money. The 'Price level' is an aggregate concept. Increase in the price level denotes less of the goods and services for one unit of money. Since the price level is concerned with many differently priced goods and services, use of index numbers required to measure changes in the price level. An index number is a device for comparing the general level of the magnitude of a group of distinct, but related, variables in two or more time periods. There are different types of measures for price index.

We also learnt that inflation is a sustained rise, or a tendency towards it in the general level of prices. Inflation affects the fixed income earners and the poor most adversely. It also slows down the growth rate of an economy. The types of inflation that a country may face range from moderate to hyperinflation depending on the rate of inflation.

Inflation can arise from either the demand side or the supply side. The former would include factors like increase in government expenditure, lowering of tax rates and increase in supply of money. The latter would include factors like increase in the prices of basic inputs or mass consumption goods, or increases in wage rate. We also read that these causes of inflation in many cases do not explain experience of the less developed countries. Structural inflation tries to do that by taking into account structural peculiarities typical to these countries.

For applying anti-inflationary policies it is necessary to know the source of inflation, i.e., demand-side or supply-side, for getting the best results.

We have learnt that deflation is a situation where prices are continuously falling or have a tendency to fall. Stagflation, on the other hand is a situation where inflation is coupled with a stagnating real output and a high rate of unemployment

Finally, we have seen that the Phillips curve tries to trace out a relation between unemployment and inflation. But other than some decades over which the Phillips curve relationship applies, it is more or less redundant. In the developing countries unemployment has come about historically with little to do with inflation.

19.13 KEY WORDS

Deflation : is a situation where prices fall continuously or have a tendency to fall. This generally happens when the

aggregate demand falls short of the *aggregate supply* in the economy.

Gestation Period : The period between the time when investment in a project is made and when production starts.

Hyperinflation : It is characterized by an astronomical increase in the annual rate of inflation. This generally happens when people loose faith in money as a medium of exchange.

Inflation : It is defined as a persistent rise or a tendency towards persistent rise in the general level of prices. It is the rise in prices of all the commodities, which are part of the price index and should not be confused with a rise in the price of a commodity or a group of commodities such as food products.

OPEC : refers to the Organization of the Petroleum Exporting Countries. These include Indonesia, Kuwait, United Arab Emirates, Qatar, Iran, Iraq, Tanzania, Venezuela, Algeria, Libya, Nigeria, and Saudi Arabia.

Piece-rate System : A fixed amount of money is paid on the basis of the number of units of a good produced/stitched/collected etc.

Price level : is an aggregate concept and it relates to the price of a basket of goods and services.

Stagflation : It is a situation when the rising inflation is accompanied by a falling or stagnating output. This generally happens in the conditions of some *structural rigidities* (like shortage of infrastructural facilities like transportation, power, etc.) in an economy.

Suppressed inflation : Inflationary situations, which do not get reflected in the quoted prices. This could be done by subsidizing the commodities under question. Governments do this many times to protect the weaker sections of the population to protect them from a fast rise in prices.

19.14 SOME USEFUL BOOKS

Bhaduri, A., *Macroeconomics: The Dynamics of Commodity Production*, Macmillan 1986, Chapter 3.

Gupta, S. B, *Monetary Economics: Institutions, Theory and Policy,* S. Chand and Co. 1989, Chapter 14.

Gupta, S. B., *Monetary Planning for India,* 1979, Oxford University Press, Chapter 3 and Appendices.

Shapiro Edward, 1984, *Macro Economic Analysis* (5th Edition); Galgotia Publications: New Delhi

19.15 ANSWERS/HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) Price is defined as the rate at which goods and services are exchanged for money. It is the amount of money received for selling or paid for buying one unit of a commodity or service in an exchange economy. The term price level is an aggregate concept as it relates to the price of a basket of goods and services. It does not refer to the price of a single commodity but to a group of goods and services taken as a whole. Therefore, when one talks of a change in the price level it is always in reference to a group of commodities.

2) An index number is a device for comparing the general level of the magnitude of a group of distinct, but related, variables in two or more time periods. A price index is used for comparing the changes in the general level of prices of a group of commodities. Generally the index number refers to the changes in the prices obtained over time, and it is expressed by putting a particular period (called the base) equal to 100 and expressing the other periods relatively to 100.

Check Your Progress 2

1) Inflation affects various sections of a society in different ways. The organized working class, such as government employees, etc. have fixed income but they manage to keep pace with inflation to a large extent. The businessmen and entrepreneurs, who can raise the prices of their goods and services may sometimes, gain from inflationary situations marked by scarcity of essential services. The poor and the fixed income earners, with low bargaining power vis-à-vis their employers and little savings to fall back upon, are the ones who the lose maximum during inflation.

2) On the basis of the severity or the rate of acceleration of prices we can divide inflation into three different types, namely, moderate, galloping and hyperinflation. When the general price level rises slowly but steadily, it is known as *moderate inflation*. Moderate inflation remains within a single digit and usually the annual inflation rates stays below 9 per cent. Steady and fairly high rate of increases in the general price level is known as *galloping inflation*. The rate of inflation runs into two digits (20 per cent, 40 per cent, etc.) and sometimes even as as high three digits (e.g., 200 per cent). *Hyperinflation* is characterized by an astronomical increase in the annual rate of inflation.

3) See Section 19.3.

Check Your Progress 3

1) The causes of inflation can be divided into two main categories, namely, supply-side factors and demand-side factors. Increase in government expenditure, increase in money supply, etc., are some of the causes in the demand-side factors. Crop failures, sudden increase in prices of basic goods etc., are some of the causes falling under supply-side factors. For details look into the section on the causes of inflation.

2) Till the late sixties it was believed that inflation almost always is accompanied by an increase in employment and output. But the experiences of some countries of a rising inflation rate accompanied by a rising unemployment rate, in the early seventies, came as a blow to this belief. Situations like this where a rising inflation is accompanied by a falling or stagnating output are known as *stagflation*

3) ***Deflation*** is a situation where prices fall continuously or have a tendency to fall. This can arise when the aggregate demand is lower than the aggregate supply and as a consequence the price level tends to fall. Deflation is characterized by a fall in output, increase in unemployment and general slowing down of economic activities.

4) The Phillips curve shows a trade-off between inflation and unemployment. It says that an economy can achieve a lower rate of inflation only at the cost of a higher rate of unemployment and vice-versa. But the validity of this trade-off has been questioned with the experiences of stagflation. The validity of a trade-off between the two is because the Phillips curve ignores the role of capital in sparking off inflation and unemployment. Under capitalism, the capitalists' aim to maximize profits leads them to adopt new technologies, which are more often than not more capital-intensive in nature. This tends to lower the need for labour and hence increase the rate of unemployment. At the same time these capitalists follow an aggressive mark-up pricing policy with a view to increase their profit margins. This leads to an inflationary spiral.

ignou
THE PEOPLE'S
UNIVERSITY

Indira Gandhi
National Open University
School of Social Sciences

EEC-11
Fundamentals of Economics

Block

8

MONEY AND PRICES

Expert Committee

Mr. Kalyanjit RoyChaudhary
St. Stephens College
University of Delhi
Delhi

Mr. B.S. Bagla
PGDAV College
University of Delhi
Delhi

Mr. R.S. Malhan
St. Stephens College
University of Delhi
Delhi

Mr. R.S. Bharadwaj
Shivaji College
University of Delhi
Delhi

Dr. Gopinath Pradhan
IGNOU

Dr. Narayan Prasad
IGNOU

Dr. Madhu Bala
IGNOU

Dr. Kaustuva Barik
IGNOU

Mr. Saugato Sen
IGNOU

Prof. S.K. Singh (retd.)
IGNOU

Course Editor	Course Coordinator	Block Coordinator
Mr. Bhawani Sankar Bagla	Dr. Gopinath Pradhan	Dr. Gopinath Pradhan

Block Preparation Team

Unit No.	Unit Writer	Unit Editor (IGNOU Faculty)
18	Sh. Sunil Ashra	Dr. G. Pradhan
19	Mr. S. Sen	Dr. G. Pradhan & Dr. K. Barik

Production	Cover Design	Word Processing
Mr. Arvind Kumar	Ms. Arvinder Chawla	Mrs. Rekha Mishra
Mr. Manjit Singh		Ms. Daisy Lal

March, 2003

ISBN-81-266-0766-1

Further information on the Indira Gandhi National Open University courses may be obtained from the university's office at Maidan Garhi, New Delhi-110 068.

Printed and published on behalf of the Indira Gandhi National Open University, New Delhi, by Professor Kapil Kumar, Director, School of Social Sciences.

Laser typeset by : HD Computer Craft, EA1/75, Main Market Inderpuri, New Delhi-110012. Ph: 2583 1437; 3108 1117

Printed at :

BLOCK 8 MONEY AND PRICES

Introduction

This block discusses the role of money in the economic system. Theories of demand for money are discussed by introducing classical, Keynesian and monetarist versions of quantity theory (**Unit 18**). The problem of price rise and employment is covered in **Unit 19**. Whole range of issues centred on the functions of money, supply of money, situations of consistent rise in prices of commodities and their impact on production and employment form the subject matter of the block.

UNIT 20 TOPICS IN INTERNATIONAL ECONOMICS

Structure

20.1 OBJECTIVES

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

- differentiate between absolute advantage and comparative advantage;
- explain how international trade takes place on the basis of comparative advantage;
- discuss the Ricardian model of international trade;
- describe the gains from trade;
- define the various measures of terms-of-trade; and
- explain and analyse the balance of payments.

20.1 INTRODUCTION

In this unit we will discuss the nature of international trade and the reasons why countries trade with each other. To begin with, we will examine whether there is any difference between internal and international trade. This will be followed by two questions: why trade takes place and how it does so.

Two theories will help us explain the gains from trade. This will be followed by the concept and structure of balance of payments along with the meaning of deficit/surplus in a country's balance of payments.

In this unit our attempt is to make a very simple presentation of issues related to international trade. We have steered clear of tedious and 'difficult' derivatives while bringing the essence of concepts to you.

20.2 COMPARATIVE ADVANTAGE AS BASIS FOR TRADE

Our wants are unlimited and resources are limited. We cannot satisfy all our wants. As the resources are also limited, we try to make the best use of them both

in consumption and production. One may be capable of undertaking many economic activities or able to produce a number of goods. But his efficiency or productivity will not be the same to produce all the goods. He may be relatively efficient for some goods less efficient or inefficient for some other goods. The same is true for all individuals. In view of this, persons choose their own lines of economic activities in which they are more efficient than others. This division of economic activities is also called division of labour. What is true for individuals is true for regions or nation.

Trade, whether within a country or between countries is an **act of exchange.** We exchange one thing for another. Like individuals, countries also normally do not produce each and everything. Only under restrictive assumption of a closed economy, where we do not have external trade and economic relations, almost all the commodities are produced within a country and exchange or barter takes place within the country.

Needless to say that in the closed economic scenario, we presume above, it will result neither in maximization of production nor optimum use of factors of production. In fact closed or isolated economies are unthinkable these days. Rapid changes in transport and communication have changed the meaning of place and time. Transportation has become so quick that one can take his breakfast in India and lunch in France/Germany and evening tea somewhere else. Earlier it took many days to cover few hundred miles.

In international trade we have two important questions: Why does trade take place? And how does trade take place? We will answer these one by one.

20.2.1 Comparative Cost Advantages - Why Trade Takes Place ?

An important question is why people trade? The answer is very simple. It is human nature to exchange one thing for another. People exchange things not for the sake of exchange but for the simple reason that they cannot produce everything efficiently and at a lower cost. Similarly, a country, which has, for example, mineral resources only and limited cultivable land, is bound to import agricultural commodities. Thus, we find that like individuals, countries also differ in factor endowments. As a result, some countries can produce other commodities more efficiently at a lower cost. There is a possibility that a country can produce most of the commodities at a lower cost as compared to many others. But the level of efficiency will not be the same for all goods. In some commodities, that country may be the best, in some it may be better and still in some others it may be just good. The efficiency is not equal in all lines of production. It is of different degrees. As it will not be possible for that country to produce all commodities because of the limitations of resources and also of time factor, production of these commodities where its superiority is most marked will be undertaken. We can, therefore, conclude that trade takes place because of relative differences in costs, which result from differences in factor endowments.

Let us take an example. We take two countries and assume that they produce only two commodities. [The number of countries and commodities can be more, but specialization in production will be on the basis of efficiency, that is, relative costs of production]. Let A and B be two countries and both of them produce commodities X and Y. We also assume that with certain amount of labour and capital different units of X and Y are produced. Let the following table show quantities of X and Y that can be produced in their countries with given amounts of capital and labour.

Country	Commodity	
	X	Y
A	20	10
B	10	20

Country A can produce 20 units of X and 10 units of Y. Country B, on the other hand, with the same amount of labour and capital can produce 10 units of X and 20 units of Y. In this example A is most efficient in the production of X and B is most efficient in the production of Y. The cost ratio of X and Y in A is 1:2 and B country it is 2:1. Obviously, under these circumstances, country A will gain by producing X only and country B will gain by producing Y commodity only. Country A has absolute advantage in X and country B has absolute advantage in Y. It is called the theory of absolute advantage and was propounded by Adam Smith who assumed labour to be the only factor of production. He thought that these absolute differences in advantage are the basis of international trade.

The basis of absolute cost advantage theory is that cost ratios of commodities differ in different countries, but, it has assumed that each country is absolutely superior in one line of economic activity. This appears unrealistic. Because in a group of countries, a country can be relatively more or less efficient. It is not normal to find a country absolutely superior. Superiority is always relative. Therefore, it is more practical to consider relative or comparative differences in cost. David Ricardo gave his explanation on the basis of comparative differences in the cost of production. It is, therefore, known as comparative cost theory of international trade. Before considering comparative differences in cost, let us take one example where costs are different. Again we assume, A and B, two countries producing X and Y commodities such that

Country	Commodity	
	X	Y
A	20	10
B	10	20

In this example country A has advantage in both the commodities as 20 unit of X and 10 unit of Y can be produced with a certain amount of labour and capital. While country B, with the same amount of factors, can produce 10 units of X and 5 units of Y. Obviously country A is superior and more efficient in both X and Y commodities and country B is less efficient in both commodities. Let us now consider the ratios of commodities in both the countries. In country A, the ratio of X and Y is 2:1 and in country B also the ratio of X and Y is 2:1. Under such circumstances trade may not take place as both the countries have an equal cost ratio. In country A two units of X will exchange for one unit of Y and in country B also two units of X will exchange for one unit of Y. Although country A is absolutely superior in both the commodities and country B is absolutely inferior in both X and Y trade will not take place as no country will gain because the cost -ratios are similar or identical in both the countries. Here, we find that despite 'advantage' there will be no trade. However, specialization and trade can be mutually beneficial even for such economies if there exists substantial economies of scale in both the industries. Yet, it is not possible to pinpoint which commodity a country must specialize in.

20.2.2 Ricardo's Comparative Cost Theory

Let us take an example to explain Ricardo's comparative cost theory. As usual we have two countries and two commodities and the amount of labour needed (in hours) to produce one unit each of X and Y as given below:

Country	Hours required for Commodity X	Y
A	120	100
B	80	90

From the above example, it is clear that country A is able to produce 1 unit of X with 120 hours of labour while it can produce 1 unit of Y with 100 hours of labour. Thus, X is more expensive than Y. One unit of X will cost 120/100 units of Y. In country B, it takes 80 hours of labour to produce 1 unit of X and 90 hours of labour to produce 1 units of Y. Notice that country B has absolute advantage in both lines of production because it takes less labour in B than A to produce both X and Y. However, within B, Y is more expensive per unit than X. One unit of X costs 80/90 or 0.89 units of Y. Although country B has absolute advantage in both lines of production, each country has a comparative advantage in different goods. A has a comparative advantage in producing that good whose opportunity cost is lower in this country than in the other country. The opportunity cost of 1 unit of X for Y in country A is 120/100 = 12/10 while in country B it is 80/90 or 8/9. Thus, the opportunity cost of X for Y is lower in B than A. On the other hand, the opportunity cost of 1 unit of Y for X in country A is 100/120 = 10/12, while in country B it is 90/80 or 9/8. So opportunity of Y for X is lower in A than in B. Thus, B has a comparative advantage in producing X while A has a comparative advantage in producing Y.

We saw above that in country A, one unit of X traded for 120/100 or 1.2 units of Y, while in B, one unit of X traded for 80/90, or 0.89 units of Y. If country A could import one unit of X for less than 1.2 units of Y, and if country B could import more than 0.89 units of Y for 1 unit of X, both countries would gain from international trade.

Ricardo did not discuss about the point where the actual rate will be determined. He only explained why trade takes place. The actual rate will be determined by the mutual demand of A and B countries. This was explained by J.S. Mill who propounded the theory of Reciprocal Demand. In fact, the other question, how trade takes place was explained by J.S. Mill. Both the theories together constitute the Classical Theory of International Trade.

Before we proceed further, it will be necessary to discuss Ricardo's assumptions regarding comparative cost theory. You may note that assumptions are a necessary part of theory. They simply facilitate our analysis and formulation of the theory. But if assumptions interfere with the conclusions then we have to critically examine the nature of the assumptions. Ricardo's assumptions, for the sake of convenience, can be put into two categories. The first category will consist of facilitating assumptions. In this category we put the following: (1) Two countries, (2) Two commodities, (3) Labour cost and not money cost, (4) No transport costs, (5) Free

trade, (6) Gold standard, (7) Perfect Competition in the production of the goods. The assumption of two countries and two commodities is just to facilitate the analysis. We can increase the number of countries and commodities. Trade will take place only on the basis of comparative cost. We are giving below three countries and three commodities case :

Countries	Commodities		
	X	Y	Z
A	10	11	12
B	9	11	10
C	11	10	9

Figures in table show number of units of respective commodities that a given amount of labour can produce. Now in this case also the trade will place only on the basis of comparative advantage or disadvantage. Country A is most efficient in Z commodity, country C is most efficient in X commodity and country A and B are equally efficient in the production of Y commodity. The basis of trade will remain the same irrespective of the number of countries and commodities.

Next, the labour cost can be converted into money costs. We can express the costs of products in terms of the currency of the participating countries. Instead of giving cost in terms of units of commodities with fixed amount of labour, we can give unit cost in terms of money. The country which can produce larger units of a commodity with the given amount of labour will have lower prices (the lower cost in money terms) as compared to a country which is producing lesser number of units with the same amount of labour. The fact is that it will not invalidate the theory. Even if we consider money cost the terms of trade will remain the same. The same logic applies to transport cost. Transport costs will change the cost ratios (as we will have to add the cost of transport to money cost). Trade according to Ricardo's theory of comparative cost will take place on the basis of comparative or relative difference in cost.

If we drop the assumption of gold standard, it will not change the basis of trade. Even under paper currency, the trade will take place on the basis of cost differences.

Assumption of free trade is necessary. If countries adopt restrictive trade practices like tariff and quotas, then the normal pattern of trade is disrupted. There will be distortion. It will not be a normal situation.

Ricardo's labour cost approach has been severely criticised by many economists, particularly, Ohlin. Ohlin in his 'Interregional and International Trade' in Appendix has given his criticism of the comparative cost theory. His argument is that labour cost approach is not the correct tool to find the cost of a commodity. Nobody can dispute this. Because these days labour cost can vary roughly from 30% to 60% depending upon the nature of the commodity. Capital is equally important. Once we accept the importance of capital, rate of interest assumes importance.

If capital accounts for a larger share in the cost, then even rate of interest can also bring about a change in the cost of production, the comparative advantage will be affected. These days technology has also assumed much importance. The fact is that labour cost is not the correct index of measuring cost of production. Labour cost alone cannot determine cost. Further, labour is not homogeneous. There are different types of labour like unskilled, skilled, trained and technical. All these have different wages or rates of remuneration. It will be difficult to find comparative

costs even if we accept labour costs.

The second category of Ricardo's assumption include,

1) both the countries are almost of same size,

2) their patterns of consumption are also similar,

3) that there will be complete specialization's.

These assumptions are really very restrictive in nature. Perhaps his overriding concern with complete specialization must have been responsible for incorporation of the other two. But overall implications of these three will be that trade may not be possible if two trading partners were vastly different in geographical sizes and had different pattern of consumption.

Heckscher-Ohlin Approach : Heckscher and Ohlin tried to analyses the question of why do different nations trade in a different manner. They tried to bring in factors other than labour explicitly into the trade model. They further tried to relate specialization to the factor endowments of the countries explicitly. These models also admit possibility of less than complete specialization. Thus, several of the shortcomings of Ricardo's formulations can be addressed under Heckscher-Ohlin framework. Take, for example, considerations such as: (1) trade is not determined by labour cost differentials alone, (2) the endowments of different resources processed by the trading countries determine the direction of spcecialisation, (3) the countries need not be of the same size, (4) they need not have similar patterns consumption either and finally, (5) trade need not lead to full specialization.

The assumtions of their model is that different countries have different endowments. Each country, even when faced with same kind technology as opened to other, picks up those technique of production which uses the factors available to it in relative abundance more intensively. The factor abundence must be reflected in relative factor prices in the society. Thus, a labour abundant society uses relatively labour-intensive technology and a capital abundant economy uses capital intensive technology. Not only that, the ones with more labour will have a relative advantage in expansions of labour intensive industry. Similarly the capital rich country finds it easier to expand its capital intensive industry. This way, factor abundence as reflected in relative factor prices, determines the direction of specialisation.

These things, we must admit, were not considered by Ricardo. Does it mean that Ricardo's comparative cost theory is wrong? No. It only means that Ricardo's theory was not comprehensive and also sophisticated. But 'why trade takes place' can only be explained by relative abundance or scarcity, resulting in comparative costs differences. Ricardo's measuring unit of cost was not correct. In fact, at the time, when Ricardo propounded his theory, labour cost approach was generally rather universally accepted. In fact it was the only measuring unit. Ricardo was justified in using the existing tool. Tools of analysis have changed and therefore refined version or a super structure can be created on the foundations laid down by Ricardo using new tools of analysis.

Check Your Progress 1

1) Is internal trade basically different from international trade? (Explain in one sentence)

.....

.....

.....

2) What is absolute difference in cost ? (Give example and explain)

.....

.....

3) Will there be trade with equal differences in cost ?

.....

.....

4) Explain the advantages of comparative cost principle.

.....

.....

5) Discuss one important shortcoming of comparative cost principle.

.....

.....

6) Analyse the effects of differences in endowments and specialization.

.....

.....

20.3 TERMS OF TRADE

We have so far answered only one question, why trade takes place. The other aspect of the problem of trade is, how trade takes place, which is unanswered. It will depend upon how much of one country's commodity the other participating country is willing to accept against a certain amount of his own commodity. The fact is that we must know the terms of both the participating countries at which they are willing to exchange their commodities.

The determination of terms of trade is also an important problem. The comparative cost differences provide the limits within which terms of trade can be determined. The actual point is the outcome of the effectiveness of the demand of both the countries. J.S. Mill's theory of reciprocal demand explains the determination of terms of trade between two countries.

20.3.1 Theory of Reciprocal Demand and Terms of Trade

J.S.Mill's theory of reciprocal demand explains how trade takes place. The problem is very simple. A comparative difference in cost-ratios sets the limits within which participating countries can import and export goods and commodities. How the terms of trade are determined between two countries? One may like to know how much, say, one country will export and how much it will import. The Ricardian theory does not answer this.

Determination of terms of trade was discussed by J.S. Mill. Because unless we determine terms we cannot explain how many units of goods and commodities will be imported/exported. This important aspect of trade was discusses by J.S. Mill on the foundations laid down by David Ricardo. Given the difference in cost, the demand in both the countries will determine terms of trade i.e., how much of a commodity one country is willing to purchase in exchange for a certain amount of other country's commodity. It is, therefore, the demand for each other's goods or the reciprocal demand, which will determine the terms of trade between the participating countries. J.S. Mill, on the basis of the cost ratios (that is, accepting

the differences in comparative cost as the cause for trade as suggested by Ricado) explained the determination of terms of trade.

Let us again take two countries, A and B, producing two commodities X and Y.

Country	Commodity	
	X	Y
A	120	100
B	90	80

Again to repeat, with the given amount of factors, a country can produce, 120 units of X or 100 units of Y. Similarly country B can produce 90 units of X or 80 units of Y. It may be pointed out that although country B in isolation is better equipped in the production of X as compared to Y but as compared to country A, it is inferior in the production of both commodities. According to the comparative cost theory, country A will specialize in the production of X and country B in the production of Y commodity.

Let us consider the cost ratios of both the countries as given above. In country A the ratio of X and Y is 12/10. It means 12 units of X will exchange 10 units of Y, or vice versa. In country B it is 9:8, that is, 9 units of X can be exchanged for 8 units of Y, or vice-versa. As indicated above, country A will be happy to get 10 units of Y by exporting anything less than 12 units of X. Similarly country B will be happy to get more than 9 units of X by exporting 8 units of Y. In between these extremes the terms of trade will be determined as there is gain to both the countries.

A is willing to give 120 units of X for 100 units of Y. B is willing to accept 90 units of X for 80 units of Y or 112.5 units of X for 100 units of Y. Thus, if they agree to a rate between 112.5 to 120 units of X for 100 units of Y, both will be better off.

If the rate of exchange so determined is equal to A country's internal ratio, it will not be willing to trade. So there will be no trade. If the rate is closer to rate of transformation in country A, most of benefits from trade accrues to B.

The concept of reciprocal demand is also criticized by a number of economists. Reciprocal demand considers only the demand of both the countries. But demand alone cannot help us in determining the terms of trade. The supply side, that is, the cost aspect is also important. In fact, there are two sets of demand and two sets of supply. Marshalls analysis overcomes this limitation. Marshall uses the concept of offer curves for analyzing the reciprocal demand. The offer curves of a country tells us, how much of a commodity one country is willing to exchange for another commodity. The point at which offer curves of respective countries cut each other will be the equilibrium point and terms of trade will be determined at that point.

20.3.2 Types of Terms of Trade

Before we discuss gains from trade, it will be better to have some idea about different types of terms of trade. In general, the terms of trade refers to a rate at which goods or commodities are exchanged between two countries. If for example, 5 units of X of country A can be exchanged by 10 units of Y of country B then the terms of trade will be 5X: 10Y or X:2Y. It means that terms of trade can also determine gains from trade. For example, if the price of export commodities

increases and the price of import commodities decreases or even remains constant, then this country will relatively gain more form the other participating country. It can buy more goods by exporting less goods. The opposite will happen if import price rises and export price either remains constant or declines. In view of this, it is desirable to discuss types of terms of trade.

1) **Commodity Terms of Trade**

It is also called net barter terms of trade. As the term indicates, terms of trade are determined on the basis of exchange of commodities. When commodities exchange between two countries, it is actually barter. It is a ratio between the export and import prices of a country. In simple terms commodity of trade at a point of time is equal to

$$T_c = \frac{P_x}{P_m}$$

(where Tc is commodity terms of trade, P is price and X and M refer to exports and imports, respectively). When we consider a period of time, then the changes in commodity terms of trade can be explained. We can write O for the base year and I for next period. If we want to find out the change in period I as compared to the base year, then it can be put as follows:

$$TC_1 = (PX_1 \quad PX_O)/PM_O.$$

2) **Gross Barter Terms of Trade**

Gross barter terms of trade as the term indicates is a modifies version of commodity terms of trade. It takes into account total exports and imports of a country instead of a particular commodity export and import. The gross barter terns of trade at a point of time will indicate the relation of ratio between total quantity of imports and total quantity of exports. In the form of a formula we can put it as follows :

$$T_g = Q_M/Q_X$$

Where, Q_m = total quantity of imports, Q_x = total quantity of exports.

If this ratios is higher then it is favorable to the country concerned and if the ratio is lower then it is unfavorable to the country. Like commodity terms of trade, we can consider gross barter terms of trade over a period of time. We can compare between two points of time. As above, the base period will be indicated by O and the next following period by 1. The formula can be modified a little :

Base Year : $TG_0 = Q_{MO}/Q_{MO}$

Current Year : $TG_1 = QM_1/QX_1$

3) **Income Terms of Trade**

G.S. Dorrance has formulated the concept of income terms of trade. It is an improvement over commodity or gross barter terms of trade. In this concept of terms of trade, we take into account index of export prices along with export quantity and in case of import we consider only index of import prices. As the import and export prices are determined by total imports and exports, it can rightly be considered as a modified or improved version of gross barter terms of trade.

Income terms of trade establish a ratio between index of export prices and export quantity on the one hand and index of import prices on the other hand. To put it into a formula

$$T_y = \left(\frac{P_x}{P_m}\right) Q_m$$

Here T_y is the income terms of trade (it may be pointed out that in economic analysis normally Y and not I is written for income. Therefore Y has been written). T_y measures country's export based capacity to import.

A change in income terms of trade indicates the nature of changes in trade (import and export). If there is an increase in terms of trade, then it shows that a country can further increase its imports. If there is decrease then the reverse will be true, that is, the capacity to import will reduced. Incidentally, it can be added that if income terms of trade are favorable, it does not necessarily follow that barter terms of trade are also favorable. This is because of the concept of prices are more important and price-variation can make a change without causing a change in barter terms of trade.

4) **Single Factoral Terms of Trade**

Trade is a continuos process spreading over a period of time. It is possible that during the time period some changes can take place in the utilization of factors, which will result in an increase in productivity. Single factoral terms of trade takes into account of these changes in productivity in export sector. The index of the productivity in export industries has to be multiplied by the commodity term of trade giving,

$$T_y = \left(\frac{P_x}{P_m}\right) F_m ,$$

Where T_s = Single Factoral Terms of Trade, P_x = Price index of exports, F_x = productivity.

Any improvement in productivity will be reflected in an improvement in terms of trade. At the same time there is disadvantages also. International trade implies open economy, that is, participating countries are maintaining economic relations with other trading countries. If we want to have the real and correct picture, then we must consider the changes in the productivity of the other country's export side also. When we consider changes in the productivity in both import and sectors and make necessary adjustment in the formula, then it is called **Double Facotoral Terms of Trade.** In this we take account of changes in productivity of both importing and exporting countries. In view of this, we can modify the equation of single factoral terms of trade to take account of changes in the other country. It will be

$$T_d = \left(\frac{P_x}{P_m}\right)\left(\frac{F_x}{F_m}\right) = \frac{P_x F_x}{P_m F_m} , \quad \text{where}$$

P_xF_x = price index of exports multiplied by productivity index of exports, P_mF_m = price index of imports multiplied by the productivity index of imports.

In both cases it will indicate relative changes in productivity.

These are some of the main types of terms of trade. Besides Prof. Viner and Professor Kindleberger have also talked about 'Real Cost Terms of Trade' and 'Utility Terms of Trade.'

Terms of trade occupy an important place in international trade. It indicates the purchasing capacity of a country to purchase goods from other country. Every country wants to have favorable terms of trade because it shows its capacity to make purchases in the international market. On the other hand, unfavorable terms of trade indicates the limitations in making purchases in the international market. In fact, terms of trade can be considered as means of gain from trade.

Check Your Progress 2

1) What are terms of trade?

.....

.....

2) What is the meaning of reciprocal demand?

.....

.....

3) Will a country gain much if her cost ratios and terms of trade are identical?

.....

.....

4) Discuss the main limitation of J.S. Mill's theory of reciprocal demand.

.....

.....

5) Explain the concept of offer curves.

.....

.....

6) Explain the meaning of commodity terms of trade.

.....

.....

7) Fill in the Blanks:

a) T_C = is ……. terms of trade

b) T_y = is ……. terms of trade

c) T_g = is ……. terms of trade.

20.4 GAINS FROM TRADE

The discussion of why and how trade takes place clearly indicate the gains from trade. International trade is international division of labour. Just as individuals gain by division of labour, similarly countries also gain by international division of labour. Trade is an extension of division of labour between countries. Some of the important gains are as follows.

i) **Optimum use of Resources**

With division of labour there will be optimum allocation of resources and maximum production within a country and also between countries. Like individuals, when all the countries specialize in those lines of economic activities where they have comparative advantage, then naturally best use of economic resources will be made. And when best use of economic resources is there, production will be maximum.

Further, resources relatively speaking, are not abundant; continuous, prolonged and indiscriminate use will exhaust them soon. Besides, the maximum outcome will not be there. Therefore, a proper and careful use will be essential. International Trade makes it possible because of the comparative advantage in costs. A country which can produce a commodity at a lower cost than other countries will make a better use of existing resources. In an isolated country when resources are used to produce most of the required commodities, then naturally the optimum allocation of resources will not be feasible. Thus, we can say that if the countries trade then not only within a country but even at the international level, there will be desirable use of resources.

ii) **Advantage of Large Scale Production**

Division of labour is limited by the size of the market. If a country has limited demand then production will be less. International trade removes this limitation of the market. Now, a country will produce not only for his consumers but for the consumers of different countries. The size of the market has increased. As a result economies of large scale production will be operating. These economics can be listed below:

a) **Economy in Large Scale Buying and Selling**

A country now producing for a bigger market, will purchase inputs in large quanties. It is a common experience that when we make bulk purchases, there is economy in expenses. Similarly, the cost of selling per unit will decrease. Thus, there will be economy in large scale buying and selling.

b) **Gains accruing due to Indivisibility of a Factor of Production**

Every machine has an optimum capacity. If the market is limited then production from that machine will be less and the unit cost will be higher. The cost of commodity can be divided into fixed and variable cost. Total variable cost will increase with an increase in output. Fixed cost per unit will decline till the optimum point is reached. In view of this as the production will increase, the unit cost will decrease till the optimum point is reached. Thus, we can say that due to trade, production will increase and machines or productive units will be producing to the optimum level and as a result the cost will decrease.

c) **Improvement in the Quality**

Because of large scale production and competition in the market, the quality of commodities will increase. In fact, consumers will buy goods of better quality with a lower price. Therefore in order to secure the market, entrepreneurs (producers) will like to improve the quality of commodities. Continuous research and development will become a part of the business unit. The cold wind of competition in the international market will force producers to renovate or innovate. There may be limitation of the best technique available in the world. Besides, the competition will force producers not only to improve the quality but related development in the sources of raw material, banking facilities etc., Thus, there will be many gains from trade.

The gains from trade are not limited to optimum allocation of resources and advantages of large scale production only. Trade directly influences developmental effort of a country. It assumes the role of a leading sector in the process of economic development. It has also been called the engine of economic growth. The fact is that rapid development of an important sector can promote development of all other sectors in the economy.

20.4.1 Factors Determining Gains from Trade

The extent of gain from trade is determined by many factors. We can discuss them under the following heads:

Relative Differences in Cost Ratio

The extent of gain from trade is determined by the relative differences in cost ratios. If a country has greater differences in cost ratios it will gain more because if the differences are marginal then gains will also be marginal. Thus, gains are directly related to productivity and efficiency conditions prevailing in a country. Higher the productivity and efficiency greater will be the gains from trade.

Reciprocal Demand

Reciprocal demand also determines the extent of gain. If, for example, country A's demand is more and country B is not willing to supply at the existing rate, then rate will change in favour of B. Or, if country A's demand is less and country B is willing to supply more then the terms of trade will favour country A. The relative strength and elasticity of demand of both the countries will determine the gains from trade. In other words, the extent of gain from trade is determined by reciprocal demand.

Ultimately two factors figure out importantly in determining the gains from trade: Higher the efficiency in production, greater will be gains. Further, income and nature of the commodity, which will influence the demand, will influence the gain. As more than one country is involved in trade, we have to consider the relative capability and demand of both the countries.

Sometimes, size of a country also influences the gains. It can be said that the gains to a small country will be relatively larger, because, a small country faces many obstacles and limitations in large scale production. In this way size of a country also influences the gain. On the other hand, in a very small country, availability of domestic reasons will be limited in size and variety. This may have adverse effects on efficiency. Therefore, we cannot make any generalization and relate the gains to the size of an economy.

Check Your Progress 3

1) Will there be gain from International Trade if the cost ratios are the same in both the countries? (Answer in two sentences)

   .....

   .....

2) Name two important gains from trade.

   .....

   .....

3) How cost ratios affect gains from trade.

.....

.....

4) Is demand also important in determining gains from trade?

.....

.....

5) Will the size of the market influences the gains from trade?

.....

.....

20.5 BALANCE OF PAYMENTS

In this section, we will introduce you to the monetary aspect of international trade. When countries trade with each other, they pay in foreign exchange or foreign currencies. So, we will be talking here of trade among countries which is not barter (that is, not in terms of physical commodities). Each country has its own currency, it own monetary system. This complicates matters. We are paying attention to some of them here.

The fundamental tool for depicting international economic transaction is the balance of payments accounts. The balance of payments accounts are a summary accounting statement of the dealings of a country with the outside world. The balance of payments accounts constitute a record of the transactions between a country and the rest of the word in a specific time interval, usually a year. Hence the balance of payments are a flow, not a stock. This is important to remember. The balance sheet of a firm, or even the national balance sheet are stocks.

There are two fundamental things to keep in mind while talking of the balance of payments. The first is that a distinction be made between debits and credits. Secondly, debit and credit items are put in proper sub accounts.

Let us begin with the second point. The balance of payments account (presented as a table) is usually divided into two main parts and each part has several subdivisions. The two main parts are:

1) **The current account**: This account shows all flows that directly affect the national income accounts of the country. It includes:

   i) exports and imports of merchandise

   ii) exports and imports of services

   iii) inflows and outflows of income

   iv) income from investments

   v) grants, remittances and other transfers.

2) **The capital Account:** This account shows all flows that directly affect the national balance sheet. It includes:

   i) direct investment

   ii) portfolio investment

   iii) other private capital

   iv) changes in cash balances, which include:

a) changes in official reserves assets, that is, changes in the reserves held by official monetary authorities, resulting from intervention in foreign exchange markets.
b) changes in cash balances held by banks and other foreign exchange dealers.

Item (a) above is actually a balancing item. Particularly it is treated separately from government capital. The main reason is that official reserve transactions are undertaken by the authorities in order to manage the country's balance of payments (BoP).

Let us now come to the other main point about the balance of payments. This principle is the distinction between debits and credits. Here, remember two things: First, since the balance of payments table is an accounting statement, the items balance in the aggregate. The total debits always equal total credits (any difference is shown as statistical discrepancy). Every individual transaction in the balance of payments table appears twice because the BoP accounts are built on the principles of double entry book-keeping. Each individual transaction appears for the first time in the current or capital account and then for the second time in the cash component of the capital account.

The second thing to remember is, that is, a general sense, any action that creates an obligation for a foreign economic agent to pay a domestic economic resident is a credit, whereas any action that creates an obligation for a domestic resident to pay a foreign agent is a debit.

Let us have a closer look into structures of balance of payments account of countries.

20.5.1 Structure of Balance of Payments

The terms balance of trade and balance of payments are often used in international trade. These two terms have different meanings. Balance of trade as compared to balance of payments has a limited and narrower meaning. Balance of trade refers only to goods, commodities and services' trade. It is also called visible and invisible import and exports. Goods and commodities can be seen and therefore they are called visible items. Services like shipping, freight and insurance are called invisible items. Balance of trade has, thus, restricted meaning. Balance of payments on the other hand is more comprehensive. It is an account of total transactions of a country with the outside world for the period of time under consideration. When we call it an account of total transactions it takes account of not only visible and invisible items of trade but other transactions as well like capital transactions.

The principle of double entry book-keeping is followed in making an account of balance of payments. But there is one difference. In business debits are shown on the left side and credit on the right side. In balance of payments, credit is shown on the left side and debit on the right side. But a question naturally arises. What do we mean by credit and debit in international trade? All transactions of a country involving receipts are credit. It includes exports of goods, commodities and services and also the capital and other things which a country is receiving. All these things are kept on the credit side as they are receipts. On the other hand, when we purchase some items, we have to make payments for them. It is the debit side. All imports of goods commodities and services require payments to the other country. These import items are kept on the debit side. Besides these, a country can give capital assistance to another country or payments of loans and interest to international institutions like World Bank and International Monetary Fund or from some other countries. They are all payments and therefore included on the debit side.

Thus, we can group all transactions of a country either as receipts or payments. Receipts are kept on the left side and payments on the right side.

Now the next question that follows is about the items of balance of payments and their classification. Important items of balance of payments are imports and exports of goods and commodities. Normally a country imports a number of commodities which cannot be produced cheaply or otherwise. Similarly, a country exports those goods and commodities which can easily be produced efficiently and economically. These commodities can be perishable goods, durable goods and also machines etc. A country will import goods only according to its requirements and needs. As suggested earlier all these items are included in visible items. These items can be either exports and imports.

The other items are services like shipping and insurance etc. These are different than visible items because services cannot be seen. For example when we export some goods through either air or water, services involved in such processes cannot be noticed. Therefore, invisible items include exports or imports of services. Normally a less developed country has to pay more for invisible items as services are not well developed in these countries.

These two items, visible and invisible, are included in **trade items**. Besides trade items, balance of payments also includes items which are called **transfer items.** These transfer items deal with financial transactions i.e., the transactions of money. The financial transactions can also be divided into two parts **unrequited transfers** and **capital transfers.** Unrequited transfers are basically different from capital transfers. The difference is of nature and type. To-day many people of different countries are working in other countries. Why they are working is not our concern in the context of balance of payments. These people remit a part of their earnings to their home country. Our concern is with their transaction. Sometimes people send gifts etc., to persons in other countries. It can be a gift check. Besides these, we have indemnities which countries may individually pay to other countries. For all these items in the case of which nothing is given in exchange, nor is there any commitment to return such items at a future date, are included in category of unrequited transactions. In fact these are the items where we do not bother about the cause but we consider only its effects. These transactions take place and are in the form of payments or receipts. Our only concern is to keep these under unrequited receipts or payments. It can be added here that generally a large number of people from less developed or developing countries go to other countries where they can earn more. Their remittances are normal as they send the money to their family members. We have to take account of these transactions also as they are included in balance of payments. Another important component of transfer item is capital transaction. These transactions include borrowings from foreign countries and international institutions, payments of loans and interest etc. In fact, international borrowing has come to stay. Countries take loans for development and other purchases. These can also be divided into receipts and payments. The amount which a country is receiving from other will be included in capital receipts and similarly capital payments will be those which a country is paying to other countries These are the main items of balance of payments. We can make a hypothetical account of developing countries' balance of payments.

This balance of payments account has been divided into parts A and B for the sake of convenience. Part A included trade items only which has visible and invisible exports and imports of goods, commodities and services. Part B deals with transfer items. In transfer items money transactions take place between countries. These money transactions have also have been divided into sub-groups -unrequited

transactions and capital transfers. These are the four main groups under which we can include all transactions of country. This account is presented in Table 20.1.

If we carefully go throughout the table, we can make some important observations about the balance of payments.

Table 20.1 : Balance of Payments

(In. Rs. crore)

Credit Side		Debit Side		Balance Surplus or deficit
Part A				
TRADE TRANSACTIONS				
(i) Visible Exports	750	Visible Imports	950	–200
(ii) Invisible Exports	200	Invisible Imports	150	+50
Total of Trade items	950		1100	–150
Part B				
TRANSFER TRANSACTIONS				
(iii) Unrequited	150	Unrequited	100	+50
(iv) Capital Receipts	200	Capital Payment	100	+100
Total of Transfer Receipts	350		200	+150

First of all, it is important to note that the total receipts are equal to total payments. On the credit side we have Rs.1300 crore and on the debit side, we have Rs.1300 crore. There is balance between credit and debit sides. Total payments of this country are equal to total receipts. It means there is equilibrium in balance of payments. We can say that balance of payment always balances (in the account sense).

Secondly, if we consider different groups or heads, there is difference between receipts and payments. Let us first consider trade items. Under trade items,visible exports are of Rs.750. While imports are of Rs. 950 crore. The value of imports is Rs.200 crore more than the value of exports. In visible items, there is deficit as imports exceed exports by Rs.200 crore.

It means that the country under consideration is importing more than what it is exporting. This can also mean that the country is unable to meet its requirements or there are certain goods/commodities which cannot be produced are being imported. As exports are less than imports there is deficit. There can be many factors for lower level of exports. For example, the goods which are produced by this country are less in demand or its quality is inferior and price higher. We can say that it has limited export capacity and therefore, there is a deficit in visible items.

Invisible items give a different picture. Invisible exports are greater than invisible imports. The table shows that exports are of Rs.200 crore while imports are of Rs.150 crore under this heading. There is export surplus. Exports are greater by Rs. 50 crore as compared to imports.

We can say, now, that in trade items visible imports are higher by Rs. 200 crore and invisible exports are greater by Rs.50 crore only. If we consider the total receipts and payments under trade items only, we find that total receipts are equal to Rs.950 crore and total payments are equal to Rs. 1100 crore. The net gap between payments and receipts is of Rs.150 croroes. Total payments are greater than total receipts. There is overall deficit in trade items. It can also be said that there is disequilibrium in balance of trade and the deficit is of Rs.150 crore.

Let us now consider transfer items. Transfer items are further sub-divided as unrequited items and capital items. We will consider receipts and payments under these two groups. It is clear from the balance of payments table that receipts under unrequited items are greater than payments. Receipts are of Rs. 150 crore and payments (debit side) are of Rs. 100 crore only. In this group receipts are greater than payments. We can say there is surplus in unrequited items. This surplus is of Rs.50 crore. The capital items also show that country is getting more than what it is giving. May be as a loan/grant from international institutions like IMF, World Bank or from some developed countries. This country is getting more assistance and its payments are limited to repayment of loans and interest thereon. Capital transactions like the unrequited transactions show greater receipts than payments. Under capital transactions, receipts amount to Rs.200 crore while payments are only of Rs. 100 crore. It means receipts are greater by Rs.100 crore than payments.

Let us now consider total surplus and total deficit in balance of payment. It will be easier to explain if we rearrange of balance of payment according to deficit/ surplus.

It can be observed that in trade items, total receipts amount of Rs.950 crore while payments are of Rs.1100 crore, thus, leaving a gap of Rs.150 crore. We can say that there is deficit in balance of trade items, which is equal to Rs.150 crore. But this does not mean that there is deficit in balance of payments of this amount. We have to take account of transfer items also. Both the groups under transfer transaction, namely, unrequited transactions and capital transactions show greater receipts than payments. The table shows that unrequited receipts are greater than unrequited payments. There is surplus of Rs.50 crore. Similarly capital receipts are higher by Rs.100 crore than capital payments. Thus, total transfer items indicate a surplus of Rs.150 crore. Thus, we find that the deficit of trade items is wiped out by the surplus of the transfer items. The total receipts are equal to that payment. It is clear from the balance of payments schedule that the total receipts of this country are equal to Rs.1300 crore and total payments are also Rs.1300 crore. There is neither deficit nor surplus. We can say, therefore, in an accounting sense, balance of payments always balances.

Now, there is another important question. Very often we say that a country's (particularly less developed or developing country) balance of payment is unfavorable. It means there is deficit in total transactions. But in our example, we find that there is equilibrium in balance of payments i.e., there is neither surplus nor deficit. Apparently these are contradictory statements. At a time only one statement can be true. If there is deficit in balance of payments, it cannot be in equilibrium. And if it is in equilibrium there cannot be deficit (or even surplus) in a balance of payments.

We can easily resolve the contradiction. First of all, when we say that balance of payments is in equilibrium balance of payments always balances. It is in the accounting sense. Because, whenever our total payments are greater and total receipts are

lesser, we have to make payment somehow. We can borrow from other countries/ international institutions to make the payment. When we borrow, it is kept on the credit side because it shows receipts and therefore total receipts are equal to total payments. It is clear that with the classification and items of a balance of payments given above we cannot give a meaningful idea of surplus or deficit in a balance of payment. It will not indicate the pressure on balance of payment. It will not indicate the pressure on balance of a payment unless we specifically know the amount borrowed for making payment or a similar action adopted by this country.

Professor J.E Meade, in trying to explain this, has rearranged the balance of payments in terms of autonomous transactions and accommodating transactions. All the terms of a balance of payment may be autonomous or accommodating or partly autonomous and partly accommodating. And on this basis we can explain deficit or surplus on a country's balance of payments. What are autonomous and accommodating transactions? Autonomous transactions arise out of normal trade consideration i.e., a country is in a position to export in the international market. All these transactions are autonomous. Or for example, if we import some commodities from a country and that country is the cheapest and best source. It means we are purchasing purely on market/trade considerations. This is autonomous transaction. There can be some other consideration also. For example, suppose we want to help a friendly country. We can purchase some items from that country or give financial assistance. These transactions are different from autonomous transactions. They are called accommodating transactions. Here the main objective is to accommodate a country. Reasons may be political, economic, social or even religious. The demarcating line is that these transactions involve a deliberate intention to 'accommodate', 'assist', 'help' the other friendly country. Thus, all transactions can be either autonomous or accommodating. Accommodating transactions of a country will indicate the extent of imbalance in a country's balance of payments. So, Professor Mead has defined deficit in a country's balance of payments as follows: 'An actual balance of payments deficit as the actual amount of accommodating finance used in any period of time', and a potential balance of payments deficit as the amount of accommodating finance which it would have been necessary to provide in any period in order to avoid any depreciation in the exchange rate without the employment of exchange control, etc.

It is of course this potential deficit (or the corresponding potential surplus) that is the proper measurement of balance of payments disequilibrium.

Check Your Progress 4

1) What is balance of payments? (Write in one sentence)

.....

.....

2) What is the difference between balance of trade and balance of payments.

.....

.....

3) In which sense, if any, balance of payments always balances.

.....

.....

4) What are invisible transactions?

.....

.....

5) Name any two important items of balance of payments.

.....

.....

6) Explain the meaning of deficit in balance of payments.

.....

.....

7) Is the deficit in balance of trade the same as deficit in balance of payments.

.....

.....

20.6 LET US SUM UP

This unit dealt with various aspects of international economics. It started by discussing the reasons for countries to trade with each other. Among other theories, the unit discussed extensively the theory of absolute advantage propounded by Adam Smith, and the theory of comparative advantage most commonly associated with David Ricardo. It also discussed the modern variant of comparative advantage put forward by Heckscher and Ohlin. The unit then went on to discuss the potential gains from international trade. Some of these are: optimum use of resources and advantages of scale economies. Concepts pertaining to various types of terms of trade such as the commodity terms of trade, gross barter terms of trade, income terms of trade, single factoral terms of trade were then explained and discussed,. This was also related to the theory of reciprocal demand. Finally the unit looked at monetary aspects of international trade and discussed the meaning of balance of trade and balance of payments. It also described the components of and the items that constitute a country's balance of trade and balance of payments

20.7 KEY WORDS

Balance of Payments : An account of total transactions of a country in a given period of time with the outside world.

Balance of Trade : An account of total merchandise trade including visible and invisible items of a country.

Capital Transaction : This is transfer of capital from one country to another. It may be from a country , international institutions like IMF or IBRD.

Comparative Cost Advantage : Difference in the cost ratio of different countries in different commodities gives comparative advantage in costs. One country may enjoys advantage in relative cost ratios. As a result one country can be more efficient in some commodities and another country can be less inefficient in some other commodities.

Division of Labour : One man cannot be equally competent in all the lines of economic activities. One person may be more efficient in some lines of economic activities, another person may be more efficient in other lines of activities. In view of this every person chooses that line of economic activity in which he is relatively more efficient.

Equilibrium : A point where two opposite forces meet. Where receipts are equal to payment.

Invisible Items of Trade : It means items of services etc., which are traded between countries and includes shipping, insurance etc.

Leading Sector : In the process of growth, during its third stage which Professor Rostow terms as 'the take off stage', one sector develops rapidly. It has linkage effect that favorably affects other activities also. It leads to the development of other sectors. Hence the name, 'Leading Sector'.

Offer Curve : It means the amount of a country's commodity which that country is willing to offer for a certain amount of other commodity of the other country.

Reciprocal Demand : The demand for each others' goods in exchange for ones' own output by two countries.

Terms of Trade : The terms (or the rate) at which two commodities are exchange between countries.

Unrequited Transactions : There is no quid pro quo, immediate or otherwise for these item.

Visible Items of Trade : It means goods and commodities which are traded between countries. As goods and commodities can be seen they are called visible items of trade.

20.8 SOME USEFUL BOOKS

Sodersten, B. (1980), *International Economics*, Mcmillan, London

Kenen, Peter (1994), *The International Economy*, Cambridge University Press, Cambridge, UK.

Krugman, Paul R. and Maurice Obstfeld (1997), *International Economics: Theory and Policy (Fourth Edition)*, Addison-Wesley, Massachusetts, USA

20.9 HINTS/ANSWERS FOR THE CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

For all, Read Section 20.2

Check Your Progress 2

1) Read Section 20.3 and answer.

2) Read Sub-section 20.3.1 and answer.

3) Read Sub-section 20.4.2 and answer.

4) Read Sub-section 20.3.2 and answer.

5) Read Sub-section 20.3.2 and answer.

6) Read Sub-section 20.3.3 and answer.

7) Read Sub-section 20.3.3 and answer.

Check Your Progress 3

1) Read Sub-section 20.4.2 and answer.

2) Read Sub-section 20.4.1 and answer.

3) Read Sub-section 20.4.2 and answer.

4) Read Sub-section 20.4.2 and answer.

5) Read Sub-section 20.4.1 and answer.

Check Your Progress 4

1) Read Section 20.5 and answer.

2) Read Section 20.5 & Sub-section 20.5.1 and answer.

3) Read Sub-section 20.5.1 and answer.

4) Read Sub-section 20.5.1 and answer.

5) Read Sub-section 20.5.1 and answer.

6) Read Sub-section 20.5.1 and answer.

7) Read Sub-section 20.5.1 and answer.

to inflationary pressures, specially if the deficit is small in relation to national income and is not persistent. Also, deficits can sometimes help the economy to recover from a recession.

21.8 MARKET BORROWING

In the previous section, you got some idea of the various measures of deficit. We saw that a deficit, howsoever defined, is an excess of spending over government revenue. In this section, we look at a related concept: market borrowing.

Public Debt

The overall debt and obligation of the government, measured at a point of time, is the public debt. The public debt has been defined in various ways depending on the items that are thought appropriate to be included in the definition. To get an idea of the public debt, let us look at the various obligations of the government. First, the government creates currency. Often, a part of the currency may be issued by the central bank, but usually the central bank in most countries is part of the government so that the total currency issued and obligation may be considered a government liability to the rest of the economy.

The second set of obligations is the short-term debt, normally with a maturity of less than a year at the time of issue and consists of items such as Treasury Bills and short-term loans from the central bank. There are some debts that do not have any specific date of maturity and are called floating, and part of these may be paid of at various times and are subject to various terms and conditions. These include provident funds, small savings, reserve funds etc. In India, the government has issued certain special securities to meet its obligations towards international institutions like the World Bank and the International Monetary Fund (IMF). These special securities are sometimes called special floating debt.

The importance of market borrowings lies in the fact that in some cases, such as Indian public finance, market borrowings are excluded in the estimation of budgetary deficits. Market borrowings are long term borrowings, where the maturity period is over a year. The reason given for excluding market borrowings from budgetary deficits is that it is felt that since these are long term obligations, they merely divert investible funds from the private sector to the government and hence do not raise the purchasing power and the quantum of currency. Consequently, inflationary pressures in the economy do not build up by the market borrowings. This view need not be correct, as the RBI itself takes up a large portion of market borrowings. The effect of both short term and long term loans taken up by the RBI is the same, in increasing the amount of currency.

Check Your Progress 4

1) Distinguish between public deficit and public debt.

.....

.....

.....

2) Explain the effect of deficit financing on the capital market

.....

.....

.....

21.9 LET US SUM UP

This unit has acquainted you with the basic elements of public economics including several basic concepts from traditional public finance. Our point of departure was a description of market failure. We discussed the various types of market failure. Next we explained the concept of Pareto improvement and Pareto optimality. In this connection we mentioned the link between perfect competition and Pareto optimality. Here we also talked about the two fundamental theorems of welfare economics.

We then took up the explanation of the concept of a public good. We saw that non-rivalry and exclusivity characterize a public good. A good like this is actually a pure public good. We took up a brief discussion of impure public goods, two types of which are club goods and goods in the presence of congestion. We also considered some issues in the provision of public goods and saw the central difference in obtaining a social benefit function from individual functions. It was pointed out that the individual curves are added vertically instead of horizontally, as happens in the case of a private good. We saw, too, what Lindahl equilibrium means.

We next went on to a discussion of externalities. We saw that an externality is an example of a market failure. After considering the four basic types of externalities, we discussed ways of coping with externalities. Here we discussed Pigou's suggestion of correction using suitable taxes and subsidies. The other way, as we saw, is granting property rights and letting private parties arrive at mutually beneficial deals. We explained very important Coase as well.

These broad topics of market failure, public goods and externalities constitute elements of what may be called an extension of some the previous units that you studied in this course. The next three broad topics- public revenue, deficits and market borrowing are central themes in traditional public finance. We began our discussion of public revenue by pointing out a distinction that some have made, between revenue and receipts after broadly classifying revenue into tax and non-tax. We then turned to a discussion of the engaging an important topic of deficits and their financing, briefly touching upon the distinction between debt and deficit. We discussed and compared various measures of the government deficit.

21.10 KEY WORDS

Balanced Budget : A government budget in which the total revenues exactly match the total expenditure.

Capital Expenditure : Expenditure incurred on items that have a long term basis.

Capital Receipts : Receipts accruing from items of a long term character.

GNP Elasticity of Public Expenditure : The ratio of the percentage rise in government expenditure to a percentage change in GNP.

Perfectly Competitive Market : A Market situation where no buyer or seller is able to influence prices, where the goods sold by all sellers are homogeneous, and where there is full and perfect information and free entry and exit to and from the market.

Pure Public Good : A good characterised by both non-rivalry and non-exclusion in consumption.

Revenue Expenditure : Expenditure incurred on short-term activities

Revenue Receipts : Receipts incurred on activities, usually for a period more than a year.

Short-term Debt Instrument : Those instruments with maturity of less than one year.

21.11 SOME USEFUL BOOKS

Bhatia, H.L, *Public Finance* (Latest edition) Vikas: New Delhi

Browning, Edgar K. and Jacquelene Browning (1994), *Public Finance and the Price System* (Fourth Edition) Prentice Hall: Englewood Cliffs New Jersey

Musgrave, Richard A. and Peggy B. Musgrave (1989), *Public Finance in Theory And Practice* (Fifth Edition) McGraw-Hill International Edition: New York

Stiglitz, Joseph E. (1994), *Public Sector Economics*, Third Edition, W.W. Norton & Co.: New York

21.12 HINTS/ANSWERS TO CHECK YOUR PROGRESS EXCERSISES

Check Your Progress 1

1) Read Section 21.2 and answer.

2) Read Section 21.2 and answer.

3) Read Section 21.2 and answer.

Check Your Progress 2

1) Read Section 21.3 and answer.

2) Read Section 21.3 and answer.

3) Read Section 21.3and answer.

4) Read Section 21.4 and answer.

Check Your Progress 3

1) Read Section 21.5 and answer.

2) Read Section 21.5and answer.

3) Read Section 21.6 and answer.

Check Your Progress 4

1) Read Section 21.7 and answer.

2) Read Section 21.8 and answer.

UNIT 21 PUBLIC ECONOMICS

Structure

21.0 OBJECTIVES

In this unit of the course on the fundamentals of economics, we consider some situations where the market process may not allocate scarce resources in an optimal manner. You will also be introduced to some concepts and processes of the government's economic functioning. After going through the unit you will be able to:

- list the situations of market failure;
- state the two basic laws of welfare;
- identify and describe situations of externalities;
- define a public good, and discuss its properties;
- explain the concepts and components of government revenue and expenditure;
- describe the various measures of deficit and discuss their impact; and
- explain what is market borrowing and list the other methods of deficit financing.

21.1 INTRODUCTION

As we discuss through this unit, we shall be concerned with two broad themes. The first, which we discuss in the latter half of the unit, comprises concepts from what is called Public Finance. These concepts include government revenue, government expenditure, budget deficit and the ways of financing these. In other words, these are about how governments raise funds, how they spend these and on what, and how the deficit or shortfall is met. The earlier portions of the unit are an extension of standard microeconomics— the concepts that you studied in the first few blocks of this course.

21.2 MARKET FAILURE

For much of the initial section of this course, you were told about the functioning of the various market forms— perfect competition, monopoly and so on. In this section of this unit, we will look at some situation where the market is not necessarily the best allocator of resources. The idea is that under certain conditions (which we

shall come to in a moment, and which some consider being too stringent), the market mechanism happens to be the most efficient way to allocate resources, to carry out production, consumption and investment. When these conditions are fulfilled, the market should be left alone.

What are these conditions? The basic requirement is that perfect competition should prevail. In other words, no buyer or seller should be in a position to influence prices and there should be perfect and complete information, which is virtually costless to acquire. Another requirement is that there are no externalities. In a later section we will discuss externalities in detail but for now it is enough to know that it means the action of an economic agent confers unintended costs or benefits on some other economic agent(s). Thirdly, there should not be increasing returns to scale in production. Although an earlier unit on production functions has explained this concept, it is worth repeating here. Returns to scale deals with a situation where all inputs are increased by the same amount and we look at what happens to output. If output increases by more than the proportion in which the inputs are increased, we have increasing returns to scale. Suppose we double the amount of all inputs. If output more than doubles, we have increasing returns to scale. What happens if there are increasing returns to scale? The basic answer is that in such a case, even after all inputs have been paid their remuneration equal to their marginal product, the total product is not exhausted, and this means the presence of extra-normal or super-normal profits. This goes against the basic idea of perfect competition.

The important thing to understand here is that perfect competition satisfies an important criterion of welfare that is put forward in the literature. This is known as Pareto-optimality. Let us see what this means. Any change that makes someone better off while not making someone else worse off is known as Pareto superior or is called a Pareto improvement. When we get to a situation where it is impossible to make someone better off without making someone else worse off it is called a Pareto optimal condition. There are two statements made in welfare economics, which are known as the two theorems of welfare economics. The first theorem states that any perfectly competitive situation is a Pareto-optimal situation. The second theorem states that any situation which is Pareto-optimal can be seen to be equivalent to a perfectly competitive situation. There is, thus, a close relationship between Pareto-optimality and equilibrium in a (perfectly competitive) market. Pareto-optimality also automatically ensures that resource allocation is most efficient and wastes are avoided.

Check Your Progress 1

1) What are the conditions under which the market economy performs as the best allocator of resources?

.....

.....

.....

2) Explain the concept of Pareto optimality. When do we say that a situation is Pareto-superior as compared to another situation?

.....

.....

.....

3) State the two laws of welfare economics.

.....

.....

21.3 PUBLIC GOODS

When we hear the words 'public' and 'goods' together, we may be led to think that these goods are provided by the public sector. This is not necessarily the case. While it is true that the government largely provides public goods-and later we will discuss the reasons-this type of goods may also be provided by the private sector. There are two characteristics that a public good possesses, regardless of whether it is provided by the government, or not. These two characteristics are **non-rival consumption and non-exclusion**. Let us see what these concepts mean

Non-rival Consumption

This concept means that for a given level of production, consumption by one person does not diminish the quantity left for someone else to consume. In other words, a good is said to be characterised by non-rivalry in consumption if, once it is produced, several people can simultaneously consume it. We can give some examples. Suppose it is a chilly winter evening and you are keeping yourself warm by sitting by a fire. If I come along and sit beside you to get some warmth by the fire, the warmth that you get from the fire does not diminish. Another classic example is national defence. National defence 'consumed' by one citizen of country does not reduce the amount left over for others. Yes, it is true that people living in border areas may feel threatened more by an external attack; nevertheless, military defence *per se* as a good is characterised by non-rivalry in consumption. Other examples are pollution control measures, and many public health programmes.

In much of economics, we deal with goods, which are rival consumption goods. For a given level of production of shirts in an economy, the amount consumed by one person reduces the amount left over for others to consume. The basic implication for rival consumption is that a way has to be found to allocate or ration the good among the consumers. With non-rival goods, there is no such problem because consumption by one person does not reduce the consumption by another person. But this is not the case with rival goods and the usual mechanism applied or suggested by economic theory to allocate the goods is the price system or, in other words, through demand and supply.

Non-exclusion

This is the second characteristic of a public good. To understand non-exclusion we can look at private goods which are characterised by the exclusion principle, and see what this means. Then we can grasp the meaning of non-exclusion. Take a private good like an apple. If you eat an apple you can exclude me from eating the apple. So for a private good, it is possible to confine the consumption. On the other hand, for public goods, it may be impossible or extremely expensive to confine the benefits of the good to a few persons. A person will derive benefit from the production of the good, regardless of whether *or not he or she pays for the good*. A classic example of a public good, much cited in the literature, is a lighthouse. When the light of the lighthouse is on, it is difficult to prevent any nearby ship from seeing it and being guided by it (non-excludability). Moreover, one ship's use does not reduce the light left for other ships, or, in other words, one ship's use does not affect the ability of other ships to use the light (non-rivalry).

It is important to carefully distinguish between non-rivalry in consumption and non-exclusion. The former concept captures the idea that the amount consumed by one person does not reduce the amount left over for consumption by others. A good may be non-rival and yet it may be possible to *exclude* some people from using that good. A classic example is cable television broadcast. If a programme is being broadcast, then its viewing by one person does not diminish the amount left for others. However, it is possible to exclude some people from viewing the programmes.

Those who do not have access to a TV set, or those who do not subscribe to some channels may not be able to view certain programmes. Remember, there is no diminishing of the *quantity* available for consumption. Another example is cinema in a theatre. For the people inside the hall, no one gets to watch more of the movie than anyone else does. It is, however, possible to exclude people who do not pay for the ticket from watching the movie. So the movie is a good which is non-rival but excludable. Goods that are both non-rivals in consumption and having the non-exclusion property are called *pure* public goods. To qualify as a public good, a good has to have at least non-rival consumption. In the examples given above, although there is non-rival consumption, the goods are excludable.

Club Goods, Congestion and Impure Public Goods

We mentioned above that *pure* public goods are characterised by both non-rival consumption as well as exclusion. There are, however, some goods for which consumption is non-rival but where exclusion can be applied as in the cable television or movie theatre examples above. These goods are sometimes called club goods. This is actually what happens in the case of members of a club, who have joint and hence non-rival consumption, but where nonmembers are excluded. Because of this kind of nature of an actual club, goods which possess the characteristic of non-rival consumption coupled with exclusion possibilities are sometimes called club goods.

There is another class of impure public goods. This class of goods have rival consumption but in their case it is very difficult or extremely costly to carry out exclusion in consumption. A congested road is a prominent example. One person driving a vehicle does not mean others can be excluded from driving their vehicles but clearly, the space taken up by one person on the road reduces the space available for others - hence use of road space is a rival good. A similar example is that of a beach. People cannot exclude others from enjoying a beach, but it is possible that the beach gets crowded and hence the space available for each person on the beach gets reduced. Thus a beach has rival consumption. From these examples we see that cases of congestion are yet another type of impure public good in a sense opposite of club goods - which exhibit rival consumption but no or negligible exclusion.

Before we discuss more about public goods, especially regarding their provision, let us pause to look at the various types of goods. At one extreme are private goods that have both rival consumption and exclusion. At the other extreme are pure public goods that have non-rival consumption as well as non-exclusion. In between the two extremes are impure types of public goods: the first type characterised by non-rival consumption and exclusion (club goods) and the other type which has rival consumption but non exclusion (goods in situations of congestion).

The various possibilities regarding the presence or absence of exclusivity and rivalries are shown, with examples, in the following table.

Rivalry and Exclusion	Non-rivalry and Exclusion
Rivalry and Non-Exclusion	Non-rivalry and Non-Exclusion

Provision of Public Goods

The private sector often does not do a good job at providing public goods. The main reason is that since it is difficult to exclude people from consuming the good, it becomes difficult to price the good. This in turn robs private entrepreneurs of the incentive to produce the good.

Even if some people could be excluded from consuming the good, the non-rivalry in consumption would mean that it is costly and inefficient to exclude anyone. Once the good is produced, the marginal cost of providing the good to an additional person

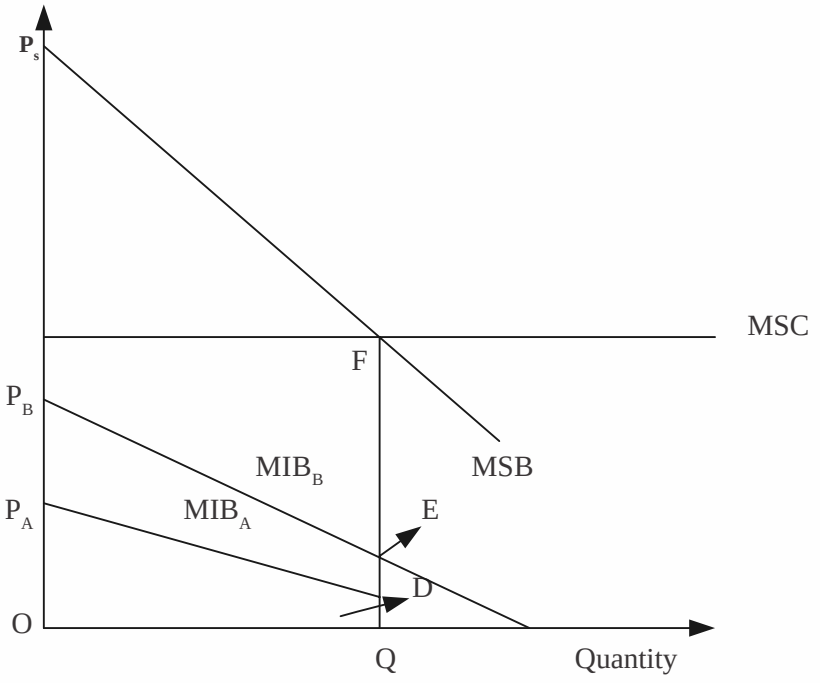
is zero. The basic issue is that due to the presence of non-exclusion in the case of public or social goods, the market mechanism, that is, the principle of demand and supply regulated by the price system may not be an efficient method for the provision of public goods. This is quite unlike the case of private goods, which are characterised by exclusion, so that only the person who pays for a good gets to enjoy it. A related and equally serious problem arises in the case of public goods, specially when that good is to be financed by the consumers or users themselves, like security in a residential colony in an entry. This is the so-called **free rider problem**. In essence this means that since for a public good it is difficult to carry out exclusion, even those who do not pay for the good, or do not put forward their share of finance can avail themselves of or enjoy the good. This provides an incentive for an individual consumer to get a 'free ride' or be ' free-rider' at the expense of others, anticipating, often correctly in the case of public goods, that others will in any case pay for the good. The problem arises when a large number of customers begin to think on these lines. As is evident, in those situations, where projects (public goods) are to be financed by the consumers themselves, there is every possibility that such projects will not get financed. For public goods in general, the free rider problem leads to situations where individuals have incentives as well as the opportunity to enjoy the goods without paying for them. Hence the good may be under-supplied.

We know that a perfectly competitive market will provide the optimal quantity of a private good because production will be expanded to the point where demand equals supply. In perfectly competitive markets, the demand curve represents the marginal social benefit, that is, the full social benefits of additional units (let us denote it by MSB) whereas the supply curve reflects the marginal social cost of production (denote it by MSC).

In the above paragraph we assumed that there were no externalities. Let us now suppose the good is a public good. Even now, output should be expanded to the point where MSB = MSC. But the important point is that with public goods the market demand curve is no longer obtained from individual demand curves by horizontal summing of the individual curves. The reason is that *all* individuals *simultaneously* consume *each* unit of the public good (non-rival consumption).

The following Fig.21.1 depicts the derivation of MSB curve for a public good from individual curves.

Fig. 21.1



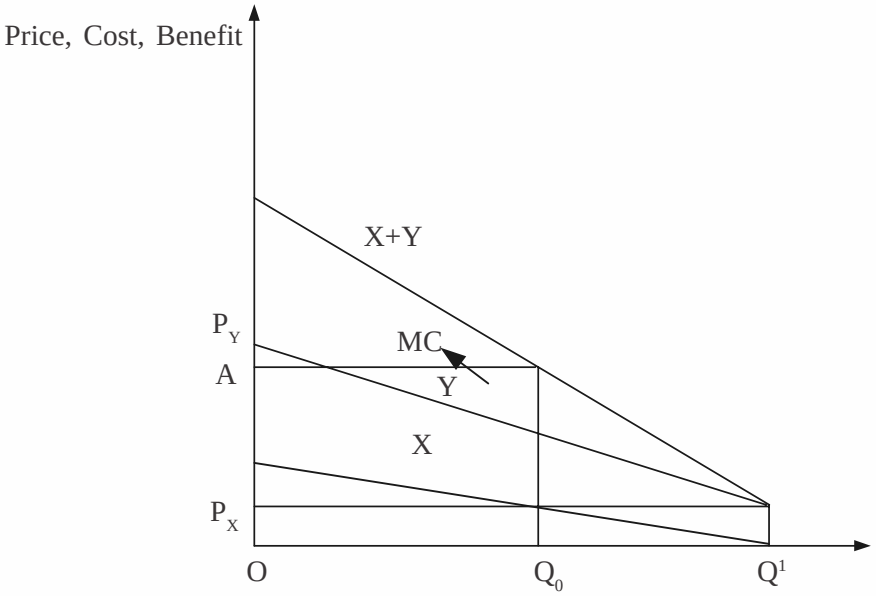
We have considered only two individuals for simplicity. Let the two persons be A and B. MIB^A represents the marginal individual benefit curve for A and likewise MIB^B for B. We also assume a horizontal marginal cost curve for simplicity. At any given quantity the demand price reflects the marginal benefit for the consumer. To get marginal social benefit, we have to add the marginal benefits to all the individual consumers. Thus in this case we add the price for consumer (P_A) with the price for consumer B (P_B) to arrive at social benefit (P_s). For the Qth unit of the good A receives MIB equal to QD and B receives MIB equal to QE. But, since it is a pure public good and hence non-divisible, both A and B consume Q units of the good and both receive the benefits. Thus the MSB curve, shown by the thick line is derived by vertically summing the MIBs of A and B because of non-rivalry. At the point Q, MSB equals QD + QE = QF. The optimal output now turns out to be OQ units of the public good because at the corresponding point on the MSB curve MSB intersects the MSC curve. To the left of Q society gets social benefit is more than costs, that is, MSB is greater than MSC and hence it is beneficial to increase production. Thus Q is the most efficient point because only at this point is the sum of A's marginal benefit and B's marginal benefit equal to the marginal social cost.

Of course, our analysis is quite simplified and we could make it more complex. For instance, in many situations, the MSC curve need not be horizontal but may be upward sloping. Also, the good need not be a *pure* public good so that the consumers need not all consume the same amount of the good. An example of this is education.

In actual practice, obtaining marginal individual benefit curves of individuals is very difficult, especially if the number of individuals is large. Frequently, in many societies, bureaucrats and government officials decide on behalf of society the quantity of public goods to be produced and distributed. The danger is that an underestimation of social costs and/or an overestimation of social benefits can result in overproduction the public good. Consequently, there can be a net social loss.

There is another problem with regard to the supply of public goods that we consider now. This arises out of a uniform pricing of public goods. Suppose there is a non-rival but excludable public good. Suppose the government follows an optimal social policy and provides an amount for which MSB = MSC. Suppose that all users of the public good are to be charged a uniform price and let the price be one at which the users are willing to buy the full amount provided. In such a case, most users will be willing to buy more units than they can, while no one wishes to buy fewer units. This is shown in the following diagram (see Fig. 21. 2).

Fig. 21.2



We have, for simplicity, taken only two individuals, X and Y. Curves X and Y show the marginal value for the two individuals. Since each unit of the good can be consumed by both X and Y, the total social value of a unit of the good is given by the curve X+Y, which is obtained by vertically summing X and Y. Suppose OA is the (constant) marginal cost. Then the MSB should be equated to the MSC (which in this case is equal to the marginal cost). To induce X to consume the quantity OQ_0, the price must be OP_X. But with price $P_{X,}$ individual Y would want to consume the quantity OQ_1, which is larger than OQ_0. Thus with the same price being charged for all consumers, some individuals will want to consume more than is provided. The solution lies in being somehow able to charge OP_X for individual X and charge price OP_Y for individual Y. In such a case, both individuals consume quantity OQ_0 which is the quantity at which MSB = MSC. Also, for this quantity, production costs are totally covered as MC = OA = Average costs (because the MC curve is horizontal) and this is equal to $OP_X + OP_Y$ which is the entire average revenue. Such an equilibrium or situation where MSB =MSC, where each individual is charged a price equal to his marginal individual benefit (MIB) curve and where prices cover production costs is called a *Lindahl equilibrium.* A basic problem relates to the elicitation of actual responses and preference of all individuals. This is the problem of designing a so-called *demand revelation mechanism.* Some economists have suggested ways of designing *incentive-compatible mechanism* whereby even self-interested individuals are induced to reveal their true preferences.

Merit Goods

In all the types of goods— private goods, impure public goods and pure public goods we have encountered till now there has been a common feature that can be found: It is the consumer who decides whether she wants to consume the good or not. Even for a free rider the decisions is hers alone. However, remember few exceptions like military defence- a public good where the decision to 'consume' does not rest with the consumer.

So we see there are goods which seem worthy of consumption and it is decided by an external agency, frequently the government. This type of goods is called merit goods. An example is the case of the government deciding that rider of motorised two-wheeler vehicles must wear helmets, for their own good and safety. Helmets would be private goods, but since it is an external agency, the government which decided that people have to wear helmets, helmets become a merit good as well.

Merit goods are goods, which are consumed on the social interest. These are reflected in community wants, and individual wants are made subservient to the common good in such cases. The assumption made tacitly here is that long association among members of a society leads to the development of some common goals, interests and values of that society. In modern democratic societies, the State is supposed to represent social preferences. There is also a sense of paternalism inherent in the concept of merit goods. Merit goods are not only those which are provided by the State. If certain individuals or groups in society act as donors or display altruism but at the same time determine the form that the charitable item is to be provided, say, in kind rather than cash, it becomes an example of merit good.

21.4 EXTERNALITIES

'Externalities' means 'of or from outside'. An externality is a situation where a consumer or producer is affected, either positively or negatively, by the consumption or production by another agent. Simply put, externalities are situations where the behaviour of some economic agent(s) affects the welfare of others.

Externalities are side effects. Examples of externalities are abound. Immunisation against a contagious infection confers the benefits directly to the person getting inoculated. But it also indirectly benefits others because the chances of their contracting the disease from the affected person are reduced when he gets inoculated. This indirect benefit is an externality.

There are externalities in the production side also. Suppose firm A spends a lot of money in imparting skills and high level techniques to its workers. If subsequently many of these workers join firm B, then firm B (which is a consumer of the services of these workers) reaps the benefits of the services of trained and skilled workers. Another example, which is frequently cited, goes as follows: There is a beekeeper, and next to his plot is the plot of an owner of an orchard. But the nectar from the flowers helps the bees to produce honey. This is an external benefit for the beekeeper.

There are external costs as well. Pollution is an example. If a firm's smoke increases the risk to the people living around the factory of catching respiratory diseases, or if its effluents ejected in a river flowing nearby poison a lot of fish and again increases the danger to the people eating the fish, it is an external cost imposed on the people. Sometimes psychic effects are put in the category of externalities. Suppose you maintain a beautiful garden and I as your neighbour derive pleasure from looking at the flowers, I get external benefits. If I play music very loudly at night, and it disturbs you, then it is an external cost for you.

What is the relation between public goods and externalities? At one level, the two are very similar, specially with regard to the characteristic of non-rival consumption. The inoculation example given above for an externality, can also be seen as non-rival consumption in the sense that the inoculation of a person does not decrease the amount of 'reduction of risk of disease' that is 'consumed' by other people. Thus, public goods and external benefits seem to be related. The basic relation is that all public goods are characterised by externalities but all examples of externalities are not at the same time examples of public goods. Public goods, as we saw above, mean that there are externalities of consumption. Hence externalities are a necessary condition for a good being a public good. Being a public good is not a necessary condition for a situation of externalities.

Ways of Coping with Externalities

There are two broad ways of dealing with the problems posed by externalities. One way consists of methods which do not require abandoning the working of the market and can be accommodated by the normal functioning of the market. The other way is using "traditional" solutions and consists of taxation and internalisation of costs. Let us take up the traditional methods first. These include taxes and subsidies. The British economist A.C. Pigou was a great proponent of this view.

To understand Pigou's solutions for externalities, let us study how he understood the phenomenon. First, let us talk of the various kinds of externalities: (1) *positive externalities in consumption*. An example is vaccination; (2) *negative externalities in consumption*. A noisy motorcycle disturbing neighbours is an example; (3) *negative externalities in production*. An example of this is a paper mill that dumps its waste into a river. The waste adversely affects the riverside residents and fishers; (4) *positive externalities in production.* The case of the beekeeper and the farmer cited above is an example.

Now let us come to Pigou's argument. His basic point is that in the presence of externalities, *even if perfect competition prevails*, a Pareto optimum is not attained. To see why, we must distinguish between private and social costs on the one hand and private and social benefits on the other. In the presence of externalities, there

is a divergence between social and private costs on the one hand and private and social benefits on the other. With externalities present, social benefit or cost is a combination of private and *external* benefit or cost.

We denote the relationships in the following way:

Let

MPC = marginal private cost

MEC = marginal external cost

MSC = marginal social cost

Then MSC = MPC + MEC

Also let

MPB = marginal private benefit

MEB = marginal external benefit

MSB = marginal social benefit

Then MSB = MPB + MEB

For overall efficiency, MSC should be equal to MSB for each product. As long as MSB>MSC production should be expanded and vice versa.

Now let us consider a method of dealing with externalities, which does not require abandoning the working of the market. This method was suggested by Ronald Coase. The essence of Coase's method is that private individuals can solve the problem of externalities through voluntary bargaining and the government is not needed to deal with externalities. However, for this the required conditions are perfect competition and absence of transaction costs. The Pigovian approach is to tax the agent who is creating the external costs. The person who has the cost imposed on him or her would be provided a subsidy.

By contrast, Coase suggests that the person who is imposing the costs can compensate (monetarily) the person who is adversely affected. Coase even suggests that in some cases the person who is likely to be adversely affected could pay the person who is about to create negative externalities for him by way of incentive to not to undertake the activity that would create such externalities. Thus the basic point Coase makes is that voluntary bargaining can lead to efficient allocation even in the presence of externalities. Coase suggests not only that voluntary bargaining leads to efficient outcomes but that there is a close relationship between external effects and property rights. Coase's proposition that with perfect competition and absence of trasnsction costs, voluntary bargaining can solve the problem of externalities is sometimes called the Coase Theorem For his ingenious theory, Coase got a Nobel Prize in Economics.

Check Your Progress 2

1) Distinguish between non-rivalry and non-exclusion in the consumption of goods.

.....

.....

.....

2) What is an impure public good?

.....

.....

.....

3) What is a merit good? Is every merit good a public good?

.....

.....

.....

4) What are externalities? Briefly compare the Pigovian method with the Coase method for dealing with externalities.

.....

.....

.....

21.5 PUBLIC EXPENDITURE

The Government has to undertake certain expenditures both for its own maintenance and the maintenance of society. All these expenditures are together called public expenditure. In almost all societies, public expenditure has shown a tendency to rise over time. While many economists have advocated a policy of laissez faire, in reality governments have had to incur certain necessary expenditures. In addition, over time ideas of social welfare have gained currency and almost all governments undertake social welfare measures. There are also necessary administrative costs and expenditure on defence.

21.5.1 Theories about the Rise in Public Expenditure

There are two broad theories about public expenditure. The first is by the German economist Adolph Wagner (1875-1917). He did a study of historical facts about the German economy and propounded what is called, ***The Law of Increasing State Activities.*** He suggested that activities of various levels of government have an inherent tendency to increase over time. The government sector in the economy rises faster than the economy as a whole. There is consequently a rise in government expenditure. Now, this rise in government expenditure can be expressed in many ways: (a) a rise in absolute levels of government expenditure, (b) a rise in the ratio of government expenditure to GNP, (c) a rise in the proportion of the public sector in the economy. Even in the case of (a), the absolute rise may be in nominal or real terms. One should also adjust for a secular increase in population and see the rise in per capita terms. It is not clear in which of the above senses Wagner was talking about the rise in government expenditure, though Musgrave suggests that the correct measure should be (c). Also, for (b) above one should as well look at the GNP *elasticity* of government expenditure.

Wagner's law is mainly applicable to modern progressive governments. According to Wagner, it is applicable mainly in the initial stages of modern government activities. He felt that as modern industrial society develops, there would be increasing pressure for social progress and there would be attempts to make business and industry more socially conscious. The public sector and government activities would therefore rise.

The second main theory about rise in government expenditure is by Jack Wiseman and Alan Peacock and is called the Wiseman-Peacock hypothesis. They studied

public expenditure in Britain for the period 1890-1955 and on this basis suggested that public expenditure does not increase in a smooth and continuous manner but in discrete jumps or in a step like manner. This is mainly because unexpected social disturbances and events take place and government expenditure has to rise to meet the requirements. Of course, they suggested that the existing revenue is in most cases not adequate to meet the expenditure requirements, and revenues, particularly taxes, rise to a new level. This hypothesis is about occurrence of unusual and abnormal events, but it is largely true that government expenditures rise over time in almost all modern societies. Buchanan and Tullock based on U.S. experience, have argued that there is an increasing discrepancy between government expenditure and government output, with the former tending to run ahead of the latter. They give two reasons for this. First, unlike the private sector, the expenditure on government officials increases faster than the corresponding rise in their output. Secondly, with the growth of welfare activities and social security, the proportion of people receiving transfer payments from the government increases.

What are the main reasons for the secular rise in government activities and expenditures over time? First, the traditional functions of the State were expanding. Defence was receiving greater emphasis and expenditure on it was increasing. Wages of government officials was going up. Second, state activities around welfare measures were increasing in their coverage. Third, investment activities of the State have been expanding. Fourth, population itself has been going up necessitating a hire level of committed expenditure on the part of the State. Fifth, related to the previous point on population is increasing urbanisation, which requires a much larger per capita expenditure on civic amenities. Sixth, modern governments need to borrow and thus public expenditure in the form of repayment of loans and increasing costs of debt servicing go up. Finally, increasing use of planning and consequently capital accumulation by the government tends to increase public expenditure.

21.5.2 Kinds of Public Expenditure

Not only are there several types of government expenditure but there are also a number of ways in which public expenditure can be classified. But before discussing the different types of public expenditure, let us bring out the essential difference between private and public expenditure. Are there any similarity between the two? The basic similarity is that *given the objectives* both would like to see a good return on the expenditure. This is particularly true for investment expenditure. Both private units as well as government would like to maximise some objective function.

The primary difference stems from the different objective functions that private and public units have. Private units and the government raise resources for expenditure. Following contrasting methods. While the government decides the amount of expenditure to be made and proceeds to raise the resources thereafter, private individuals keep the income at their disposal as a constraint before deciding the amount of expenditure. Public expenditure also has a greater degree of flexibility. Moreover, the state has a much longer time horizon to plan the expenditure.

Let us now discuss the various types of public expenditure. There have even been different ways of classifying public expenditure. The traditional way, existing for several centuries, is to use an accounting classification. This has been useful for the state in keeping track of expenditure and it affords some control and checks over public expenditure. It provided information about leakage, misappropriation and wastage of resources. This way of classifying public expenditure way was useful for auditing purposes and to control misappropriation but it was not useful in providing information about the effects of expenditure. Therefore, for policy making, a better way to classify expenditure was sought. An economic basis of classification was brought in, which could provide better information about the economic effects of expenditure.

There are many ways to classify expenditure on an economic basis. Two of the most useful are the classification into productive and unproductive expenditure and transfer and non-transfer expenditure. Let us first deal with the former. The basic sense of 'productive' here is investment. Broadly, investment expenditure is, according to this classification scheme, considered productive because it is seen to raise the economy's productive capacity, while consumption is considered unproductive. This view is expressed strongly under the laissez-faire philosophy. In fact, this distinction used to be strictly made during Adam Smith's time. In this view, expenditure on defence, administration, law and order were considered unproductive. The government sector was considered alien to the rest of the economy. Today we need not strictly adhere to this kind of view and hence, classification of government expenditure, for several reasons. First, the government is an integral part of the economy and has to undertake many kinds of expenditure that would have the effects on the rest of the economy. Second, there are many assets, which are not directly productive in the sense of yielding returns but are necessary for economic development. Expenditure on social items is of this kind. Some of these expenditures, such as public works, can yield future returns. Moreover, they increase the national product. Third, assets need not only be in tangible form to be called productive. Investment in human capital, for instance, can significantly raise the productive capacity of the economy. These expenditures also yield utility directly. Fourth, the economy and society cannot sustain themselves without certain necessary expenditures. National defence is one such item. There are also certain expenditures, which indirectly raise the productivity of the economy such as expenditure on research and development.

Another way of classifying government expenditures is as transfer and non-transfer expenditure. Pigou favoured this approach to classifying expenditure. Any expenditure without a corresponding transfer of real resources is called a transfer payment. Examples include gifts, pension payments etc.

21.6 PUBLIC REVENUE

Any government, to carry out activities for economic development, to run the administration, and to perform other general tasks, needs to raise funds for financing these various activities. All the funds that the government raises are together called public revenue. There are several types of revenue, important among which are taxes, market borrowings, income from currency, income from public undertakings, sale of assets, fees, fines, gifts and donations and so on. Some economists, like Hugh Dalton, have sought to make a distinction between public receipts and public revenue-with the latter being a part of and narrower concept than the former. It also does not include market borrowings, income from sale of public assets and income from printing money. There is no harm, however, in simply considering public revenue as a single concept here.

Classification of Public Revenue Sources

Now let us discuss the revenue side of the government budget and look at the various receipts of the government. Just as there are revenue expenditure and capital expenditure, there are revenue and capital receipts as well.

Let us first discuss receipts which are revenue receipts. In Indian budgets, they are placed under the revenue account. Revenue receipts are classified into tax and non-tax revenue. We will presently discuss the difference between tax and non-tax revenue. For the moment, let us just say that non-tax receipts include currency, coinage and mint; interest receipts, dividends and profits; and other non-tax revenue from various government services like administrative services, public service commissions, and jails and prisons.

Now we move on to capital receipts. The basic difference between revenue receipts and capital receipts is that the former is of a short-term duration (less than a year, usually) while the latter are receipts from activities of a long-term nature. The principal type of capital receipts are market borrowings - loans that have a maturity of 12 months or longer at the time of issue. The second category of capital receipts is external loans. The next category of capital receipts is recoveries of loans made by the government. In India, for the budget of the central government, it consists of loans made by the central government to state governments, union territories, and non-government parties. Provident Funds are another important component of capital receipts. Other capital receipts include the net effect of transactions occurring under a variety of accounts and deposits.

Now we go back to revenue receipts and the difference between tax and non-tax revenue. The main characteristic feature of a tax is that it is a compulsory levy on those who are to pay it irrespective of whether they receive any corresponding return of goods and services from the government. In other words, those who pay taxes do not receive definite and direct *quid pro quo* from the government. Thus a tax is *not* a price paid by the tax-payer and no tax-payer can claim any direct benefit from the government on the ground that he or she is paying a tax. The benefit may go to anyone irrespective of who pays the tax. A tax is a liability imposed upon the tax assessees who may be individuals, groups, or other legal entities.

Now let is discuss some concepts relating to a tax. First, the **base of a tax** is the **legal** description of the object with references to which the tax is levied. For example, the base of an income tax is the income of the assessee as defined and estimated. The base of an excise duty is the production or processing of a specific good.

The **incidence of a tax** related to the entity (person or group or other legal entity) that has to bear the final burden of a tax. The assessee on whom the tax is levied, and who is to pay the tax can sometimes shift it further to someone else. For example, sales tax is levied on the seller of a commodity and he pays it, but the burden of payment is passed on to the consumers by the seller in terms of a higher price. So when you buy a good on which there is a 7 per cent sales tax, you pay a price that includes this tax that had actually been levied by the government on the seller (all sellers of this commodity). On the other hand, income tax is a tax, the burden of which cannot be shifted. If you are to pay income tax, you cannot shift it someone else. The initial entity on which the tax is levied is called the **impact of a tax** while the final burden of a tax is called the **incidence of a tax.**

Those taxes for which the burden of tax cannot be shifted by the assessee are called direct taxes, while those whose burden can be shifted are called indirect taxes. Thus income tax is a direct tax while sales tax and excise duties are indirect taxes. Thus for direct taxes, the impact and incidence of the tax is the same, while for indirect taxes, the two are different.

Check Your Progress 3

1) Discuss the Wiseman-Peacock hypothesis about the rise in government expenditure over time.

.....

.....

.....

.....

2) Distinguish between (a) public and private expenditure, (b) productive and unproductive public expenditure, (c) transfer and non-transfer public expenditure.

.....

.....

.....

3) Distinguish between revenue and capital receipts and give two examples of each.

.....

.....

.....

21.7 MEASURES OF BUDGET DEFICIT

Very generally speaking, a budgetary deficit is simply the excess of public expenditure over public revenue. In practice, however, there are a variety of measures of the budget deficit. There is the added problem that a single measure or a concept may be called by different names by different experts, or a particular name may signify different concepts to different experts. In this section we discuss both the measures used in the theoretical public finance literature as well as the various measures used by the Government of India (GOI) in its annual budget documents.

We must, at the outset, distinguish between the concepts of deficit and debt. The deficit is the excess of government spending over government revenues in a particular year. The government debt, on the other hand, is the sum of all previous outstanding government debt obligations held by the public. So the government deficit is a flow concept which is measured over a *period* of time. In contrast, the government debt is a stock concept, which is measured at a *point* of time. Another general point that needs to be mentioned in this connection is the distinction between real deficit and nominal deficit. The former measures the deficit adjusting for inflation and is measured in terms of some base level prices whereas the latter is ascertained in terms of the current price level.

We now discuss the various measures and concepts of budget deficit as used in the literature and also by the GOI. To proceed, let us write down the items of receipts and expenditures as they appear in the annual budgets of the GOI.

We can begin by looking at the expenditures. Typically, expenditures are classified into those in revenue account and those in capital account. Basically the difference between the two accounts is that the former relates to expenditure incurred in the current financial year while the latter deals with expenditure incurred over longer time periods. There are two items in expenditure on revenue account: interest payments by the government and non-interest expenditure. Expenditure on capital account similarly has two items: loans and advances, and capital outlay.

Now let us discuss the revenue side of the government budget and look at the various receipts of the government. Just as there are revenue expenditure and capital expenditure, there are revenue and capital receipts as well. Revenue receipts consist of two items: tax revenue and non-tax revenue. Non-tax revenue, in turn, has three items-interest, non-interest revenue and grants. Capital receipts are of four kinds: recoveries, borrowings, other than through Treasury Bills, other capital receipts and sale of public assets.

The balance of total expenditure and total revenue is budgetary deficits, which is met by borrowings through Treasury Bills and drawing down cash balances.

Now let us consider revenue and expenditure items in detail.

We can first depict the above mentioned items succinctly as follows:

I) Revenue Receipts

a) Tax Revenue (net)

b) Non-tax Revenue
   i) Interest
   ii) Non-Interest
   iii) Grants

II) Capital Receipts

a) Recoveries

b) Borrowings, other than through Treasury Bills

c) Other Capital Receipts

d) Sale of Public Assets

III) Total Receipts (I + II)

IV) Expenditure on Revenue Accounts

a) Interest Payments

b) Non-interest Payments

V) Expenditure on Capital Account

a) Loans and Advances

b) Capital Outlay

VI) Total Expenditure

VII)Borrowings through Treasury Bills and Drawing Down of Cash Balances

Several measures of deficits can be constructed from the above items of revenue and expenditure. Let us now consider these.

1) **Deficit on Revenue Account or Revenue Deficit:** The excess of expenditure on revenue account over receipts on revenue account, or item IV - item I.

2) **Deficit on Capital Account:** The excess of capital disbursements over capital receipts, or item V - item II.

3) **Budgetary deficit :** 1 + 2

We now discuss an important aspect related to budget deficit or budgets in general. This is to do with the questions: should the government aim for a balanced budget and avoid deficits altogether? Is it deleterious to the state of the economy? Why is a reduction or even removal of deficits advocated by many agencies and experts?

Classical writers till about the first three decades of the twentieth century were traditionally advocating the balancing of government budget over the relevant period under consideration. In other words, the total revenue of the budget was sought to

be balanced with the total expenditure of the government. Of course, the important question is what should be the relevant considered period? The relevant period was usually taken to be a year. Hence, on the revenue side of the budget, borrowings (short term as well as long term) were not included. Moreover, the budget was considered in an accounting sense. But even in this practice, some modifications have sometimes been suggested. Some have suggested that repayment of loans should not be counted towards the current year's deficit, although interest payments should be considered.

What were and are the main arguments put forward by the advocates of balanced budgets? First, traditionally, the government budget was considered just like the budget of a private unit. Just as it is thought undesirable for a private unit, particularly a household to incur deficits, it was considered unwise on the part of the government to run deficits in its budget. These days, of course, with consumer credit and credit cards, it is not always insisted upon that private individual budgets balance. The argument that government budgets are just like private budgets stems from viewing government as external to the productive economy. The governments, too, had sometimes a tendency to be profligate.

The second argument put forward against budget deficits are that financing of the deficits means that currency and money supply in the economy increases and this puts pressure on prices and leads to inflation.

The third argument is ingenious. It is argued that people do not like increased taxes imposed on them. On the other hand, government expenditures have a tendency to increase, as we have seen in the section on expenditure. So, the government finds it easier to finance the budgets through public borrowing, because the people are happy to lend to the government, especially if the rate of interest on the loans is high. The very fact that the government finds it easy to finance deficits through public borrowing or increasing the money supply, makes the government reckless about expenditures, and spending increases even more. This increases deficits. Inflation resulting from deficit financing also has a tendency to further deficits. Thus, deficits have a tendency to feed on themselves and spiral. Therefore, proponents of this view feel that deficits should be firmly curbed as soon as they arise.

We must keep in mind that those who propose that deficits should be kept at very low levels also argue that government should keep tax rates low. Their argument is not only that lower rates of taxation would lead to higher tax compliance but is also that high rates of taxation would diminish the incentive of economic agents in the private sector to undertake profitable ventures. Consequently, output and productivity would suffer. So, if lower rates of taxation are to go hand in hand with low deficits it only means that governments should keep expenditure levels very low. Moreover, the proponents of this view also suggest that governments should not enter into areas where, there may be potential profits for government, such as hotels and tourism. Hence in summary we can say, the proponents of the view that deficits should be low tend to be conservative and believers in laissez faire philosophy.

What can be the arguments in favour of budget deficits? Remember that Keynes was an advocate of deficit financing for curing economic depressions. Moreover, during the post World War II period of growth in the American economy, that economy had, in several years in the 1950s and 1960s, deficits in government budgets. Some of the other arguments in favour of budget deficits, or rather, for tolerating deficits and not developing a phobia of them, are the following:

First, balancing of budgets should not be an end in itself. Both in aims and structure public budgets are different from private ones. Budgets should be used to help the economy. Budgets should not be neutral in their effect on the economy, unlike what the proponents of balanced budgets suggest. Secondly deficits do not always lead

to inflationary pressures, specially if the deficit is small in relation to national income and is not persistent. Also, deficits can sometimes help the economy to recover from a recession.

21.8 MARKET BORROWING

In the previous section, you got some idea of the various measures of deficit. We saw that a deficit, howsoever defined, is an excess of spending over government revenue. In this section, we look at a related concept: market borrowing.

Public Debt

The overall debt and obligation of the government, measured at a point of time, is the public debt. The public debt has been defined in various ways depending on the items that are thought appropriate to be included in the definition. To get an idea of the public debt, let us look at the various obligations of the government. First, the government creates currency. Often, a part of the currency may be issued by the central bank, but usually the central bank in most countries is part of the government so that the total currency issued and obligation may be considered a government liability to the rest of the economy.

The second set of obligations is the short-term debt, normally with a maturity of less than a year at the time of issue and consists of items such as Treasury Bills and short-term loans from the central bank. There are some debts that do not have any specific date of maturity and are called floating, and part of these may be paid of at various times and are subject to various terms and conditions. These include provident funds, small savings, reserve funds etc. In India, the government has issued certain special securities to meet its obligations towards international institutions like the World Bank and the International Monetary Fund (IMF). These special securities are sometimes called special floating debt.

The importance of market borrowings lies in the fact that in some cases, such as Indian public finance, market borrowings are excluded in the estimation of budgetary deficits. Market borrowings are long term borrowings, where the maturity period is over a year. The reason given for excluding market borrowings from budgetary deficits is that it is felt that since these are long term obligations, they merely divert investible funds from the private sector to the government and hence do not raise the purchasing power and the quantum of currency. Consequently, inflationary pressures in the economy do not build up by the market borrowings. This view need not be correct, as the RBI itself takes up a large portion of market borrowings. The effect of both short term and long term loans taken up by the RBI is the same, in increasing the amount of currency.

Check Your Progress 4

1) Distinguish between public deficit and public debt.

.....

.....

.....

2) Explain the effect of deficit financing on the capital market

.....

.....

.....

21.9 LET US SUM UP

This unit has acquainted you with the basic elements of public economics including several basic concepts from traditional public finance. Our point of departure was a description of market failure. We discussed the various types of market failure. Next we explained the concept of Pareto improvement and Pareto optimality. In this connection we mentioned the link between perfect competition and Pareto optimality. Here we also talked about the two fundamental theorems of welfare economics.

We then took up the explanation of the concept of a public good. We saw that non-rivalry and exclusivity characterize a public good. A good like this is actually a pure public good. We took up a brief discussion of impure public goods, two types of which are club goods and goods in the presence of congestion. We also considered some issues in the provision of public goods and saw the central difference in obtaining a social benefit function from individual functions. It was pointed out that the individual curves are added vertically instead of horizontally, as happens in the case of a private good. We saw, too, what Lindahl equilibrium means.

We next went on to a discussion of externalities. We saw that an externality is an example of a market failure. After considering the four basic types of externalities, we discussed ways of coping with externalities. Here we discussed Pigou's suggestion of correction using suitable taxes and subsidies. The other way, as we saw, is granting property rights and letting private parties arrive at mutually beneficial deals. We explained very important Coase as well.

These broad topics of market failure, public goods and externalities constitute elements of what may be called an extension of some the previous units that you studied in this course. The next three broad topics- public revenue, deficits and market borrowing are central themes in traditional public finance. We began our discussion of public revenue by pointing out a distinction that some have made, between revenue and receipts after broadly classifying revenue into tax and non-tax. We then turned to a discussion of the engaging an important topic of deficits and their financing, briefly touching upon the distinction between debt and deficit. We discussed and compared various measures of the government deficit.

21.10 KEY WORDS

Balanced Budget : A government budget in which the total revenues exactly match the total expenditure.

Capital Expenditure : Expenditure incurred on items that have a long term basis.

Capital Receipts : Receipts accruing from items of a long term character.

GNP Elasticity of Public Expenditure : The ratio of the percentage rise in government expenditure to a percentage change in GNP.

Perfectly Competitive Market : A Market situation where no buyer or seller is able to influence prices, where the goods sold by all sellers are homogeneous, and where there is full and perfect information and free entry and exit to and from the market.

Pure Public Good : A good characterised by both non-rivalry and non-exclusion in consumption.

Revenue Expenditure : Expenditure incurred on short-term activities

Revenue Receipts : Receipts incurred on activities, usually for a period more than a year.

Short-term Debt Instrument : Those instruments with maturity of less than one year.

21.11 SOME USEFUL BOOKS

Bhatia, H.L, *Public Finance* (Latest edition) Vikas: New Delhi

Browning, Edgar K. and Jacquelene Browning (1994), *Public Finance and the Price System* (Fourth Edition) Prentice Hall: Englewood Cliffs New Jersey

Musgrave, Richard A. and Peggy B. Musgrave (1989), *Public Finance in Theory And Practice* (Fifth Edition) McGraw-Hill International Edition: New York

Stiglitz, Joseph E. (1994), *Public Sector Economics*, Third Edition, W.W. Norton & Co.: New York

21.12 HINTS/ANSWERS TO CHECK YOUR PROGRESS EXCERSISES

Check Your Progress 1

1) Read Section 21.2 and answer.

2) Read Section 21.2 and answer.

3) Read Section 21.2 and answer.

Check Your Progress 2

1) Read Section 21.3 and answer.

2) Read Section 21.3 and answer.

3) Read Section 21.3and answer.

4) Read Section 21.4 and answer.

Check Your Progress 3

1) Read Section 21.5 and answer.

2) Read Section 21.5and answer.

3) Read Section 21.6 and answer.

Check Your Progress 4

1) Read Section 21.7 and answer.

2) Read Section 21.8 and answer.